# DOXASTIC SPACES

## A NEW APPROACH TO RELATIONAL BELIEFS AND UNSTABLE NEGLECT

**Dane Rook**

*Jesus College, University of Oxford*

Thesis submitted for the degree of:

*D. Phil. in Geography & the Environment*

# ABSTRACT

This thesis introduces and explores a novel construct for studying human beliefs in social science: doxastic spaces. These flexible analytical devices are demonstrated as capturing three key properties of beliefs which are difficult to depict through other formats: the relational, relative, and reflexive properties of beliefs. The doxastic-space paradigm developed by this thesis is likewise shown to enable new and insightful theories about belief formation and change. Two such theories cultivated herein are quantized evidence theory (QET) and entropy-based social learning (EBSL). These theories prioritise not only the evidential bases of beliefs, but also the cognitive limitations on memory and attention that people face in constructing and updating beliefs about their worlds. Such bases and limitations underscore not only the role that context has to play in sculpting beliefs, but also the reciprocal function of beliefs in helping to determine and demarcate context. Part of that context is discussed as being other people relevant in social judgment and learning situations. And interplay between beliefs and context is used to aid explanation for unstable tendencies in neglectful cognition. The work mixes theoretical and empirical investigation of the doxastic-space framework, and suggests that it may serve social science by working to not only forge deeper comprehension of belief dynamics but also to operate as a platform for interdisciplinary exchange.

Katherine Rook taught me that true belief transcends hope; this work is dedicated to her.

# ACKNOWLEDGMENTS

There are many people who deserve my gratitude in making this project both possible and pleasant to undertake. Unfortunately, space here is lacking to recognise everyone who is due thanks for their help and encouragement, and therefore my listing of appreciation must be selective. I have to acknowledge Gordon Clark for his steadfast commitment to helping me; his support enabled me to come to Oxford in the first place, and his guidance (and prodding) has helped to positively shape my experience over these last few years. For advice, insights, and inspired conversation during my time at Oxford, I wish to particularly thank Maurizio Fiaschetti, Heather Hachigian, John Krebs, Tim Schwanen, and Darek Wójcik. My friend of over a decade, Charles McClure, has also given sound counsel on several ideas in this thesis. I also would like to thank John Whiteman for some (characteristic for him) brilliant jolts to my (world)views and for his pragmatic criticisms. And through it all, Amanda Diener has been a source of support and good cheer. Finally, I doubt I would have found the resolve to complete this effort had it not been for the love, firm dedication, and understanding shown to me by friends and family, both overseas and here at Oxford. Among them, Bethan Nichols has been the greatest ear for my ideas and root of my confidence. Thanks also go to Oxford University Press for funding my doctoral research with a Clarendon Scholarship.

# CONTENTS

**CHAPTER 1**: INTRODUCTION……...…p. 5

**CHAPTER 2**: LITERATURE REIVEW……..…….....p.28

**CHAPTER 3**: MINDING MAPS IN MAPPING (NEGLECTFUL) MINDS:
MULTI-SCALE REFLEXIVITY IN NEURAL SPACE……..…p.70

**CHAPTER 4**: COGNITIVE LIMITATIONS ON BELIEF APPROXIMATIONS:
A THEORY OF EVIDENTIAL QUANTIZATION……...……....…….…p.107

**CHAPTER 5**: ENTROPY-BASED SOCIAL LEARNING……...….p.152

**CHAPTER 6**: DISAGREEING ON AGREEING…...……p.238

**CHAPTER 7**: CONCLUSION AND EXTENSIONS…….……...….p.296

**REFERENCES**……...….p.303

CHAPTER 1

# INTRODUCTION

## **1.1**\\ On the Proper Study of Beliefs: Relationality, Reflexivity, and Instability

How should human beliefs be studied? A reasonable answer would seemingly require clarifying (if only temporarily) two ambiguities nested inside this normative question. Firstly, any judgment on the suitability of an approach to studying beliefs should take into account the particular definition of beliefs used. And there are an astonishing number and variety of such definitions, which range from those highly formal and explicit, to those very informal and vague. Both the age of the concept of belief, as well as diversity of definitions for it, suggest that no single conceptualisation of human belief has yet emerged which is universally fit for all purposes. Secondly, determinations about the proper study of any beliefs seem inherently steered by purpose: the matter of *who* is studying such beliefs and *why* they are doing so (or may wish to do so) appears practically inseparable from normative concerns regarding how beliefs should rightly be studied.

The foregoing considerations should serve to introduce what is one general (and generally fundamental) property of human beliefs: their relational character. Under many interpretations of belief, human beliefs serve as bridges connecting the believer with the believed: beliefs are, in most cases, artefacts of the operation of believing (or the potential for such an operation), which requires both a subject, the believer, and some (often focal) object: the believed. The relational character of human beliefs therefore not only determines the suitability of any candidate approaches for studying beliefs, but is also a fundamental part of the function of beliefs for people.

Yet, as will be treated extensively in what follows, human beliefs are also relational for connecting believers with subsets of their environments in which beliefs are formed and may be said to exist. In short, many forms of belief are context-dependent due to contextual-sensitivities of the people who formulate and use them. Nonetheless, this dependency is not unidirectional because what comes to qualify as ‘context’ may itself be determined (partly or wholly) by a persons’ beliefs.

Hence, many human beliefs are not only relational, but also reflexive; they emerge from complex and dynamical interdependencies between a believer, the believed, and circumstances of believing.

Moreover, it is this simultaneously relational and reflexive character of human beliefs that appears to have most sincerely challenged predominant conceptions about beliefs in social science over recent decades. In psychological and economic research especially, a substantial amount of evidence has accrued over the past half-century which demonstrates that human beliefs exhibit some curious forms of instability.[1] Such instabilities have frustrated many researchers in those and other disciplines because they play no conventional role within the two hypotheses on judgment and choice which (arguably) remain the primary cornerstones (or at least leading benchmarks) for models of human behaviour within social science: expected utility; and subjective expected utility.[2] Early efforts at reconciling observed belief instabilities with these expected utility hypotheses largely focused on crafting models and theories wherein human decision-makers could wield unconventional forms of preferences, yet otherwise possess and make use of stable forms of beliefs.[3] Such efforts continue today and have generated some remarkable insights into the operation and limits of human cognition. Nevertheless, some uncomfortable incongruities and unresolved problems persist with such preference-centric remedies that stubbornly impose strict

[1] These instabilities take various forms, but primarily concern: 1) both observed under- and overconfidence in beliefs; 2) peculiar asymmetries in the treatment of the balance and weight of evidence (see, e.g.: Griffin and Tversky (1992)); 3) both subadditive and superadditive treatments of probabilities; 4) belief reversals (Fox and Levav 2000); and general instability concerning the nature of beliefs on events which have extremely-small probabilities of occurrence. These belief instabilities will be discussed more extensively in the next chapter, as well as in the others which follow.

[2] Although the former and latter hypotheses are typically attributed to von Neumann and Morgenstern (1944; 1947) and Savage (1954), respectively, there are so many other contributors to these ideas, both modern and ancient, that such singular attributions appear downright unfair. Indeed, to better understand the durability of those expected utility explanations of behaviour, one must appreciate the prevailing zeitgeists and mind-sets that facilitated the development of those ideas in the first place. A brief tour of the backdrop for development of those ideas appears in the next chapter. It should also be noted that whereas subjective expected utility clearly deals with beliefs (Savage 1954), expected utility generally may be understood to deal with what might be better described as 'knowledge', as it concerns exogenously-given probabilities. While there are many views that take knowledge as true belief (e.g.: Ramsey (1929)), there is also substantial thinking and evidence suggesting people do not always take probabilities entirely as 'given' when forming judgments (see: Kahneman and Tversky (1979); Viscusi (1989)) which may preserve beliefs' relevance.

[3] See, e.g.: the recursive preference framework (Epstein and Zin 1989).

normative requirements upon decision-makers' beliefs, but allow them to harbour increasingly complicated preferences. These difficulties will be enumerated and dissected in detail in subsequent chapters. For now it suffices to recognise that critical reconsideration of perspective might be widely beneficial for researching decision-making, judgment, and choice in the social sciences: rather than accepting stable beliefs as typical in human cognition, and seeking to explain empirical instabilities as exceptions to that norm, it might prove fruitful instead to recast beliefs as characteristically unstable, and aim to explicate observable stabilities in their formulation and use.

## **1.2**\\ What Place for Randomness in Social Science?

And such a proposed reframing would not stand altogether exceptional in modern social science, for instabilities are invoked regularly as explanations under the guise of (typically parameterised) randomness. Particularly, injecting stochastic elements into theories as error terms in models has become a staid and accepted practice in many branches of economics and psychology. Nor is such a practice essentially wrong-minded: models are necessarily compressions of reality (if it were otherwise then they would fail to be models). Thus, errors arising from models' predictions are often inescapable. Indeed, in many strands of social science assessments of a model's fitness is often dictated by relative acceptability of the behaviour of errors which it generates. Nevertheless, some models and theories of human judgment and choice are placing evermore reliance on randomness as an explanatory driver, rather than as a stopgap between prediction and observation.[4]

[4] Within psychology, ideas on cognitive sampling present perhaps the clearest illustrations of granting primacy to semi-structured randomness as an explanatory engine behind human behaviour. Although some researchers pursuing ideas in that vein acknowledge that assumed randomness in the structuring of beliefs from evidence is merely a placeholder to be turned deterministic in future work (see, e.g., Stewart, Chater, and Brown (2006)), others (discussed in the next chapter) appear durably-committed to randomness as central to cognition; see, e.g.: Erev, Wallsten, and Budescu (1994); Wallsten and Gonzalez-Vallejo (1994); Budescu, Erev, and Wallsten (1997); Budescu, Wallsten, and Au (1997); Brenner (2003); Brenner, Griffin, and Koehler (2005).

But to bestow pride-of-place on stochastic processes as explaining human decision-making would seem to run the risk of conflating *epistemic* and *aleatory* forms of randomness. The former species of randomness pertains to limits on comprehension. It is the sort of randomness generally invoked when someone (or, when more universal: everyone) fails to understand the behaviour of a force or system: it is then described as *random* because there are no recognisable (to the particular observer) rules governing its behaviour. Still, to assert epistemic randomness is not to claim that such rules fail to exist; when something is pronounced as entailing epistemic randomness the intention is simply that such rules are not *yet* well understood (which does not preclude the possibility that they may never be). Epistemic randomness seems a necessary expedient for incremental science. One explains what one can by using deterministic rules, and then invokes epistemic randomness (i.e., impermanent ignorance) to describe observable departures from predictions made by the deterministic components. The hope in this protocol is that the temporarily-assumed ignorance (epistemic randomness) will eventually become eradicated and supplanted by other deterministic structure (which will typically be imperfect and then require other forms of epistemic randomness).

Aleatory randomness, meanwhile, is of a more permanent persuasion and often gets interpreted as 'genuine chance' – a process for which either: 1) no set of rules can accurately predict individual outcomes; or else 2) the requisite set of starting conditions cannot be measured with adequate precision, and the relationships between these conditions estimated with sufficient ease, such that the process is treated as being driven by pure chance for the sake of convenience and cost. This second form of aleatory randomness depicts the approach to chance in many popular casino games, such as roulette; the underlying mechanics are well-understood, but the necessary calculations and measurements to score a perfect prediction prove either unwarranted, or somehow impermissible.

In the limit, then, it seems that the distinction between epistemic and aleatory randomness is mostly about permanence and convenience.[5] But for purposes of the philosophy of science, and the cultivation of improvements to theory, the difference between these two types of randomness may entail serious implications. The view espoused here is that aleatory randomness in the social and physical sciences should be reserved for either: the well-understood (such as the mechanics of a roulette wheel); or the inexplicable (which is itself not the province of science). Human beliefs seem to meet neither standard. And to assume chance as a core explanation driving the behaviour of human beliefs appears tantamount to privileging aleatory randomness in theories about beliefs.

Truly, human beliefs may never be perfectly understood (especially as the purposes for studying them and definitions of them may well continue to evolve along with both societal and scientific advances, such as furthered understanding of processes and relationships in the human brain). Nonetheless, it seems defensible that extant models and theories of belief perform admirably enough to suggest that randomness is not *the* primary driver of generation and dynamics of beliefs. And, observably, people seem capable of purposed and intentional behaviour that appears far more reasonably described as grounded within determinism, but influenced by persistent instabilities which appear to emanate not from genuine randomness, but from epistemic randomness that people experience in their own lives: limits to personal understandings of the world around them.

### **1. 3**\\ Context, Salience, and Memory Structure

This assertion returns us to the original matter of reflexivity between belief, context, and believer. Many simple forms of reflexivity (i.e., systems subject to mutually-influencing feedback) have been numerously proven as capable of mimicking randomness within static, and artificially-sparse,

[5] Still, in some cases this limit-conversion can also concern *fundamental* feasibility (e.g.: the Heisenberg principle).

environments, even though the governing rules are exclusively deterministic (cf. Wolfram (2002)). It is perhaps unsurprising, then, that human beliefs, judgments, and behaviour should oftentimes appear to engender randomness; after all, humans occupy (for the most part) a world that is neither static nor sparse, but rather is ever-changing (partly due to the beliefs, judgments, and actions of people themselves) and flooded with countless intricate and interacting elements that both come and go through time. The general human fixation on contexts, or subsets of the entire environment of the world in which people find themselves, therefore seems a pragmatic coping mechanism for dealing with the rampant enormity, diversity, and reflexivity of the complete and unfiltered world.

But context appears something that can rarely be perfectly imposed upon a thinking person in any solely-exogenous way. For example, when researchers attempt to experimentally control the 'contexts' to which experimental subjects are exposed, there invariably persists some potential for mismatching. That is, the contexts *actually utilised* by the subjects in their judgments and ensuing behaviours may be: some subset of that which is supplied by the researchers; a superset of the supplied contexts (e.g., subjects go "beyond the information given"; see, classically, Bruner (1973)): or else some mixture of both (i.e., subjects selectively ignore some elements of the experimentally-supplied contexts, but also selectively utilise other elements – presumably from memory, whether through recall or imagination – which are outside of the experimental contexts). In short, as far as human cognition is concerned, the particular context utilised by a decision-maker generally may be partly-endogenous, and ergo reflexive, to that decision-maker. And it is this reflexivity that seems a central driver of instability among not only human beliefs, but also neglect.

Realise that the very notion of context requires neglect. That is, context is both a spatial and temporal bounding of the larger environment, and contextualising something entails a procedure of exclusion (whether implicit or explicit) which removes some possibilities from consideration.

(In such a way, contextualisation can be seen to mitigate some potential ambiguities, and thereby preserve scarce cognitive resources, which could explain people's general attraction to concreteness; nevertheless, people occasionally prefer less context (see Fox and Tversky (1995)).) Those excluded possibilities thereby get neglected. And, depending upon particular circumstances, such neglect may ultimately be either beneficial or detrimental to the decision-maker. The main concern here, however, is that such neglect may also be unstable for the very reason that human beliefs can be unstable: reflexivity. And salience seems to be the key link between instabilities and reflexivity of beliefs, contexts, and macroscopic environment to which beliefs and context pertain.

Like beliefs, salience is relational. Something can only be salient *to someone* or *for someone*, even if that someone is implicit because people are universally presumed to find that thing salient. But, demonstrably, not all people find the same things to be salient, as is evident in the distinctions between novices and experts in many fields; seasoned and expert clinicians, for example, might find very different symptoms to be salient when diagnosing a patient than would a medical student.[6] Likewise, enduring public enthrallment with fictional doyens such as Sherlock Holmes appears to stem significantly from fascination with what that mastermind detective finds salient compared with what readers (or, in more recent reincarnations of Holmes: viewers) might. The relative character of salience sheds light on yet a third fundamental property of beliefs: *relativity*.

Attention will turn to this third belief property shortly, but that discussion will be aided by first noticing that salience is vital for dictating neglect. What is salient is characteristically not neglected, although something being salient typically implies that something else will be (at least partly) neglected. Salience and neglect thus share a rough complementarity. But, as was just

[6] See, e.g.: Ericsson et al. (2006).

mentioned, salience is relative to the observer, and some things all people find salient, but other things are found salient by only a few (and still further things are salient to none). Yet what drives this relativity of salience? The answer seems to reside in the link between memory and recognition.

Recognition requires memory. Something may be seen for the first time by a person without that person recognising the thing because that person has no memory of the thing for never having encountered it previously. But a person may also fail to recognise something she has seen previously because that thing is not immediately accessible to her from memory. If we reasonably assume (as has been demonstrated immensely by cognitive science) that the functional organisation of memory has structure to it, then we should be convinced that the propensity for recognising any particular thing should rely (at least partly) on such a thing's position in memory.[7]

And herein appears to lie a key to the relativity of salience: things which people find universally salient are those that either do not require memory proper (e.g., some forms of exogenous stimuli, such as loud noises or bright lights; yet recognise that differences in memory may make people still differently-attuned to some forms of such stimuli, depending on cumulative previous experience), or for which the particular thing occupies relatively similar positions in the memories of all people. We might call this form of salience *exogenous salience*. *Endogenous salience*, then, might be characterised by salience that is memory-driven, or different across people on the basis of meaningful differences in the structures of their individual memories. A thing might be endogenously salient if it is prominently recognisable by those for whom it is accessible from memory, but less recognisable (even unrecognisable) for those who cannot access it from their memories. But there is also a dimension of activity involved in the distinction between these two

[7] Or else components of that thing, if constructive memory is involved (cf. Schacter, Norman, and Koustaal (1998)).

saliences. Something which is exogenously salient is so whether a person is actively searching for it, or merely passively encounters it. Endogenous salience is trickier, and less stable, in this sense in the respect that some things which are endogenously salient may only be recognised when actively sought, whereas others may become recognised more passively, despite no initial expectation of their being recognised. Endogenously salient things are subject to unstable neglect.

And in social settings this instability seems fuelled to a significant degree by the interplay between the relational and relative properties of beliefs. The relative nature of beliefs is well-established in much of social science: specific beliefs (under most definitions) are constructed, changed, and even abandoned relative to accessible *evidence* (broadly defined, and discussed later) that supports them in relation to other, alternative beliefs. The relativity of belief therefore is twofold, and may also be understood as based on the entire system of beliefs that an individual may be said to possess (or else possibly could possess) at any point in time relative to the evidence accessible by her at that time. Yet beliefs are also relational (in a sense additional to those defined earlier) in that some of the evidence accessible to a person in a social setting might be the beliefs held by other people.

## **1.4**\\ Doxastic-Space Framework for Studying Beliefs

Among the social sciences, not many approaches to studying beliefs are able to balance the three main properties of beliefs discussed thus far: relationality (including connecting both the believer with the believed, as well as believers to other believers), reflexivity, and relativity. It is perhaps thus unsurprising that dealing with instabilities of belief and neglect has posed difficulties for many extant models and theories on beliefs, as well as modes of analysing beliefs. Especially in studying social situations wherein people's beliefs may be influenced by others' beliefs, many approaches to beliefs have faltered for presupposing improper modes of stability among people's beliefs (e.g., consider the case of rational expectations equilibria (REE) in economics). And similarly the

processes for learning rooted in the Bayesian paradigm which has been mated with these models and theories of belief stability have generally underappreciated the simultaneously relational, reflexive, and relative properties of beliefs. There thus appears evident need for an alternative approach to studying beliefs in social science which embeds these three properties centrally into its conceptual and analytical treatment of human beliefs, while providing for their unstable nature. This thesis aims to establish several building blocks for just such an alternative by presenting a new framework for understanding beliefs that is anchored in the flexible notion of *doxastic spaces*.

Before delving into what doxastic spaces are and what a framework built upon them can achieve, some definition of the types of beliefs covered in this thesis must be furnished in order to avoid self-contradiction (see this chapter's opening paragraph, which concerns definition-dependence in studies of beliefs). To make the framework serviceable for as broad of a swathe of the social sciences as possible, a highly-encompassing definition for beliefs in social science, that of beliefs as subjective expectations about propositions about the world, is adopted. This definition claims its modern heritage from philosophy, and remains dominant in economics today.[8] Given the parallels between mathematical expectation and probability, this definition oftentimes compels researchers (at least many economic researchers) to treat beliefs and probabilities as synonymous.

[8] "In epistemic logic, any belief can be expressed according to one of two approaches, the syntactic (propositional) one…and the semantic (possible-world) one…These two contrasted approaches are logically equivalent" (Walliser 2008, p. 9). According to Walliser's foregoing distinction, the approach to beliefs embraced here is largely syntactic, whereas economics often utilises a set-theoretic, semantic approach. As Walliser observes, however, the two approaches can be transmuted from one to another. Specifically, in the semantic approach, propositions are 'events' (collections of worlds), which are propositional in nature. In the doxastic-spaces approach here, propositional events are usually hypotheses, or event descriptions (see Tversky and Koehler (1994)), which stresses a role for communicational detail (or vagueness) in defining beliefs. While this stance may be unpalatable to those wishing to universalise beliefs, and divorce them from context, much psychological evidence (to be reviewed in the sequel) confirms the detail-sensitivity which this approach can enable. Moreover, another motivation often supplied for favouring semantic approaches is that syntactic ones lend themselves to intractability. Recent advances, like Gilboa, Samuelson, and Schmeidler (2013), however, help to make syntactic setups more manageable. The present approach of doxastic spaces looks to further continue that trend.

The definition of beliefs here aspires to be not so confining on the basis of the colossal volume of results from many fields (especially cognitive psychology and behavioural economics) which demonstrate clearly that people often encounter hardship in constructing and using their beliefs in ways that obey the mathematical laws of probability. Thus, the definition favoured here is one that is more encompassing in its attitude toward acceptable ways in which beliefs may become quantified. And while the situations studied in this thesis generally concern those wherein people can manage some quantification of their beliefs, no presumption is made that all beliefs can effectively become quantified: some beliefs may not be migrated from the qualitative domain; these unquantifiable beliefs are those whose general location in a doxastic space is unidentifiable.[9]

And in defining beliefs for the purposes of this thesis and the framework of doxastic spaces which it presents, some further specificity is warranted about the substrates upon which beliefs necessarily rely: propositions. Propositions are the objects in belief syntax (i.e., they are the Z in: "Person A believes that Z"), and the 'degree' of a person's belief is her subjective distillation of the amount of support provided for the truth of the relevant proposition by evidence accessible to her. It is this subjective distillation of evidence for (or against) the proposition in question which is the aspect of belief most open to potential quantification, and which is often taken to be representable as some probability of truth of the proposition. Moreover, the level of detail involved in a proposition, as well as its relationship with other propositions, is of utmost importance in studying beliefs (as they are defined herein). Logically, two propositions with identical accessible evidence in support of their truths *should* generate identical degrees of belief in them when such propositions are entertained in the mind of a single person. If it were otherwise then such a person could be labelled guilty of inconsistency. This point seems almost incontestable. But here lies a

[9] Contrast this approach against "probabilistic sophistication" (Machina and Schmeidler (1992; 1995); Grant (1995)).

conundrum for humans, who wield limited, and even fluctuating, cognitive resources: identical accessibility of evidence for any pair of propositions is not guaranteed; nor is it assured for any particular set of accessible evidence that all relevant propositions upon which that evidence might bear can be constructed or summoned to mind. Yet the state-space orientation that neoclassical economics uses in its treatment of subjective beliefs (and the Bayesian learning that updating such beliefs entails) demands both. The exhaustive nature of the state space seems too demanding for capturing beliefs of people with finite cognitive resources.[10] The idea of a doxastic space provides a manageable alternative to exhaustive state spaces, but can encompass them too as extreme cases.

As a conceptual construct, a doxastic space is both simple and flexible. Doxastic spaces are abstract analytical spaces comprised of anywhere between one and an infinite number of dimensions (although the number of dimensions in the doxastic spaces with which this thesis is concerned will typically be finite). Each dimension in a doxastic space represents a proposition about the world (in some comprehensible language) that is entertained by a specific person, or could hypothetically or potentially be entertained by her, at some point in time. If the person is limited in the number of propositions which she may bring to and hold in mind at once, then we might readily see that her personal hyperspace could be in a constant state of flux, with dimensions being exchanged continually as she moves from considering some proposition to then neglecting it in order to consider another; or she may consider a variable number of propositions through time and thus the dimensionality of her space might shuttle (perhaps irregularly) between collapsing and expanding.

[10] It is interesting to note that, in his seminal book, Savage (1954) seemed to demand very exacting specifications for the definitions of states themselves (i.e., knowledge of which state the world occupies should resolve all subsequent uncertainty), but envisioned far less expansive state spaces than have since become commonplace within economic research. This ballooning of the state space seems to have no single justification in the economics literature, but instead appears to have developed through an organic embrace of the idea of an all-encompassing state space (Gilboa 2010).

But shifting dimensionality is not the only time-variant feature of doxastic spaces. What makes doxastic spaces doxastic is that a person's beliefs are represented by her position along all dimensions in her doxastic space. And it is this representation of beliefs through location within a doxastic space that grants the construct richness and flexibility: a wide variety of types of belief (under the unifying banner of expectation, whether mathematically-obedient or not) can be depicted by altering the scale along any dimension in such a space. That is, one might represent continuous beliefs by allowing a person's belief to be placed along any of an infinite number of points within a dimension, but alternatively permit discrete beliefs by tolerating her beliefs to reside at only a finite number of positions along a dimension. In comparable fashion, one can allow, for example: ordinal as well as non-ordinal beliefs; linear dimensions wherein positions within a dimension are uniform distances apart from one another, as well as curved dimensions with disjoint subintervals of the dimension being of non-uniform length; and dependence or independence between dimensions, such that fixing the location of the beliefs of a person in one dimension restricts (or even completely determines) her beliefs in another dimension. Furthermore, mixtures between these possibilities across dimensions are also feasible. And other possibilities exist, but will not be specifically treated in this thesis. Furthermore, each dimension is in the work here presumed to be bounded and have established endpoints (generally total belief and total disbelief, respectively). Consequently, movement of a person's beliefs along a dimension can be seen as reflecting belief dynamics, whether appropriately described as learning, or else instability.

Yet the doxastic space construct is made more powerful by the possibility of it being social in nature. That is, each doxastic space might accommodate not just the beliefs of one individual, but conceivably many individuals simultaneously. In the most basic of applications, we might consider a situation wherein all people concerned consider an identical collection of propositions, and that

their beliefs about those propositions are scaled in identical fashion, such that they can be placed in a shared doxastic space. Then differences in the positions between different people's beliefs when placed in such a shared doxastic space may be analysed through a host of methods, each suited to a distinct purpose and apropos under particular conditions. For example, to measure total dispersion of multiple people's beliefs across several dimensions one might utilise a measure of entropy in beliefs. Alternatively, one might detect 'outliers' using Mahalanobis distance, or look for inter-dimensional patterns with relational class analysis (see: Goldberg (2011)). Or to examine total 'distance' or 'overlap' between two particular people's beliefs one might use any of a number of measures and metrics, such as Euclidean norms, Jaccard distance, or Hamming distance (to name but a few). Each of these tools carries its own interpretational and practical advantages (and difficulties), but comparisons between them are enabled via a common framing in doxastic spaces.

Still, even further representational and analytical richness is made possible if one allows the subjects of study to entertain different conceptualisations such that they scale the dimensions in their doxastic spaces in different ways, or else entertain different sets of propositions from one another. In such a way one person's doxastic space may not match another's, and comparisons between their respective spaces can help to trace and distinguish differences in their ensuing perspectives and consequent behaviours. Moreover, as is commonplace in social settings, the beliefs of one person might be influenced by others' beliefs, but others' beliefs might be only imperfectly visible; then the beliefs of a person about others' beliefs can be represented in a doxastic-space format with only a shift in interpretation: the various points along any dimension in the doxastic space for one person then do not represent other people's actual beliefs, but the person's beliefs about others' beliefs (which may be accurate or not, and may also influence the

position of her own beliefs along that and other dimensions). Hence, the doxastic space is an inherently social construct, and can be applied to both the study of individual and collective beliefs.

## **1.5**\\ Going from State Spaces to Doxastic Spaces (and Back Again)

Nor is the doxastic space entirely incompatible with the state-space approaches of neoclassical economics. Indeed, those approaches typically confer upon people a shared space comprised of all relevant states of the world (which may be either finite or infinite, depending upon the application) – which are effectively propositions of a highly detailed nature (in the well-accepted view of Savage (1954), knowledge of which state the world is in resolves all meaningful uncertainty for a person). Still, in those approaches people do not form beliefs about states, but rather collections of states: events. States then merely serve as binary 'tokens' in formation and maintenance of beliefs; people understand particular states as either true or false, and their beliefs are then formed by relative (that is, normalised) tallies of truths and falsities among collections of states (i.e., events).

The doxastic-space approach can accommodate such a setup quite readily, and therefore becomes useful for comparative exercises wherein, e.g., predictions from non-state-space formulations might be compared against and judged on equal footing to conventional state-space formulations. But, as some chapters of this thesis demonstrate, a doxastic framework has value for being able to easily accommodate setups and explore problems that prove difficult in the state-space paradigm. One problem that doxastic spaces are better-equipped to investigate than conventional state spaces is that of the genesis of beliefs. That is, state spaces, and the expected utility approaches that rely on them, generally require that people formulate 'prior' beliefs, yet do not have much to say on how prior beliefs might come into initial existence. Nor are conventional state spaces ideal substrates for studying departures of belief dynamics from what would be predicted by Bayesian

models of belief change. Doxastic spaces, meanwhile, can provide insights about both problems. Chapters contained in this thesis demonstrate several ways in which those insights can be derived.

## **1.6**\\ Unifying Framework for Interdisciplinary Exchange

While the primary purpose of this thesis is to supply a new framework for studying human beliefs, its secondary, and farther-reaching, ambition is to bring into closer contact ideas, methods, and researchers from across the social and physical sciences. Although interdisciplinary collaboration, communication, and interest is perhaps greater now than it has been in modern scholarly history, considerable further progress remains to be achieved in promoting potentially fruitful marriages between theories, methodologies, and perspectives among social-science disciplines. It is little exaggeration to say that many of the most transformative advances in social and physical science have come from interdisciplinary 'borrowing'. But, given gains realised by such episodes of transference, it is disheartening that more involved (and reciprocated) exchange is yet to be found.

Part of the paucity of such fluid and multilateral exchange is surely barrier-oriented: there are high costs to fluency in the *lingua franca* of any academic discipline in terms of non-superficial understandings of its canonical theories, methods, and findings. But an even higher cost is borne for familiarity and penetrating comprehension of any discipline's frontier work: the latest cutting-edge empirics and theorising that creates buzz (and controversy) and inspires others to partake in mimetic prospecting. Yet the most fruitful interdisciplinary collaborations would seem to be at the intersections of such frontiers: in areas where theory and analytical tools have not yet had time to ossify, and might be most benefitted from the novel perspectives of 'outsiders' because such frontier pursuits are novel (and often unorthodox) within their own parent disciplines. It is to this end that shared frameworks for theorising and analysis might prove most useful for social science. This thesis intends the framework of doxastic spaces to eventually serve as one such shared nexus.

As proof of the doxastic space framework's potential for interdisciplinary serviceability, the subjects treated in this thesis merge ideas from a cross-section of fields within the social and physical sciences, as well as humanities. The three traditions of thought which are most represented are economics, psychology, and human geography, but sociology, physics, neuroscience, philosophy, statistics, and computer science also are present. Within the four chapters forming the body of the thesis, the following research topics (each of which resides at the frontiers of its parent discipline) all find a place: approximate number systems (also known as approximate magnitude systems) from cognitive science; entropy, which is common in physical science, but less so in social science; neuroeconomics, which represents a mosaic of disciplines; cognitive hierarchy theories of non-equilibrium strategic thinking from behavioural game theory; quantization methods from signal-processing; and rare-disaster theories from financial economics.

## **1.7**\\ Structure and Conventions

The essential structure of this thesis is as follows. The next chapter provides a selective curation of perspectives on beliefs in the social sciences over (approximately) the last half-century. That tour of the literature emphasises primarily economic and psychological treatments of belief, as they have been the branches of social science to have made the greatest progress in study of beliefs (although they have not always agreed with one another). Chapters 3, 4, 5, and 6 represent works all connected to topics of beliefs and doxastic spaces, albeit in diverse respects. Each of those chapters is intended as a self-contained work, and offers its own suggestions for possible future extensions of specific ideas covered therein. Nonetheless, Chapter 7 presents some routes for extending the doxastic-spaces concept more generally and identifies several potential ally projects.

It may assist the reader to have a brief overview of each of the four body chapters of this thesis. Chapter 3 ("Minding maps in mapping (neglectful) minds: Multi-scale reflexivity in neural space")

explores, from a neurophysiological standpoint, how stable or unstable beliefs might actually be capable of being. This chapter thus frames succeeding ones by suggesting the extreme flux of beliefs on the basis of their ultimate source of generation: neural activity, which is itself limited by neural structure. Chapter 3 deals specifically with the topic of neuroeconomics, and states a case for engagement by economic geographers with that emerging discipline. The foundation for that argument is thus: economic geography possesses a generic advantage in appreciating many of the concepts (such as dynamic reflexivity and interdependencies across multiple scales of space and time) which underlie brain function and are therefore necessary for both judgment and behaviour by human beings. Chapter 3 is substantially forward-looking, in that it deals with the topic of beliefs and doxastic spaces indirectly by questioning a necessary substrate on which beliefs (as cognitive elements) must rely: neural spaces. 'Conventional', single-scaled approaches (that favour time-space localisation) in neuroeconomics are critiqued, and a rationale supplied for why economic geographers may gain from mutualistic engagement with neuroeconomics. Part of that rationale is a deeper understanding for the topic of neglect, which is essential to both the notions of context and myopia (both of which are of native and strengthening interest for economic geographers). While a theoretical exposition, Chapter 3 is based on, and marshals, a substantial corpus of research in cognitive and neuroscience.

Chapter 4 ("Cognitive limitations on belief approximations: A theory of evidential quantization") introduces the doxastic space in its simplest format: with a single dimension and decision-maker: it presents a substantially psychological approach to the formation and distortion of beliefs at the level of the individual. That chapter offers a cognitively-motivated theory about how decision-makers (DMs) formulate beliefs based upon evidence accessible to them, but may experience distortions of those beliefs due to constraints on their working memory capacities. Empirical work

on so-called *approximate number systems* and the ratio-dependent nature of human judgment are major components of that descriptive theory, called quantized evidence theory (QET). In particular, Chapter 4 shows how QET can make predictions on situations where DMs might exhibit either subadditivity or superadditivity in their probability judgments due to the influence of their confidence in evidence accessible to them (as a function of both strength and weight of evidence). It also demonstrates how QET may be capable of resolving some longstanding difficulties in explaining the 'description-experience gap' phenomenon that remains disputed within psychology.

Furthermore, QET is certainly not the only theory which might be studied within the doxastic-space framework, but is one which is particularly amenable to such examination, especially when it is extended to learning by multiple DMs simultaneously over time. Such extension is pursued in Chapter 5 ("Entropy-based social learning"). Chapter 5 uses a more complicated doxastic space than does Chapter 4, and involves multiple DMs and dimensions. It provides a theory of social learning based upon the concept of agglomeration in beliefs (doxastic agglomeration): entropy-based social learning (EBSL). EBSL exploits a link between the idea of agglomeration within a belief space (i.e., concentration by DMs in terms of belief density) and a decomposition technique for Shannon's (1948) measure of informational entropy.[11] EBSL aims to describe how people may learn about phenomena through changes in consensus when they cannot observe one another's beliefs directly, but have some inkling about how relatively dispersed or concentrated beliefs are. The EBSL model embodies some pleasing properties (e.g.: endogenously self-tuning parameters; a clear cognitive foundation, which is based upon QET; and the ability to handle substantial diversity in the beliefs of DMs which it models). The EBSL model is simulated in Chapter 5 using a novel, large-scale dataset of investors' beliefs about the resolution of the 'Fiscal Cliff' matter in

[11] A static study of beliefs based on thermodynamic entropy appears in Rook (2012).

the United States in late 2012. These exercises with EBSL show how it is capable of explaining sudden, unexpected shifts in a belief distribution through a process (belief echoing) that is comparable to, but generally distinct from, informational cascades. Finally, Chapter 5 shows how belief echoing and EBSL-like cognition may drive some larger anomalies within financial markets.

Chapter 6 ("Disagreeing on agreeing") utilises the QET-based EBSL model and the doxastic-space framework to analyse situations wherein DMs disagree about whether or not they actually have agreeing beliefs (and they cannot communicate those beliefs directly) resulting from cognitive heterogeneity among them. Such disagreements about agreement are termed "asymmetric consensus misperceptions". Chapter 6 delves into the topic of belief consensus, as well as common knowledge and common beliefs, and offers a practical 'shortcut' for exploring the issue of commonality (which was greatly popularised in social science by Aumann (1976)) in ways that involve finite verification steps. This shortcut is based on belief agglomeration. The chapter also investigates (using the doxastic-space framework) some results on asymmetric consensus misperceptions when DMs are allowed to vary in the ways by which they can exhibit cognitive diversity. A general finding from this analytic investigation is that cognitive diversity, and asymmetric consensus misperception, may not be rare, and implications for rational expectations equilibria, as well as so-called 'cognitive hierarchy' models of behaviour are treated in some detail.

Before wading further, a brief mention of conventions on enumeration used throughout this thesis is warranted. First, all sections and subsections in each chapter are prepended with the chapter number, as are all figures, tables, and appendices. Equations, propositions, and examples, however, retain a 'hermetic' numbering scheme (i.e., they are specific to each chapter, and do not carry over sequentially from one to the next). Footnotes, however, remain sequential throughout.

Additionally, Chapters 3, 4, 5, and 6 are worded in American English, per the requirements and/or preferences of the journals to which the papers that those chapters represent have been submitted.

## **1.9**\\ Boundaries and Methodologies

Thus far, focus has squared on what this thesis 'does', but it seems responsible to remark briefly on what it 'does not'. First and foremost, it does not, and, given restrictions on space and time, cannot, provide an exhaustive and complete treatment of doxastic spaces. Indeed, it only inaugurates that mission, and the fact that it cannot be finished in the same document in which it begins hints that it harbours substantial future promise as a topic for continued expansion and pursuit. Second, this thesis does not deal with the titanic topic of human preferences, except only tangentially. Such omission is intentional, and triply-motivated in that: 1) the topic of preferences appears to have already garnered more scrutiny and interest in many of the social sciences (particularly economics and psychology) than beliefs; 2) new thinking about preferences seems to be hampered by a paucity of new thinking about beliefs; and 3) once again, limits in both space and time anyway preclude a significant foray into preferences. Nonetheless, whether beliefs and preferences are fully separable has become questioned recently in social science (based on plentiful experimental findings; see references listed later in this thesis), and thus the doxastic-spaces approach may remain incomplete until it is more closely wedded with understandings of (particularly the instability among) human preferences. Chapter 7 poses several such possibilities.

Finally, the papers contained herein rely substantially on the concept of 'accessible evidence' as a basis for human beliefs. Still, all are rather silent about the exact mechanics behind the selection and acquisition processes for this accessibility in evidence (Chapter 2 does treat several elements of this matter; likewise Chapter 7 provides some indirect insights on this issue.) Yet, once more, my election to focus on what is done with accessible evidence instead of how its accessibility is

secured has been a driving motive shaped by limits on resources of space and time. Assuredly, future exploration of accessibility processes should promote further discoveries about the natures and behaviours of human beliefs, and (as with preferences) a framework of doxastic spaces seems not fully complete without a deeper understanding of such evidential accessibility. Chapter 7 supplies what impress me as paths that wield considerable promise for investigating accessibility.

Also, an aside on methodology is needed. This thesis is mostly a theoretical contribution, but one propelled by descriptive intent; I have strived to the best of my abilities to generate models and constructs that fit the facts relatively closely, but are also manageable enough to permit some generalisations and portable insights. The length of the reference list should testify to the effort devoted to ascertaining that this fitness with empirical demonstration and experimental results elsewhere is reasonably close. Mathematical analysis, computational simulation, intuition, and narrative logic also play their respective roles in the methods employed by this thesis. A substantial empirical contribution, however, emanates from the survey-derived data treated in Chapter 5. More will be said on the novelty of that instrument and its findings later. Nonetheless, while this thesis is skewed towards theory, the reader should recognise that it engages in a variety of theoretical styles, from the formal to the more qualitative and informal (on the latter, see Chapter 3). In all, the contribution made is one which has tested the author's abilities in critical, rigorous thought, and should perhaps be assayed not only for what it does and does not do, but what it might foster.

# CHAPTER 2

# REVIEW OF LITERATURE

## **2.1**\\ Overview

Each of the four chapters following the present one contains its own separate review of literature relevant to the specific material which it covers. As there is little need to repeat in this chapter the more focused coverages undertaken in those chapters, the immediate ambition is to provide a more macroscopic treatment of developments in the conceptualisation and study of human beliefs within the social sciences over the last (roughly) half-century. Of course, to give an exhaustive exposition of such a topic as broad as this one is impossible within the space allowable for this thesis. The focus of this chapter will therefore be more selective, and aim to illuminate paths both more- and less-travelled by social scientists over recent decades in their endeavours to understand and predict (and even influence) human beliefs. Moreover, the emphasis here is skewed towards work done in economics and psychology, as those fields have made greater advances in the formal study of beliefs (both quantitatively and qualitatively, as well as theoretically and empirically) than others within the social sciences (at least with respect to the definition presented in the previous chapter of beliefs as subjective expectations on propositions about the world).[12] Where appropriate, pertinent literature and developments in other fields (especially geography, political science, mathematics, and sociology) are also discussed. The goal is to lead the reader through a guided tour of where the scholarly study of beliefs has ventured in recent years, as well as where its present combination of momenta within various academic disciplines might take it in the near-term future.

[12] Once again, this embraced definition conforms more strongly to the syntactic approach to beliefs, as discussed by Walliser (2008). On semantic approaches to beliefs (in game theory), see the review by Battigalli and Bonanno (1999).

## **2.2**\\ Backdrop to Expected Utility

It is scarcely possible to discuss the study of human beliefs in the modern social sciences without treating the topic of expected utility.[13] Indeed, the very trajectories of conceptions about what human beliefs *are* within various branches of social science (and especially economics and psychology), have been indelibly shaped by the flows and ebbs in support for the two main strands of the expected utility hypothesis: that attributed to von Neumann and Morgenstern (1944; 1947) (henceforth 'expected utility' (EU)); and that due to Savage (1954) (henceforth 'subjective expected utility' (SEU)). A core distinction between these two strands resides in the nature of the probabilities which they employ.[14] The former deals with 'objective' probabilities – those which are exogenously given and (ostensibly) not disputed – and the latter with 'subjective' probabilities – those which the decision-maker must formulate herself and are "personal" (in Savage's words).[15]

Yet to properly understand the potency of these two variants of the expected utility hypothesis as innovations in social science, we must respect the fact that, like all innovations, they entailed two separable components: *invention* and *implementation*. As inventions, both EU and SEU are elegant

[13] Schoemaker (1982) claims that: "It is no exaggeration to consider expected utility theory the major paradigm in decision making since the Second World War" (p. 529). See that paper for an excellent discussion of expected utility. It is also startling to realise that the roots of expected utility trace back to Bernoulli (1738) as an attempt to resolve the St. Petersburg paradox; that is, expected utility began its early conceptual life as a direct attempt at *descriptive* theory.
[14] It seems wholly plausible that theories of expectation in economics (e.g., Shackle (1952)) which deviated from more conventional ideas about probability (i.e., the Kolmogorov axioms) were disadvantaged for shunning orthodoxy. There are, nonetheless, a great many 'unorthodox' theories of probability that have emerged (consult Gillies (2000)).
[15] See Camerer and Weber (1992). Anscombe and Aumann (1963) also provided a much-cited bridge between objective and subjective probability, yet their approach requires mathematical devices extraneous to Savage's (1954). (See also Machina and Schmeidler's (1992) "robust" approach to defining subjective probability.) Within the social sciences there seems to be rather substantial variability from one particular field to the next in terminology which distinguishes among different types of probabilities. The convention that will be used here is that: ***risk*** describes situations of objective probability (wherein probabilities are 'given' by some exogenous source); ***uncertainty*** describes situations of subjective probability (wherein probabilities are not exogenously supplied and decision-makers must themselves construct subjective estimations of probabilities); and ***ambiguity*** describes situations of subjective probability wherein a decision-maker cannot characterise a judgment problem by a single probability distribution (with one likelihood per conceivable outcome). This taxonomy can be understood as applicable from the point-of-view of the decision-maker herself (rather than the analyst studying such a decision-maker). And, while this taxonomical convention may be criticised for imperfections, it appears to accord with popular approaches elsewhere.

feats of mathematics and logic. But it seems that the relative fitness of these ideas within the intellectual landscapes from which they emerged has played at least as significant a role in their durability and popularity as has their inventiveness; that is, the manners in which EU and SEU could be implemented after their creation found them in considerable demand, and saw them used as foundations for numerous novel theories and fields of research. Nonetheless, of both EU and SEU, it was SEU that saw more challenges soon after its introduction (and, in some cases, even prior to its formal publication; see famously, e.g., Allais (1953)). To appreciate the literature on human beliefs in the social sciences, then, it becomes incumbent upon one to comprehend the intellectual backdrops from which SEU emerged, and in which it swiftly gained uneasy hegemony.

Savage's (1954) approach to subjective expected utility was erected in the midst of sweeping transformation in the social sciences and mathematics. The zeitgeist of change in mid-century scholarship in those areas is hard to overstate, and included, among a host of other movements and intellectual sea changes: the rise of logical positivism (propelled vocally in economics by Milton Friedman (1953), a collaborator and early mentor of Savage's; cf. Friedman and Savage (1948)); the recent overthrow of cardinal approaches to utility by ordinal utility; the groundswell support for revealed preference approaches (credited to Samuelson (1938)); struggle between 'frequentist' and Bayesian philosophies of probability and statistical methods; the implications in mathematical logic from Gödel's (1931) incompleteness theorems and the ensuing Church-Turing thesis (Church 1936; Turing 1936); and rise of statistical definitions for information (Fisher 1925; Shannon 1948).

There is not space here to elaborate fully on the imprints of each of those developments upon both Savage's work and its subsequent welfare in social science; but a cursory sketch may well suffice. Recall that Savage was, per his own confession, mainly a statistician. And indeed one gains a keen sense in the opening of his 1954 treatise on subjective expectations for his concern with

enlightening the reader about key distinctions between the major treatments of probability in that day. Specifically, he casts his focus on relative merits and demerits of frequentist versus Bayesian (also: 'subjectivist' or 'personal') perspectives on probability and statistics (so-called 'logical', approaches to probability, such as Keynes' (1921), had since fallen from the forefront of consideration in mainstream statistical scholarship), and effectively supplies mathematical and philosophical foundations for a Bayesian view (hence the title of his 1954 masterpiece). Among Savage's critiques of the frequentist approach (which he terms as "objectivistic") is that, under it:

> …probabilities can apply fruitfully only to repetitive events, that is, to certain processes; and (depending on the view in question) it is either meaningless to talk about the probability that a given proposition is true, for this probability can be only 1 or 0, according as the proposition is in fact true or false. Under neither interpretation can probability serve as a measure of the trust to be put in the proposition. Thus the existence of evidence for a proposition can never, on an objectivistic view, be expressed by saying that the proposition is true with a certain probability. (1954, p. 4)

Thus, like Knight (1921), Savage found troubling the application of frequency-driven statistical approaches to the probabilities of "unique" events; but he was also concerned about the binary approach adopted to belief under the frequentist paradigm. Instead of looking at propositions as either true or false, Savage favoured Ramsey's (1926) stance on probabilities as "degrees of belief", which could range between zero and one. Furthermore, Savage also espoused Ramsey's notion of belief as able to be operationalised through "willingness to bet".[16] Indeed, equating a decision-maker's beliefs with her acceptance of wagers on those beliefs placed beliefs on comparable footing to preferences as suitable objects for 'scientific' study; beliefs could thus be 'revealed' in a fashion similar to preferences. Indeed, Savage's setup takes preferences as primitive, and (conditional on a decision-maker obeying a series of specified preference axioms) derives her beliefs *from* her preferences. This approach to simultaneous derivation of preferences and subjective probabilities was seen as a major triumph in the theory of human choice (Kreps

[16] Fox and Tversky (1998), however, observe that Émile Borel seems to have anticipated the idea of measuring beliefs based upon methods of elicited preferences.

(1988, p. 120) refers to it as the "crowning glory of choice theory"). And in the rising tide of logical positivism, which stressed empirical verification and proof over normative rhetoric about suitability, the ability of Savage's approach to expose beliefs and preferences to hypothesis testing immediately endeared it to many social scientists (particularly economists). Moreover, Savage's approach meshed cleanly with (the by then well-known) Bayes' rule for updating probabilities; social science now possessed a structured core of dynamic theory for 'rational' human behaviour.

The dynamic applicability of Savage's SEU approach was of pronounced importance at the time of its development due to a recent evaporation of hope for deductive completeness in mathematics after findings by Gödel, Church, and Turing.[17] These mathematical developments demonstrated that, within (most) formal axiomatic systems, there would generally remain propositions that could not be proven within the system itself. By eliminating the general possibility for an all-encompassing (singular) deductive logic of mathematics, these findings implicitly highlighted the importance of inductive methods (especially statistical inference) for theorising. And these methods came to prioritise quantification (as the mathematical 'value') of information on the heels of work by earlier statistics research (e.g., Fisher (1925)) and communications theory (specifically Shannon (1948)).[18] Savage's approach was able to show how beliefs *should* (normatively) adjust to arrival of new information if a decision-maker were to conform to specified preference axioms.

Given how snugly Savage's approach to beliefs seems to have fit the demands of its day, it is perhaps unsurprising how readily it was accepted in social science. But there was one big problem: Savage never intended his approach to be descriptively accurate. Indeed, he repeatedly mentions in his 1954 book how his treatment of subjective probabilities and preferences was meant only as

[17] See also: Copeland, Posy and Shagrir (2013). Likewise, consult extension of these findings by Chaitin (2005; 2006).
[18] For condensed introduction to the history of communications theory and statistical information, see Kullback (1959).

a normative guide. It should thus not be shocking to witness how quickly and extensively Savage's approach was shown to have descriptive shortcomings. Famously, both Allais (1953) and Ellsberg (1961) furnished early (eponymous) 'paradoxes' which contradicted predictions made by Savage's approach. And in psychology (which never embraced Savage's views as fully as economics did) a systematic deconstruction of Savage's axioms through a series of experiments helped to birth and popularise one of the most radical and transformative projects within modern social science: the so-called "heuristics and biases" programme inaugurated by Amos Tversky and Daniel Kahneman.

In an early paper, Tversky and Kahneman (1971) showed how people's intuitions about probability could be systematically biased. In successor studies (Kahneman and Tversky 1972; Tversky and Kahneman 1973; Kahneman and Tversky 1974) the pair showed how people's beliefs could be contextually manipulated to produce systematically-biased patterns of behaviour. Their explanation for the pervasiveness of this manipulability was that people often engage in incomplete thinking, and employ expedients (or "heuristics") to preserve scarce cognitive resources. The power of this idea of heuristically-guided thinking for many psychologists was sweeping, and over the next few decades identifying and cataloguing 'new' heuristics and biases became something of a sport in the discipline (see, e.g.: Kahneman, Slovic, and Tversky (1982); Kahneman and Tversky (2000); Gilovich, Griffin, and Kahneman (2002)), even if not all psychologists were convinced that heuristics necessarily lead to biased or detrimental judgments (see, e.g.: Gigerenzer and Todd (1999)). These findings invigorated the idea that people face meaningful limitations on their cognitive abilities, and that constrained cognitive resources can alter beliefs and preferences.

But these ideas of finite cognitive resources were new neither to psychology nor economics. They had already been (e.g.) demonstrated in experiments by Miller (1956), and articulated in clever 'models' by Simon (1955; 1956) (indeed, even those well-cited works cannot claim full credit for

the origins of such ideas; nonetheless, they remain among the most-revered works for the gripping imagery which they supply).[19] Simon in particular promoted an algorithmic orientation to the modelling of human thinking, and expressed as early as 1974 his hopes for parameterising the human mind.[20] Such procedural approaches were not quick, however, to meet the approbation of economists, and, despite descriptive deficiencies in Savage's approach to beliefs, the SEU view of behaviour came to gain deep importance within economics due to three complementary trends. Interestingly, these three trends all have direct connections with the study of financial markets, which seems to have become far more concretely entrenched as one of the leading topics on the global economic research agenda since the advent of SEU, and hint at the degree of reciprocity (and mutual reinforcement) which SEU and financial theorising have enjoyed in recent decades.[21]

### **2.3**\\ Entrenchment of SEU

The first trend involved portfolio theory, which was developed by Markowitz (1952a) and initially used objective probabilities.[22] Portfolio theory would eventually become modern mathematical finance's cornerstone, and give rise to such powerful ideas as the capital asset pricing model (CAPM) (see: Treynor (1961; 1962); Sharpe (1964); Lintner (1965a; 1965b); Mossin (1966)), and, in turn, the Black-Scholes-Merton model of derivatives pricing (Black and Scholes 1973; Merton 1973).[23] Yet these prominent constructs in academic finance which are rooted in portfolio theory

[19] For a collection of some of Simon's many, many models of human thinking, see Simon (1979).
[20] Still, in many of his works, Simon suggests that the parametric structure of the human mind may never be completely determined in quantitative terms. He nonetheless often advocates the identification of "invariants" in human reasoning, judgment, and choice, and stresses the importance of building at least qualitative models of thought (see: Simon 1990).
[21] For a detailed chronology of connections between SEU and finance, see Rubinstein (2006).
[22] A contemporary alternative to Markowitz's theory was Roy's (1952) less-popular "safety first" approach. One can hardly help but wonder what sort of state the modern financial economy might be in had that popularity been reversed.
[23] For a perspective on the Black-Scholes-Merton (BSM) model's history in application (i.e., as used by traders rather than academics), see Haug and Taleb (2011). Nonetheless, one gains a sense of the influence which the BSM model and other theories enabled ultimately by Savage's approach to probability carried over the academic community (of economists at least) by inspecting how many of those credited with their invention earned 'special trips' to Stockholm. Moreover, reflexivity between such academically-originated theories and practice seems legitimate (Mackenzie 2006).

(and formalisation of the principle of financial diversification that it captured) do not rely on the original version of portfolio theory (i.e., Markowitz (1952a)) so much as they do on its successor (i.e., Markowitz (1959)), which integrates Savage's (1954) approach to uncertain probabilities into the original version of portfolio theory, and allows progression from *given* to *uncertain* beliefs.[24]

The second trend came from the rise in formal thinking about informational market efficiency and the implications for the randomness of observed price behaviour in financial markets.[25] This genus of thinking, in which informational efficiency implies randomness, was enabled by formalisations of the concept of information, as in Shannon (1948).[26] In Shannon's setup, information is defined on the basis of one's beliefs about some state of the world (i.e., some proposition) relative to other possible states. Essentially, Shannon's perspective on information is that it should be defined on the basis of *ex ante* beliefs (formulated as mathematical expectations); this formulation effectively makes beliefs a primitive of information (but still leaves the awkward question of how one is to construct beliefs in the first place). A similar perspective on information was readily adopted (using SEU-beliefs) among mathematical economists, as were ideas about the efficient use of information in market settings (i.e., that the prices of assets and securities traded in financial markets would rapidly come to reflect all information available about their prospective future values, such that the current prices of such assets and securities would generally and continually represent the 'most informed' forecasts of their values when properly discounted). The writings of (e.g.) Fama (1965) and Samuelson (1965) were particularly influential in propagating these ideas

[24] See Markowitz (1991; 1999) on the early history of portfolio theory.
[25] Ideas of market efficiency date back at least to the dissertation of Bachelier (1900). For a deeper overview of the history of informational efficiency in markets, see: Dimson and Mussavian (1998).
[26] See also Fisher (1925). It is notable that Shannon's version of information also ushered in the age of digital communications, upon which modern financial practice now relies heavily. Nonetheless, it has been remarked that Shannon distrusted serious application of his ideas outside of the intended communications theory (see Luce (2003)). It is interesting, then, to note that 'information economics' (which has been widely associated with Stiglitz (1961; 2000) has focused mostly on asymmetries across market agents (instead of diversity in interpretations) for information.

in the economics community.[27] And Samuelson's acclaimed "proof" that asset prices which follow martingales will fluctuate randomly helped to propel interest in mathematics of efficient markets.

Vital for much of the mathematics of informational efficiency in markets is what has become known as the '*First Fundamental Theorem of Asset Pricing*': the principle of 'no-arbitrage' (for some key mathematical developments of this idea, see: Harrison and Kreps (1979); Harrison and Pliska (1981); Delbaen and Schachermayer (1994)). No-arbitrage concepts rely on the notion that immediate and risk-free profit opportunities are not possible within information-efficient markets. Of course, absence of arbitrage opportunities relies upon the underlying condition of market participants having 'correct' beliefs about the price-value relationships between all relevant assets and securities in a market. This notion of correctness in beliefs ties directly to rational expectations.

Rational expectations and their associated equilibria represent the third trend in economics which was impacted significantly by ideas of SEU, but, of the three trends mentioned here, was that which possibly had the strongest effect upon how subjective beliefs came to be treated in economics. An early formulation of the theory of rational expectations was presented by Muth (1961), but the later embrace of the idea by scholars like Radner (1972) and Lucas (1972) were the true accelerators of that theory within economic research. It is surprising to note the diametric opposition of the main messages of the two schools of thought represented by rational expectations (RE) and the heuristics and biases (HB) programme of psychology. RE views (conventionally) maintain that people are not systematically biased in their beliefs, and that the mistakes they make

[27] Many elements of the ideas of the informational contents of prices, and their role in the market-wide dissemination of information appeared earlier in Hayek (1945). One must remember further two facts about the initial popularity of efficiency in market-dictated prices as a guiding theory of economic behaviour: 1) the bulk of the most-watched work on informational efficiency was being done by American economists; and 2) this work was being done in the depths of the Cold War, which may have created non-negligible pressure to defend Capitalist ideologies centred on markets.

in anticipating some element of the future are only idiosyncratic.[28] The HB paradigm, meanwhile, underscores the perspective that systematic biases in people's beliefs do frequently arise, and that, although idiosyncratic errors are also common, these systematic components are highly influential.

While undertaking a formal tally of empirical results favouring the HB view over that of RE, and *vice versa*, would surely prove daunting, at first suspicion one might suppose that evidence for either must detract from the other (or, more scientifically, evidence against either should favour the validity of the other). Some resistance against this zero-sum perspective has, however, surfaced in corners of economic theorising. And a main fuel for this resistance seems to lie in convictions that competitive forces (painted in evolutionary, survival-of-the-fittest rhetoric) should eradicate systematic biases from market settings. Moreover, criticisms by some RE advocates that HB studies are typically conducted in non-market settings claim that evidence of systematic biases at the individual level need not transfer to the aggregate level of collectives and markets when forces such as arbitrage might discipline subjective beliefs. Furthermore, RE occupies a highly-defensible position in the respect that it is (unfairly) difficult to disprove in a universal sense because it requires not only that market participants wield accurate models of the phenomena in question (characteristically referred to as the particular data-generating process, or DGP), but so too must the researchers studying them. Thus, 'tests' of RE become almost Sisyphean tasks in practice, as one might always suppose that the researcher's model and beliefs are inaccurate, while those of the market participants being studied remain truly accurate. More will be said later in this chapter about methodological aspects of the study of RE, but for now it suffices to recognise that other forces apart from scientific best-practice appear to motivate durability of RE views in economics.[29]

[28] RE implies that people are stable in the ways in which they are unstable in their beliefs, because, under RE logic, people cannot switch between systematically-flawed beliefs, and are thus susceptible to unstable forms of instability.
[29] By scientific best-practice is intended the concept of 'falsifiability' that is popularly attributed to Karl Popper (1935).

Perhaps most clearly identifiable among these other forces is the mathematical simplification that RE extends to much economic theorising about market function. Specifically, RE approaches to modelling beliefs greatly enable usage of representative-agent setups, whereby theorists need not consider all market participants separately, but can instead treat them as a single entity by presuming homogeneity among their beliefs. And while some economic researchers have made inroads in thinking about belief heterogeneity (see, early on: Harrison and Kreps (1978); Kirman (1992)), only recently has diversity in beliefs started to become a serious topic for mainstream study (see Hong and Stein (2003; 2007)), and even still the extent of diversity permissible remains typically limited (further discussion of recent developments in belief heterogeneity appears in later chapters). And these limits to belief diversity are driven considerably by concerns for tractability and practicality; as economics generally prizes parsimony in its theorising, models which feature heterogeneous beliefs as centrepieces are often disadvantaged against those depicting homogenous beliefs, particularly when the latter group can justify the general disappearance of diversity in beliefs with arguments anchored in such powerful concepts like learning and common knowledge.

The concept of common knowledge has become closely linked with RE in much economic thinking. Indeed, the seminal treatment of common knowledge due to Aumann (1976) mentions (in a footnote) RE as one of the potential fields of application for common-knowledge results (Aumann specifically mentions works by Radner (1972) on RE, as well as Grossman and Stiglitz (1976) on informational efficiency). Common knowledge is both a simple and challenging idea as formulated by Aumann. At its core, it merely involves agreement by people; but it is the infinite extensiveness of such agreement that engenders some hardships (both theoretically and practically). Nonetheless, Aumann's result furnishes a potent equilibrium concept: given identical shared prior beliefs, and common knowledge of one another's posterior beliefs, two agents cannot

agree to disagree if they are "rational"; they must achieve some consensus in their posterior beliefs. In essence, given a shared starting point in their beliefs, people must achieve a 'belief equilibrium' (amounting to a consensus) if they are all rational and have common knowledge of their posteriors.

The potency of such an equilibrium concept in beliefs for economic theorising should not be underestimated. Under the appropriate conditions – identical priors (also called, confusingly, 'common priors'), common knowledge of posteriors, and rationality – one can essentially disregard belief heterogeneity, as it can be presumed to be only transient because permanent disagreement among beliefs becomes impossible (it should be noted that, by the time of Aumann's results on common knowledge, Harsanyi's (1967; 1968a; 1968b) arguments for common priors, which centred on the notion that all differences in beliefs should stem ultimately from differences in information, were already well-established and widely accepted among the economics community).[30] And such an equilibrium concept became even more advantaged by subsequent findings that both common knowledge and rational expectations equilibria could be 'learned' by rational agents (see especially: Geanakoplos and Polemarchakis (1982); Bray (1982)). This combination of common knowledge and learning meant that belief homogeneity could be both a permissible and fruitful (in terms of elegant extensions for theory which it enabled) assumption.[31]

[30] Yet, despite practical usefulness from assuming common priors, the philosophical justifications for this common priors assumption (CPA) have been variously debated by many notable authors (see especially: Berheim (1986); Aumann (1987); Morris (1995); Gul (1998)). Interestingly, even some of the most ardent supporters of the CPA question its descriptive validity; Aumann (1998, p. 929), for example, remarks: "To avoid misunderstanding, we stress that we do not consider the CPA "true;" the concept of truth does not apply here". This curious remark seemingly emanates from disputes over the dynamic nature of information, and how accessibility of information matters to belief. In financial economics, considerations of the CPA have focused more on issues of pragmatism and concern over implications, rather than philosophical roots or foundations (see, e.g.: Varian (1985)), but still limit extent of diversity.

[31] Of course, one can manipulate beliefs (as expected probabilities) to be any arbitrary value by simply altering the propositions upon which they rely. The common state space construct (wherein economic agents share a single, identical, and expansive state space) seems to have emerged in economics to address just that problem (Gilboa 2009).

## **2.4**\\ Fissures and Uncertain Directions

Nevertheless, some rather uncomfortable problems soon arose from this lack of diversity in beliefs, both theoretically and empirically. To begin, a series of theoretical results in the early 1980s implied necessity of some disagreement (i.e., diversity in beliefs) in order to explain reality. For instance, Grossman and Stiglitz (1980) effectively demonstrated necessity of some 'noise' (and absurdity of perfect informational efficiency) in financial markets, whereas several 'no-trade' theorems (Milgrom and Stokey 1982; Tirole 1982) illuminated the point that, under common knowledge of beliefs, market participants should not engage in trade with one another.[32] But trade, and substantial amounts of it, characterise real-world markets, which hints that consensus and homogeneity of beliefs seems not to represent a suitable description for the usual affairs in markets.

Moreover, during the same timeframe, novel econometric analyses of price movements (beginning most notably with renowned tests by Shiller (1981)) began to bolster the case that RE explanations of price dynamics were not in fact as plausible as might have been initially suspected (or hoped).[33] A doorway was opened for heterogeneous beliefs to re-enter economic theory about market behaviour under the banner of 'noise' (see the acclaimed piece by Black (1986) on causal roles for noise in markets). Some inventive economists embedded this idea of perpetual noise in markets into theories and models of price behaviour driven by a fraction of market participants exhibiting systematic biases in their reasoning and beliefs; these systematically-biased agents were termed 'noise traders' (see especially: De Long et al. (1990a; 1990b); Shleifer and Summers (1990)). Coupled with demonstrations that opportunities for arbitrage in markets might be serially-limited

[32] These theorems have also playfully been termed 'Groucho Marx' theorems in reference to that famous actor's quip that he was suspicious, and desired not to be a part, of any club that would be willing to accept him among its members.
[33] Similar attacks on informational efficiency (primarily through demonstrations of the non-random, predictable nature) of features of asset prices also surfaced and helped erode the hegemony of RE (see Lo and MacKinley (1999)).

(Shleifer and Vishny 1997), the noise-trader approach in finance opened floodgates for incorporating findings from the HB programme into economic research about financial markets. That is, the noise traders in many behavioural finance models were specified to capture one or two of the biases identified in HB research by psychologists, and then contrasted against their unbiased ('smart-money') counterparts in models that allowed for the beliefs of both types to then compete.

Some problems surfaced for this approach, however. As Gennaioli and Shleifer (2010) observe: "models of heuristics have proved elusive" (p. 1399). That is, much of the work in psychology under the HB programme did not outright formalise the heuristics which give rise to biases as mathematical models, but instead sought to characterise them in ways that, while empirically convincing for many within the emerging behavioural finance and economics research paradigms, were difficult to implement as formal models capable of precise economic predictions. As a consequence, many of the proposed models in behavioural finance did not 'model' heuristics, but merely attempted to emulate biases that result from them, and then pointed to interpretational analogies to insist that the model could generate approximately-equivalent results to the intended heuristic 'as if' (using Friedman's (1953) sense) the model were of that heuristic more directly.[34]

What many of those models underappreciated, however, was the procedural natures of most heuristics. That is, heuristics are, in general, *processes* of thought utilised by human thinkers, and that such heuristic processes tend to exhibit substantial context- and path-dependencies. Failed sensitivity to this procedural aspect of heuristics caused a considerable amount of descriptive potency to be 'lost in translation' from the robust psychological findings to the mathematical attempts at recapturing those findings as models in behavioural finance. Additionally, this

[34] Many of these models are contained in Thaler (2005). For earlier work in behavioural finance, see Thaler (1993).

approach of endowing noise traders with only one or two biases at once ran into yet another difficulty which had previously been identified within the HB programme by psychologists: multiple heuristics are often deployed (and in competition with one another) at once within single individuals; and different individuals may employ different (but overlapping) combinations of heuristics. In market settings, therefore, the number of heuristics at play might be monumentally larger, and the standard tack of capturing all of these heuristics through a single class of biased noise trader was starting to become transparently deficient. In short, the diversity in this pioneering class of behavioural-finance models was becoming noticed as excessively-constrained in terms of the cognitive heterogeneity which they were aspiring to capture. The next stage for such approaches, however, seemed to be relatively clear: do as the psychologists had done and try to combine multiple heuristics and biases all at once into single models of bounded human judgment.

Among the clearest candidates for such models of combined biases and heuristics were prospect theory (PT) (Kahneman and Tversky 1979), which applies to the domain of given probabilities (risk), and its sibling relevant for the domain of uncertain probabilities (or non-additive "capacities"): cumulative prospect theory (CPT) (Tversky and Kahneman 1992).[35] But, despite attempts at applying CPT to financial modelling (see: Barberis and Huang (2008)), there have been serious challenges toward operationalising CPT in behavioural finance. As Barberis (2013b) notes:

[35] CPT generalises the class of so-called rank-dependent models of utility, which include, e.g., those due to Quiggin (1982), Segal (1987; 1989), and Yaari (1987) for situations concerning risk, and Schmeidler (1989) for situations that involve uncertainty. For a theory of uncertainty similar to CPT, consult the work due to Luce and Fishburn (1991). Furthermore, note: capacities are non-additive set functions that permit a general concept of probabilities; see, originally: Choquet (1955). (Gilboa (2009, p. 159) notices that rank-dependent models for utility are special cases of Choquet integration, in that they all entail the usage of unique probability distributions; 'multiple priors'-type models (e.g., Gilboa and Schmeidler (1989)), meanwhile, entail the coexistence of more than one (i.e., non-unique) probability distribution simultaneously.) Also recognise that the prospect theories and rank-dependent theories belong to the so-called *non-expected utility* family of theories about judgement. Other members of that family include (but are not limited to, and will not be discussed further here; see Thaler (2003) and Gilboa (2009) for references on the following): weighted-utility theory; implicit expected utility; disappointment aversion; regret theory.

> One might have thought that, more than 30 years after the publication of Kahneman and Tversky's (1979) paper on prospect theory, we would have a clear sense of how important a role their theory can play in economic analysis. This is not the case. Because of the difficulties inherent in applying prospect theory in economics, it is only in the last few years that we have made real progress in doing so. Indeed, this research effort is still in its early stages. (p. 191)

The "difficulties" noted by Barberis are not encouraging because (as the migration of Markowitz's ideas from given (1952a) to uncertain (1959) probabilities emphasised) in finance it seems that uncertainty is often a more applicable concept than risk. And the appropriate version of prospect theory under conditions of uncertainty, cumulative prospect theory, is generally more difficult to work with than basic prospect theory due to its use of capacities and other complicating devices.[36]

But translating PT and CPT from psychological to some economic applications (especially financial markets) may be an errant mission in several respects. To begin, PT and CPT both presume degrees of stability in judgment that seem unreasonable when compared with reality (Shleifer 2012). Other approaches, such as salience theory (Bordalo, Gennaioli, and Shleifer 2012) have begun to emerge which are better able to capture instability and contextual dependencies in ways that PT or CPT are unable to do.[37] But, at present, most such alternatives generally remain applicable only in the domain of risk – with given probabilities – rather than being usable under conditions of uncertainty. Nonetheless, many of these emerging approaches continue to capture a feature of belief distortion that is central to both PT and CPT: the notion of probability weighting.

## **2.5**\\ Distorted Beliefs

A key consideration in both PT and CPT is that people can distort their beliefs (when expressed in the form of probabilities) in response to their attitudes toward either risk or uncertainty.[38] That is,

[36] Tversky and Kahneman (1992) themselves recognise some of these difficulties (see, e.g., their page 311).
[37] The theory of "decision by sampling" due to Stewart, Chater, and Brown (2006) is another such alternative approach. It is discussed along with other psychological theories of cognitive sampling and adaptive thought later in this chapter. See also: Stewart, Reimers, and Harris (2014).
[38] It stands difficult to mention prospect theory (whether the original or cumulative rendition) without acknowledging some of its other major contributions to advancing economic (or at least behavioural economic) thinking on cognition

according to PT and CPT, most people tend not to (unless they are neutral to risk and uncertainty) treat probabilities linearly, but treat such values as more, or less, than they mathematically should. Distortions of probability judgments by decision-makers as a reaction to (anticipated or experienced) uncertainty and risk was not a concept newly-developed by PT or CPT, but had been earlier proposed elsewhere in psychology and economics (see, e.g.: Preston and Baratta (1948); Mosteller and Nogee (1951); Fellner (1961); Handa (1977); Karmarker (1978)). PT and CPT were novel, meanwhile, for jointly explaining such distortions – or 'weightings' – of probability as being driven by perceptual limits (they prefer the term "diminishing sensitivity") *and* decision-makers' attitudes, as well as deriving specific weighting functions which can be parameterised and tested.[39]

Yet probability weighting functions (PWFs), and the decision weights (distortions of probabilities) which they generate, seem to amass some perplexity when it comes to psychological interpretation. In both papers on prospect theory Kahneman and Tversky (and in other papers thereafter wherein Tversky was a co-author) are at pains to ensure that readers do not misunderstand decision weights as a person's beliefs proper, but rather as transformations of such beliefs (see also Barberis (2013a)).[40] This stressed distinction seems to have achieved its desired effect; many contemporary

about risk and uncertainty. A complete dissection of PT/CPT is avoided here, but it should be mentioned in passing that perhaps the strongest contributions of PT and CPT were to *condense* many disparate ideas that were already familiar to economists, and (for some peculiar reason) more palatable collectively than separately. Consider, for example, the fact that the proposal by PT/CPT that decision-makers tend to emphasise changes to utility (or wealth) more strongly than levels of wealth (i.e., they focus on relative gains and losses more than absolute utility or wealth levels) was actually first introduced by Markowitz (1952b) (Wakker and Tversky 1993) but was little-used before PT.

[39] It is notable that the form of such weighting functions for distorting probability have empirically been found to exhibit a characteristic 'inverse-S-shape' with both concave and convex portions over the unit interval. A similar shape was initially proposed by Friedman and Savage (1948) and Markowitz (1952b) as an explanation for why people might engage in both purchasing of insurance and gambling (i.e., be simultaneously risk-avoiding and risk-seeking). As Tversky and Wakker (1995) point out, however, utility proved an inadequate explanation for this combination of seemingly contradictory behaviours, and nonlinear probability distortions then became a more preferable explanation.

[40] Contrastingly, Karmarker's (1978) approach to probability distortion was careful to note that such transformation was generally due to constraints on "information processing performance". While that view might be criticised for not placing more emphasis on the influence of attitudes to risk (preferences), it seems to move in the opposite cautionary direction as prospect theory in its justification for the deviation from more pure-form probabilities in human judgment. One should, nonetheless, notice the statement by Kahneman and Tversky (1979, p. 281) wherein they assert that both preferences and beliefs may serve to distort probability perceptions by decision-makers when *rare events* are involved.

authors still feel compulsion to enunciate that decision weights are not the same things as beliefs, even when the two might coincide (i.e., when decision weights do not distort a decision-makers beliefs). Confusion appears to arise, however, as to what psychological forces actually influence decision weights. Specifically, admonitions about need to distinguish between beliefs and decision weights seem to have had the negative impact of discouraging (especially among economists) considerations about decisions weights as being influenced from misperceptions of probabilities. That is, many economists continue to understand decision weights as transforming probabilities solely on the basis of decision-makers' preferences, and not due to any traceable properties of beliefs themselves.[41] Hence, decision weights are commonly seen as due (potentially) to attitudes about risk, uncertainty, or ambiguity, but not as partial functions of the misestimates of probability.

In a paper (which he began with Amos Tversky before Tversky's death in 1995), Wakker (2004), however, advocates the possibility (especially under conditions of uncertainty) that decision weights may distort probabilities due to influences from *both* preferences *and* belief-based components of resource-constrained cognition. Wakker suggests a decomposition technique for separating the influences of preference and belief on decision weights (while courteously noting that the germ for such technique was first suggested by Fellner (1961), and that an alternative method was devised by Epstein and Zhang (1999)), and further speculates that the scholarly emphasis on describing the phenomenon of ambiguity aversion highlighted by Ellsberg (1961) may have detracted from focus on the problem of distinguishing the belief and preference components of probability distortions. Nonetheless, Wakker's ideas seem to remain thinly-noticed in the economics community. Yet oversight might not be a main cause for this neglect. As Barberis

[41] For example, this point goes unmentioned in the relevant reviews on joint empirical and theoretical developments in economics and psychology (i.e., behavioural economics) by Rabin (1998) and DellaVigna (2009). And it is mentioned only tangentially in the review by Starmer (2000) in discussion of the so-called 'longshot bias' in gambling.

(2013a) states, "the distinction between beliefs and preferences is important: overestimation is a mistake; it is less clear that overweighting is a mistake" (p. 614). Such a perspective is characteristic of economic treatments that take decision weights as influenced only be preferences and encapsulates the more general stance in economics of greater acceptance towards the ranges of preferences that a decision-maker may reasonably hold relative to the range of allowable beliefs.

Nonetheless, a few economists (and psychologists bold enough to engage in economic research), have taken notice of the need to separate the influences of belief and preference on probability distortion (see especially: Kilka and Weber (2001); Snowberg and Wolfers (2010); Sydnor (2010); Barseghyan et al. (2013a; 2013b)). One intriguing stream of research on distorted probabilities is the optimal expectations (OE) programme (Brunnermeier and Parker 2005; Gollier 2005; Brunnermeier, Gollier, and Parker 2007). OE thinking posits that decision-makers derive utility from optimism as well as the ultimate accuracy of their beliefs (or, rather, the benefits that such correctness may bring); and therefore decision-makers 'select' distortions in their beliefs that strike an optimal balance between the veracity of their expectations, and the interim pleasure derived from having an optimistic outlook. While innovative, the OE perspective seems to suffer from some shortcomings: 1) it is decidedly asymmetric and does not provide a satisfying explanation for why people might engage in pessimistic distortions of their beliefs (which they arguable do); and 2) it requires even greater cognitive powers for decision-makers than without belief distortion.

Thus, although the OE view may be an attempt at appeasing economists unwilling to believe that people carry serially-flawed beliefs, it does not appear to represent any genuine resolution. More cognitively realistic treatments of probability misperception seem to arise from prospective reference theory (Viscusi 1989), although some tests of that approach have been discouraging (see Harless (1993)). Gayer (2010) offers a compelling recent approach rooted in similarity perception.

But just because much of (behavioural) economics has not been careful to separate the effects of beliefs from preferences in probability weighting does not mean that it has been empirically, or theoretically, disinterested in characterising probability weighting functions. For example, efforts by economists and psychologists alike have sought to parameterise PWFs and furnish psychological interpretations for the relevant parameters (see: Tversky and Wakker (1995); Wu and Gonzalez (1996; 1999); Prelec (1998); Gonzalez and Wu (1999); but note that those treatments apply variously to situations of risk and uncertainty, and that they focus heavily on preference-centric interpretations). The main research, however, tends to accentuate two forms of probability distortion: *curvature* (nonlinearity in probabilities), and *elevation* (shifts upward or downward when 'true' probabilities are plotted against distorted probabilities; this operation is generally attributed to attitudes toward risk or uncertainty). Other transformations, such as shifts, are usually not discussed.[42] Of added importance is a characteristic finding in such research that much intersubjective (and even intra-subjective) heterogeneity appears to exist among best-fit PWFs.[43]

This heterogeneity in probability weighting is likewise noted by the authors in both presentations of prospect theory, and serves to underscore yet another service performed by the introduction of those theories into economics: highlighting of the role of relativity in human decision-making.[44] This relativity is deeply embedded in prospect theory in the devices of reference points and frames. Both reference points and frames can be construed as forms of contextualisation of the problem facing a given decision-maker. PT and CPT both utilise a two-stage framework, and in each theory the first stage ("editing" in PT, and "framing" in CPT) involves problem-contextualisation that

[42] Shifts, or slanting (also *translation*), of probabilities was nonetheless suggested many years back by Fellner (1961). (Uneasy dispute, such as is evident in Brewer and Fellner (1965), suggests why slanting fell out of favour in economics.) Shifts may be interpreted as related to elements of confidence in judgments, which will be treated shortly.
[43] On heterogeneity in PWFs, see recently Bruhin, Fehr-Duda, and Epper (2010), who adopt a preference-centric view.
[44] On diversity in the probability weighting functions, see especially pages 306-307 in Tversky and Kahneman (1992).

relies extensively on both frames and reference points. Important to note, however, is that neither PT nor CPT explicitly formalises this first contextualisation stage. Moreover, subsequent work in behavioural economics has encountered severe difficulties in formally modelling the crucial roles of frames and references points that are entailed in this contextualisation process in both theories.

## **2.6**\\ References and Frames as Attention and Memory

Also, a point that was not substantially stressed in the original papers on either PT or CPT is the extensive potential for interdependence of frames and reference points (although this lack of emphasis is perhaps unsurprising in light of how many elements of PT are taken as being given). That is, both PT and CPT assert that decision-makers derive value (utility) from not only the absolute levels of some good or service which they consume, but also relative changes (specifically: gains and losses). Yet these changes are 'framed' with respect to some reference point, which may be the *status quo*, or some other expected value; but the pertinent reference points may be likewise evoked by the way in which the choice alternatives are presented to, or formulated by, the decision-maker. This entangling presents obvious difficulties in terms of formal modelling, as it engenders various instabilities due to whether a frame or reference point gets established first. That is, the reference points and frames ultimately utilised by a decision-maker in her judgment in some specific problems may differ depending on whether a frame or reference point is fixed first.

And while frames and reference points may both be exogenously bestowed upon decision-makers in a limited range of settings, more realistic scenarios are those in which frames and reference points become driven at least partly by decision-makers' own beliefs (this point was alluded to, but not fully-elaborated on, in both original presentations of PT and CPT). It is worth briefly considering how the promotion of the importance of reference points and frames by PT and CPT has impacted general conceptualisations of beliefs in economics (although shortcomings remain).

Beginning with reference points, it should be mentioned that there are two general types of reference points in both PT and CPT: those which apply to utility (the value function); and those which apply to probability (specifically the probability weighting function that distorts probabilities). The first target of concern here is the latter class of reference points, although the former will be also examined shortly. Both PT and CPT assert that the probability scale contains two natural and obvious reference points: zero and one, which amount to impossibility and certainty, respectively. A general psychological construct invoked by both PT and CPT is that of *diminishing sensitivity*, which stipulates that changes are more salient and potent nearer to a reference point than further away from it. Consequently both PT and CPT take probability weighting to distort probabilities more strongly near the endpoints of zero and one (i.e., small and large probabilities) than in the middle of the probability scale (i.e., near one-half). And while empirical results generally confirm this relative distortion (although with substantial inter- and intra-subjective variability), a rather perplexing finding emerges as part of both PT and CPT: probability distortion is 'misbehaved' very close to both endpoints of zero and one (see specifically pages 282-283 in Kahneman and Tversky (1979) and page 303 in Tversky and Kahneman (1992); see also Quiggin (1982), who also discusses the problem of instability at probabilistic extremes).[45]

That the relevant reference points for decision weights may be unstable is troubling owing to a general finding that decision-makers are typically highly-sensitive to shifts in reference points. But the instability of reference points also has bearing in the second class of reference points found in PT and CPT, which typically has a stronger level of direct connection with framing: reference points in value or utility (indirectly, framing may also influence probability reference points

[45] Notice, however, that it is the inclusion of an explicit role for reference points which distinguishes CPT (and PT) from other rank-dependent utility models, such as the theory of anticipated utility that was put forth by Quiggin (1982).

through the generation of evidence from memory; this point will be discussed later in this chapter). As was mentioned, this form of reference point may rely either on *status quo* values, or else other beliefs about value. Situations in which the *status quo* serves as the main point of reference have perhaps received more (and more high-profile) treatment in behavioural economics (see, e.g.: Samuelson and Zeckhauser (1988); Kahneman, Knetsch, and Thaler (1991); Fernandez and Rodrik (1991); Masatlioglu and Ok (2005)). Nonetheless, the selection of the *status quo* can sometimes be unstable for being arbitrarily-influenced if it is selected with respect to some anchor prior to complete formulation of the choice alternatives (or else before complete generation of the frame; such a situation might occur, for example, in bidding situations); see particularly the innovative experiments by Ariely, Loewenstein, and Prelec (2003; 2006).[46] Recent and promising work on expectation-based reference points has emerged, however (see: Koszegi and Rabin (2006; 2007)).

If reference points for value and utility do rely to some extent (or at least in some important classes of situations) on beliefs (as mounting evidence suggests), then several concerns immediately arise. First, as noted in Fehr-Duda and Epper (2012), belief-dependent reference points for preferences (in the forms of decision weights) may help resolve many longstanding puzzles in financial economics. Second, belief-dependent reference points for preferences suggest that experience should be potentially considered a stronger determinant of behaviour, and differences in experience might then serve to generate divergences among decision-makers' beliefs, preferences, and judgments (see.: Greenwood and Nagel (2009); Malmendier and Nagel (2011; forthcoming)).

And, similarly, differences in the attentional resources of decision-makers might also generate divergences in how judgment and choice situations are framed. Limited attention (for example, in

[46] See also: Stewart, Chater, and Brown (2006); Fox and Tannenbaum (2011); Bordalo, Gennaioli, and Shleifer (2012); Stewart, Reimers, and Harris (2014).

terms of: the ability to detect differences in alternatives; or restrictions on the number of alternatives that can be called to and manipulated in mind at once) and framing are innately connected; framing occurs when a decision-maker considers only a restricted set of alternatives, and possibly even considers that set of alternatives in a restricted way. These restrictions cause some possibilities to be neglected, and thus not attended to. Of course, limited attention may be both a property of the decision environment (with, e.g., restricted informational flows) or the decision-maker (and thus a cognitive or motivational limitation). Economics has traditionally focused on the former type of informational restriction, and, when modelling bounded cognition, preferred to limit either the availability of information to (some) decision-makers, or else their abilities to process such information. Making limited attention restrict inputted information, however, has become increasingly popular as a research topic in behavioural economics lately, and can produce distinct results (and theoretical advantages) over casting limited attention as either a restriction on the decision-maker's environment, or else here informational-processing abilities. This recent view of limited attention as a cognitive restriction on inputs means decision-makers may have informational *available* to them that they cannot *access* in some way, even though they might be able to process it if it were accessible. In short, some possibilities can persist unnoticed.

Many noteworthy attempts at capturing this new interpretation of limited attention have emerged in behavioural economics, and there are too many to mention in detail here (although more coverage will be given in later chapters to select theories and approaches). But some of the more promising formulations are worth listing: rational inattention (Sims 2003); local thinking (Gennaioli and Shleifer 2010); revealed attention (Masatlioglu, Nakajima, and Ozbay 2012); salience theory (Bordalo, Gennaioli, and Shleifer 2012); sparsity-based cognition (Gabaix 2014).[47]

[47] Clark (2011) also presents a non-formalised treatment of 'myopia', which entails context-specific reasoning forms.

Furthermore, it should perhaps be obvious that limited attention also bears on the operations of search, and this new paradigm of limited attention has also begun pursuit of implications for attentional restrictions on the search process (see especially: Caplin and Dean (2011); Caplin, Dean, and Martin (2011)). A standing shortcoming of these approaches, however, is that they generally take the process of search under limited attention as exogenously-driven; that is, they standardly utilise objective (exogenous) characteristics of alternatives in some *given* search context as an ordering device for the process, and often lack a role for endogenous drivers (Bordalo, Gennaioli, and Shleifer 2012). Put differently, the unions between limited attention (as a restriction on cognitive inputs) and search that have emerged in the behavioural economics literature thus far are predominantly passive, and do not (in general) sufficiently accentuate the role of guided search driven by *ex ante* beliefs, which are in turn fuelled (in many settings) by evidence from memory.[48]

The role of memory and recalled evidence in connection with beliefs will be discussed at length soon. But first more must be said about another important type of framing effect that has been shown to be considerably important to the instability of judgments: comparative ignorance. The comparative ignorance hypothesis (CIH) (Fox and Tversky 1995) was born from the competence effect (Tversky and Heath 1991), which in turn was a critique of ambiguity aversion (demonstrated earlier in economic research by Ellsberg (1961))[49]. Succinctly, the CIH claims, based on findings from experimental investigations, that ambiguity aversion (or the preference for concrete rather than vague specifications of outcomes and likelihoods of those outcomes) is comparative on at least two dimensions: the alternatives faced; and the relative knowledgeability of other decision-makers. That is, the CIH asserts that ambiguity aversion is substantially reduced when a vague

[48] A recent exception (depending on leniency in one's definitions) is Hanna, Mullainathan, and Schwartsztein (2014).
[49] See also: Fox and Levav (2000); Fox and Weber (2002).

gamble is evaluated in isolation, as compared to when it is presented alongside a less vague alternative. Likewise, decision-makers become less aversive towards ambiguity when they consider themselves more knowledgeable than other decision-makers who will be facing the same choice, or evaluating the choice that they make; but decision-makers generally become more ambiguity averse if they consider themselves less knowledgeable than other relevant decision-makers.[50] Thus, both ambiguity in beliefs and the credibility of others' beliefs may impact a decision-makers' judgments in a fashion that makes her susceptible to two dimensions of framing. More will be said shortly on both ambiguity and other-regarding beliefs, but for now the focus returns to the roles of evidence and memory as determinants of the beliefs held by decision-makers.

It seems relatively surprising that, although economics has historically been so preoccupied with the topic of beliefs, the formulation process for 'prior beliefs' remains such an under-developed idea in economic thought. That is, a great many theories in economics take as a starting point that individuals *have* prior beliefs, but fail to specify more precisely the origins of those prior beliefs (in this sense 'prior' is relative to receiving further information or engaging in some economic activity). Psychology, meanwhile, seems to have paid far more attention to the origins of beliefs, and has generated some intriguing theories about the role which evidence has in belief generation.

[50] Güney and Newell (2014) present evidence that ambiguity aversion may dissipate with experience and familiarity. Further, Rabin and Schrag (1999) suggest ambiguity and learning may have a peculiar relationship in their model of *confirmatory bias*. They show that, given some ambiguous evidence, decision-makers may oftentimes be prone to misinterpreting such vague evidence incorrectly and in ways that support their favoured hypotheses or initial impressions. This confirmatory bias then leads to a circularity of overconfidence if additional ambiguous evidence is also misinterpreted. An ultimate implication of such possibility is that learning may sometimes fail to correct beliefs which are initially erroneous. Such flaws in learning run contra to conventional Bayesian wisdom on belief dynamics, and may promote further interesting phenomena when one moves into more realistic realms of decision-makers actively seeking evidence. For example, if decision-makers actively search for evidence, and encounter evidence that could be ambiguous, but which they wrongly interpret as confirmatory, then they may later seek out other evidence that does not push them in the direction of truth. Such circumstances would make more subjective the notion of "success" in environmental searches based upon learning cue-criterion relationships, as in (e.g.) Newell et al. (2004).

Support theory (Tversky and Koehler 1994; Rottenstreich and Tversky 1997) remains perhaps one of the more renowned theories of how uncertain beliefs (and probabilities) are generated from evidence. Support theory will be explored in detail in later chapters, but for now realise that it says nothing about the connection between the evidence which gives rise to subjective beliefs, and from whence such evidence comes. Essentially, support theory fails to stipulate how evidence is drawn from both current environment as well as memory. Yet other realms of psychology have been immensely preoccupied with the link between evidence, the storage and organisation of evidence in memory, and the role of memory structure in cognition. The works of Herbert Simon in particular aimed to study the influence of memory structure on cognition in meticulous detail (see especially Simon (1979; 1996)). Simon's work uncovered some rather remarkable properties of human memory, and his findings suggested that the idea of restrictions on the 'processing' capabilities of the human mind may in fact emanate to a significant extent from the specific structure of human memory.[51] For example, Simon's work with cognition in chess (see, e.g.: Chase and Simon (1973); Simon and Chase (1973); Simon and Gilmartin (1973); Gobet and Simon (1996a; 1996b)) revealed that grandmasters and novices do not generally possess remarkable differences in working-memory span (i.e., number of items that can be held in mind simultaneously; see Miller (1956)) or computational ability (in terms of the number of different move combinations they could explore). Instead, Simon's work revealed that the grandmasters carried an edge over novices in what combinations of (sequences of) moves they could recall and recognise from memory. And this recall and recognition was shown to have been facilitated by greater degrees of organisation and structure (as well as content) in the memories of chess experts

[51] Recent work in psychology has also produced some impressive models rooted in memory structure. See: MINERVA-DM (Dougherty, Gettys and Ogden 1999); CHREST (Gobet et al. 2001). That processing limitations may be derived from memory structure therefore suggests that such limitations may actually be forms of input limitations.

relative to less-experienced players. This memory-structure perspective on expertise has also been shown to carry over to other domains outside of chess, and has become a well-accepted criterion for cognitive expertise within a great many fields, including medicine, mathematics, music, science, navigation, and many others (Ericsson et al. 2006). This finding that memory structure influences retrieval of evidence therefore suggests that memory structure must help dictate beliefs, if beliefs rely at all on evidence drawn from memory (which they appear to regularly do). Furthermore, if attention can be guided by beliefs, then it must somehow be influenced by memory.

It is interesting to note that the psychological and economic traditions have often taken opposing levels of interest in the role of memory structure. One retrospective article by Luce (2003) suggests that the information-processing paradigm for studying cognition, which has taken such strong root in economics, has failed to become as deeply engrained in psychology partly *because* the highly-structured nature of human memory frustrates many of the requirements of an unordered space (similar to the 'exchangeability' conditions developed by Bruno De Finetti, which remain vital to the Bayesian approach to subjective probability; see Kyburg and Smokler (1964) and the papers therein). (For additional, classic references on the structure of human memory and its impact on cognition and judgment, see: Tulving and Donaldson (1972); Luce (1986)). Yet, until recently, economics seems to have stood reticent to grant that human memory's structure matters significantly as a determinant for the outcomes of judgment (Hayek (1952) is a notable exception).

Nevertheless, lately there seems to be strengthening surges in interest among economists in (modest) considerations for memory structure's role as a rudder for beliefs (see: Rubinstein (1998a); Gilboa and Schmeidler (2001); Mullainathan (2002); Aragones et al. (2005) Gennaioli and Shleifer (2010)). These developments in consideration for memory structure, albeit substantially fledgling in sophistication when contrasted against the richness of explorations of

memory structure undertaken in psychology, are still permitting powerful advances in economic theory. For one, they have both inspired and enabled (and, reciprocally, been inspired and enabled by) work on ideas of *similarity* as a foundation for beliefs and judgments; see: Rubinstein (1998b); Billot et al. (2005); Billot, Gilboa, and Schmeidler (2008); Gilboa, Lieberman, and Schmeidler (2006; 2010; 2011).[52] Surprisingly, this increased interest at the forefront of economic research in structure of memory and its impact on beliefs may be moving in opposing direction to interests in psychology within various areas, including cognitive sampling and the experience-description gap.

## **2.7**\\ Sampling and Gaps

The cognitive-sampling paradigm was accelerated by an insightful article of Fiedler's (2000) which offers an imaginative explanation for the genesis of many human biases as stemming from the "sampling" processes which people use in jointly probing both the external environment and their own minds. Fiedler's provocative idea echoes Simon's (1990) metaphor of human behaviour as residing at the nexus of two blades in a scissors, which respectively represent environment and cognition. Comparably, Fiedler's (2000, p. 661) rendition asserts:

> As a matter of principle, samples result from the dynamic interaction of the person and the environment. Sampling is not fully under the individual's cognitive control but is greatly affected by the distribution of information in the environment. Conversely, the inclusion of environmental stimuli in an effective sample depends on the individual's attention, capacity and processing goals. Within such a genuinely cognitive-ecological approach, neither the individual nor the external world can be blamed for the emergent sampling biases; they arise as a natural consequence of human-environmental interaction.

The sampling approach's emphasis on environmental-cognitive interplay as governing bias is powerfully appealing, and otherwise difficult to argue against. Such appeal seems to explain the substantial uptake of this idea, which became enshrined in the resulting metaphor of human judgment as resembling the operations of a "naïve intuitive statistician" (Fiedler and Juslin 2006).

[52] It is disappointing, nonetheless, to see that ideas on *asymmetric* similarity (initiated by Tvesky (1977)) have yet to permeate economic interest in similarity to a significant extent, despite application in fields like artificial intelligence.

At first glance, Simon's scissors metaphor and the naïve intuitive statistician seem to be one and the same, in light of Fiedler's explanation of such a statistician operating at the nexus of mind and environment. Yet one must be careful to parse the differences embedded in each metaphor. Simon's approach characteristically prioritised the role of *search* in both the external realm of environment and the internal realm of the mind (although he does stress that, to some degree, the distinctions between the two are less pronounced than they may seem, and that they are to some extent both mental recreations of one another from the perspective of the decision-maker; see Simon (1996)). While both search and sampling are procedural, and can sometimes be treated synonymously, they can also diverge meaningfully from one another, and such potential divergences largely revolve around the structure of the population of elements being searched or sampled. Simon's structural emphases tended to appreciate hierarchical richness in both external and internal fields of search (i.e., environment and mind, respectively), as well as the dynamical contingencies entailed in search (that is: the results of one step in search – whether parallel or sequential – can strongly influence other steps). Meanwhile, sampling approaches under the metaphor of the naïve intuitive statistician tend to focus on more distributional aspects of structure.

Focusing on the distributional aspects of structure tends to make sampling innately more bound to randomness than is search. That is, sampling relies on probabilistic draws from some specified distribution of possibilities, whereas search may rely on some stochastic influence, but may also (and typically seems to be) driven by deterministic rules. The extent of randomness in the sampling paradigm therefore relies on the structural richness of the distributions assumed (or empirically derived), which includes whether or not they are static. Particularly for cognitive sampling,

wherein evidence is extracted from memory, one must question whether it is reasonable to treat such a process as representable as series of random draws taken from an unchanging distribution.[53]

Many judgment situations do not seem amenable to such a representation. For one thing, one must consider that draws from memory should not be considered as being undertaken 'with replacement'; it stands highly suspect that people make repeated *and* identical calls to the same remembered evidence, especially if mental 'sampling' is sequential. But if sampling is without replacement, then the assumption of an unchanging distribution across draws becomes less tenable (and particularly so if the population of evidence from which draws occur is of rather limited size).

And many studies under the cognitive-sampling paradigm do recognise the unrealistic nature of some of the assumptions on randomness which they adopt, but insist upon their convenience as placeholders until more realistic assumptions can be made after the underlying phenomena are better comprehended (see, e.g., Stewart, Chater, and Brown (2006); Stewart (2009)). This practice reflects sound incremental science. Nonetheless, observing that the paradigm is now over a decade old, it could be expected that fewer invocations of randomness would be being made and that sampling would come to resemble deterministic search. Yet the sway of stochastic components does not seem to have subsided much in the sampling paradigm since its inception, which is worrying for the fact that it seems to be demoting the importance of some important vital structural influences on human judgment processes (especially in terms of the extent to which advances in the understanding of how memory structure may play unexpected roles via constructive thought).

[53] Hills and Hertwig (2010) find that search strategies in external environments are highly predictive of judgment outcomes and do not seem to be random; if parallels exist between internal and external sampling then their results may have import for the extent to which the usage of evidence taken from memory is describable as being 'sampling'.

Hence, the cognitive-sampling paradigm has, at its core, much potential for maturing the appreciation for how environment and memory interact to shape human judgments and choices, and some researchers allied with this paradigm have made significant strides in studying instabilities among people's subjective beliefs and probability judgments that stem from such interactions (see, e.g.: Ungemach, Stewart and Reimers (2011)). Nevertheless, caution should attend the assumptions about randomness embedded in sampling models, and efforts should be made to ensure that such randomness remains ancillary, rather than focal, in explaining judgments. That is, it seems imperative that randomness be permitted as a stopgap, and not a final explanation.

And a question germane to the mitigation of randomness in explaining human judgment seems to be the relationship between memory structure and the forms of mental representation of evidence. This question has surfaced in another psychological topic touched by the sampling paradigm: the study of the so-called 'description-experience gap' (DEG).[54] Awareness of the DEG grew out of empirical findings that putatively contradicted a 'received view' on probability weighting: people tend (*ceteris paribus*) to overweight small-probability (rare) events. While a great many studies support this tendency for circumstances wherein decision-makers have the relevant probabilities (or other likelihoods) given to them, it has been found to be regularly reversed for circumstances wherein decision-makers have to make judgments based on their own experiences: people have been shown to apparently *underweight* rare events when judgments are made from experience, but *overweight* rare events when making judgments based upon descriptions of rare-event likelihoods.

Debate about the robustness and causes of the DEG within psychology has been at times fierce, but highlights some key issues in the conceptualisation of the psychology of subjective probability.

[54] For a compact overview of the DEG literature, see Chapter 4.

For one, it prompted systematic assessment of whether the DEG was truly a phenomenon of weighting, or whether it was an issue caused by sampling biases (i.e., decision-makers act as if underweighting the rare events being studied because in experiments they often under-experience them) or estimation biases. To a large extent, it seems that, while the debate remains unsettled, sampling biases appear to be the leading candidate for explaining this phenomenon, and a substantial number of studies show that when more representative samples are experienced, the gap between rare-event judgments from description and experience typically lessens, although often does not disappear completely. Identification of the significant role of sampling bias accentuated the idea that people oftentimes engage in 'frugal search' and content themselves with relatively small samples. This contentment with small samples (which has been shown to obtain even when sampling is transparently costless in non-cognitive terms) has highlighted the general importance of limits on working memory, as well as the role of intersection between memory structure and mental representation of quantitative evidence. Early attempts at explaining the DEG generally disregarded effects and influences of memory structure (apart from exploring so-called 'recency effects', which have been shown to have occasional bearing on experiential judgments). And only recently has work on the DEG tended to be concerned with the structure of memory, and on how experiences extracted from memory aggregate to generate judgments for rare events.

Methodologically, the trajectory of research on the DEG is perhaps unique because it was influenced by the results of competitions that were hosted by interested researchers, wherein contestants attempted to construct models that could explain data pertinent to the DEG which had been agreed upon in advance by panels of experts.[55] It seems to have been this meritocratic approach which allowed exemplar-based learning models to become so accepted as among the

[55] See results of these competitions in a special volume of: *Journal of Behavioral Decision Making* (Volume 23, 2010).

leading candidates for the DEG: these models, which capture judgments as stemming from combinations of instances retrieved from memory (versus non-exemplar models, which, when applied to experience-based judgments, take judgment as being based on some 'aggregate statistic' of experience which is updated directly to reflect learning and ongoing environmental sampling), were among the best performers in these scholarly competitions surrounding DEG phenomena. These exemplar-based explanations posit that biases in judgment on rare events emanate substantially from how evidence is extracted from memory and subsequently mentally represented.

Findings on the DEG (since it has begun to take memory structure more seriously as a causal force) harbour some profound implications for study of formation and evolution of human beliefs about rare events. According to prospect theory (and other variants of non-expected utility theories which use nonlinear probability weighting), people over-accentuate rare events. But if this over-accentuation is universal, then how can so many rare events catch people unaware? For example, the recent global financial crisis which started to unfurl in 2007 should have been more prominent in people's minds as a possible threat, rather than being construed as a surprise by so many. Some theorists have proposed that such events are effectively "Black Swans" (Taleb 2007), and that they are not considered as possibilities by people *ex ante* (i.e., they are 'neglected' entirely; see, e.g.: Gennaioli, Shleifer, and Vishny (2012)). Yet while such utter neglect may be true for some events, it does not seem plausible that people always fail to consider rare events; instead, many rare threats may simply be considered, but then disregarded thereafter for becoming under-accentuated.[56] The DEG seems to offer the possibility for supplying a cognitive foundation for beliefs about rare events, and it might become usefully integrated into nascent work on rare disasters in financial

[56] See Chapter 3.

economics (which does not yet embrace psychological variability as influencing market reactions to threats by rare disasters; see: Barro (2006); Weitzman (2007); Gabaix (2012); Wachter (2013)).

Still, despite the promise of findings from research on the DEG as explanations of other phenomena related to human beliefs, much remains to be resolved regarding the DEG itself. In particular, how exactly people mentally represent quantitative information and evidence remains an open question. One candidate explanation that has so far gone unmentioned in published research is the role of *approximate number systems* (ANSs) in cognition about rare events. ANSs are covered extensively in the later chapters of this thesis, and therefore it suffices here to mention that these cognitive systems are mechanisms for the 'noisy' (but not necessarily random) retrieval of numerical evidence from memory and for its subsequent, approximated representation mentally.

Moreover, much research on human cognition demonstrates its socially-comparative nature (e.g.: the previously-discussed comparative ignorance hypothesis of Fox and Tversky (1995)); people's beliefs are often sizably swayed by what they estimate (or are told) other people believe.[57] Consensus, or perception thereof, therefore may strongly steer what people believe about rare and frequent events alike. Much more will be said on consensus in subsequent chapters, so not much needs to be conveyed here apart from acknowledging that DEG research ignores consensus effects.

Another present shortcoming of treatments of the DEG are their reductive approach toward the role of confidence in dictating beliefs about rare events (from experience or description, but primarily the former). Essentially, the vast majority of treatments (at least implicitly) presume that people can represent their beliefs about the rare event(s) in question as single point estimates (from

[57] While the interactive nature of thinking and strategy has been embraced by economists for some time (see, notably: Schelling (1968)), the cognitive foundations and occasional instability of this other-regarding thought seems to have been less well-studied. For example, the possibility that people sometimes over-infer from others' behaviour, but at other times under-infer from such behaviours remains a phenomenon yet to be comfortably reconciled within economic theory about interactive behaviours (compare, e.g., Eyster and Rabin (2010) with Eyster and Rabin (2005)).

unique probability distributions) of subjective probability. But this approach ostensibly could be a grossly incomplete reflection of the influence of people's confidence upon their subjective beliefs. That is, variable levels of confidence may cause a decision-makers' beliefs (formulated as probability estimates) to be greater or less than they should otherwise be under different alternative levels. Thus, changes in confidence amount to movements along a single dimension: the probability scale. This concept conforms to SEU reasoning that all relevant evidence should be embedded in a unique probability estimate (and presumably confidence, or lack thereof, equates to some form of subjective evidence).[58] Under situations of ambiguity this reasoning may encounter descriptive difficulties, because (as Ellsberg (1961) showed) people may resist, or even be incapable of, formulating beliefs as single probabilities if many interpretations are all plausible.

Presumably, in many of the experimental DGE setups in which people learned about rare-event likelihoods through 'sampling' procedures (again, whether such procedures are better understood as *search* is a potential point of contention), they may have felt that more than one distribution could account for the evidence which they experienced. And an inability to generate a probability judgment that could reconcile all plausible distributions would give rise to a situation of ambiguity.

## **2.8**\\ Multiple Beliefs and Confidence

In some such instances it may then become more appropriate for a decision-maker to act as if she entertained a set of subjective beliefs for the target hypothesis (i.e., a range of probabilities) rather

[58] An alternative to the single-dimensional approach to confidence is the hierarchical (or meta-belief) approach in which decision makers' wield first-order beliefs, but then have beliefs about these beliefs (second-order beliefs), and beliefs about their beliefs about their beliefs, etc. Such hierarchies may also be intersubjective (or hetero-hierarchical, rather than intra-subjective/homo-hierarchical; see Walliser (1998)), and describe individuals' beliefs about other individuals. But hierarchical beliefs both engender difficulties in implementation (e.g., questions of truncation), as well as inspire doubt as to their psychological realism (beyond only a very few levels). Belief hierarchies will be discussed in more detail in Chapter 6, as will the notion of cognitive hierarchies (for now, consider the reviews of these concepts covered in the following: Bonanno and Battigalli (1999); Crawford, Costa-Gomes, and Iriberri (2013)).

than a single point-estimate belief; for how (barring any further evidence becoming accessible to her) should such a decision-maker reasonably (i.e., non-arbitrarily) choose among possible beliefs?

Savage (1954, p. 57-58) realises that the requirement for decision-makers to always formulate their beliefs as points rather than ranges of probability may be unsatisfactory under variable confidence:

> …there seem to be probability relations about which we feel relatively "sure" as compared with others…The notion of "sure" and "unsure" introduced here is vague, and my complaint is precisely that neither the theory of personal probability as it is developed in this book, nor any other device known to me renders this notion less vague.

And while approaches to making more precise this variability in surety (or confidence stemming from ambiguity) were unknown to Savage for not having then been invented, many such theories have since been developed. Shafer's (1976) mathematical theory of evidence, which incorporates notions of 'upper' and 'lower' probabilities, is one such theory, although (per Shafer's own confession) it is mainly normative and not aimed at being psychologically accurate (which it fails to be).[59] Other approaches, however, do adopt the concept of sets or ranges of probabilities. Among such theories, the "multiple priors" (Gilboa and Schmeidler 1989) and "robust beliefs" (Hansen and Sargent 2001, 2008) approaches are popular within economics. Theoretically these theories are quite beautiful for their elegance, but, like Shafer's theory, they have very questionable validity for lacking psychological plausibility; in short, they seem to be too demanding of decision-making.

Moreover, the aforementioned approaches to non-unique probabilities say very little about belief formation: they are not particularly specific about the processes by which evidence is converted to beliefs (likewise these and related approaches that can accommodate ambiguity through multiple simultaneous beliefs are awkward dynamically: they are tricky to work with and often do poorly in reflecting learning dynamics, particularly when learning may be social; more will be said on

[59] See also the earlier work by Dempster (1967), as well as Suppes (1974) and Walley (1991).

these theories in later chapters). Having theories which directly translate properties of confidence in evidence into impacts on (potentially non-unique) subjective beliefs seems of fundamental importance to descriptively-accurate models for human judgment (even though more stable aspects of confidence that appear not directly related to evidence may often carry influence.[60] Indeed, Griffin and Tversky (1992) have demonstrated the robustness of a finding that people do take account for the volume and credibility (or "weight") of evidence, although they often do so to an unreasonable degree. Further, it is worth mentioning that the SEU/Bayesian approach makes no distinction between the "strength" of evidence expressed in a probability and the weight of evidence (e.g., if one takes subjective probability as a ratio of evidence for a proposition relative to overall relevant evidence, then SEU/Bayesian conventions do not (in static purposes) distinguish between 1/10 and 10/100). Yet the point must be stressed that the weight of evidence is a mostly relative construct psychologically (even if it fails to be so statistically), and whether samples of evidence are judged as being small or large may vary by context, or even across decision-makers. The next chapters of this thesis present a theory of responsive confidence in evidence which impacts (potentially non-unique) beliefs, and which can vary by context as well as interpersonally.

## **2.9**\\ Beliefs Elsewhere

An impression that one might glean from the literature on beliefs covered thus far is that a prevalent perspective within economics and psychology entails beliefs conventionally being either (or both): conceptualised as individualistic (as may be an enduring effect of Savage's (1954) intention for beliefs as "personal probabilities"); or appraised relative to some exogenous standard (i.e., beliefs are either 'correct' or else exhibiting various degrees of 'error'). This impression seems mostly

[60] Gender, e.g., has been shown as stably related to confidence (Barber and Odean 2001; Charness and Gneezy 2012).

accurate (apart from in social psychology, which will be covered in later chapters). But thinking from other realms of social science does not necessarily cast beliefs as being so strongly rooted in individualism, or necessarily locatable on a spectrum that spans from correct to shades of incorrect.

Indeed, in sociology and political science beliefs are far more tightly associated with concepts of identity. Such identity-centred perspectives oftentimes position beliefs as relatively more stable constructs than they are in economics and psychology, and associate beliefs with concepts such as political affiliation, cultural inheritance, and social stature. Within sociology in particular, beliefs as more permanent elements derived from (but reciprocally contributing to) identity are found in explanations of social structure as well as action (see, e.g.: Martin (2009; 2011)).[61] Such perspectives in some ways stretch the interpretation of beliefs, and mostly fall outside the definition of beliefs utilised in this thesis (i.e., beliefs as subjective expectations about the truth of some proposition).[62] Yet, the social orientation of such views will variously be drawn upon in later chapters, particularly as relates to the ideas of diversity. In fact, lack of substantial appreciation for diversity in beliefs is one of the more criticisable characteristics of economic treatments of belief.

Given this general suppression of belief heterogeneity in economics, it is altogether surprising that economic geography, which has been one of the most reliable sources of criticism of mainstream economic theory over recent decades, has failed to furnish any serious work of its own centred on the topic of beliefs. Indeed, the role of beliefs is rarely (if ever) dissected or questioned in any detail in the discipline of economic geography. An extensive review of the prominent literature in

[61] It is perhaps worth remarking that, while economics has been recently keen to embrace some ideas from psychology, it has been generally resistant to incorporate ideas from sociology (and even from social psychology). Cases of note include: the idea of 'embeddedness' which Granovetter (1985; 2005) has had difficulty transporting into economic work; and the role of social identity in decisions, which Akerlof and Kranton (2000) have encountered only limited success in weaving (in any appreciably-durable way) into mainstream economic theory and experiments.

[62] A notable exception in political science is the work done by Tetlock (2005) on expert political beliefs and judgment. Likewise, some of the work in political and social analysis by Pettit (2002) and Goldberg (2011) also merits exception.

the field, including its main texts (e.g., Clark, Feldman, and Gertler (2000)) and scholarly journals (namely: *Economic Geography* and the *Journal of Economic Geography*) failed to unearth any substantial treatment focused explicitly on beliefs' role in human judgment.[63] For the most part, beliefs in economic geography seem to go unmentioned, or else are handled in ways practically undistinguishable from mainstream economics (as is the case with the so-called *new economic geography*; see, e.g.: Fujita, Krugman, and Venables (1999)). Nor do more recent innovations in economic geography prioritise (or give justification for not prioritising) the status of beliefs in explaining or describing human actions (see, e.g.: Fingleton (2007); Bathelt and Glückler (2011)). This neglect of beliefs is surprising, and perhaps works such as the present one may supply remedy.

## **2.10**\\ Doxastic Desiderata

It is now time to take stock of the desiderata for the study of beliefs that have been exposed through this review. Aspiring to fulfil these criteria, which have proved to be shortcomings (singly or in conjunction with others) in earlier approaches to research on human beliefs, will serve as the key ambition for the doxastic approach to exploring beliefs that is contained in the following chapters. First, one should wish for the ability to capture significant belief diversity, but also allow for instances of consensus. Second, dynamics must be captured in ways that reflect various forms of belief change, whether classifiable as learning, or else various forms of instability. Third, one must allow for various forms of distortion in beliefs; it is preferable that distortions be traceable to either limited cognitive resources, or else the interference of preferences. Fourth, the ideal approach

[63] Indeed, a systematic review of both the aforementioned journals failed to uncover a single article in the entirety of the history of either that focuses in any primary way on the subject of beliefs. Instead, economic geography seems to focus more heavily on notions of acquisition and transmission of knowledge and information (and limitations thereon).

should capture social aspects of beliefs. Fifth, a suitable approach should make clear the bases upon which beliefs are formed and updated; that is, a central role for evidence should be apparent.

Capturing these desiderata should significantly aid the embedding of the three properties of beliefs identified in the introductory chapter – reflexivity, relationality, and relativity – into a robust and extensible framework for studying beliefs. It is to the formalisation of such framework that I turn.

# CHAPTER 3

# MINDING MAPS IN MAPPING (NEGLECTFUL) MINDS: MULTI-SCALE REFLEXIVITY IN NEURAL SPACE

## **3.1**\\ Introduction

Neuroeconomics studies people's decision-making by relating their choices to neural patterns through a blending of frameworks from neuroscience and economics (Camerer 2008). This young sub-discipline brings together a diverse array of researchers and techniques, which makes it (as its very name hints) a cradle for interdisciplinary fusions of ideas and methodologies. Apart from the obvious ranks of economics and neuroscience, neuroeconomics also draws its acolytes from psychology, sociology, mathematics, computer science, and philosophy (to cite just a few prominent influences). But, in reciprocal terms, neuroeconomics seems to be causing the loudest stir (among the traditional social science silos) within economics. Robert Shiller, a 2013 laureate of the Nobel Prize – primarily for contributions to empirical advances in behavioral finance – conveys his support for merging neuroeconomics into mainstream economic work (as well as captures the enthusiasm of some economists) with his remarks that:

> Economics is at the start of a revolution…Neuroscience – the science of how the brain, that physical organ inside one's head, really works – is beginning to change the way we think about how people make decisions. These findings will inevitably change the way we think about how economies function. In short, we are at the dawn of 'neuroeconomics'. (2011)

Positive sentiments about the prospects for neuroeconomics within wider economic theory and research are not, however, universally-shared among economists; Gul and Pesendorfer (2008) and Harrison (2008), for example, each sound well-cited oppositions to the incorporation of brain science into economic inquiry. Nonetheless, despite resistance by some, neuroeconomic theory, methodology, and findings continue to infiltrate top economics journals, such as: the *Quarterly Journal of Economics*, the *American Economic Review*, *Journal of Economic Perspectives*, and *Games and Economic Behavior* (to mention only several).[64] The same might be said as well in neuroscience: despite mixed receptions for the sub-discipline, neuroeconomics papers recurrently

[64] See also the special issue from 2008 of *Economics and Philosophy* dedicated to debates about neuroeconomics.

(and increasingly) appear in many neuroscience outlets (specifically, *Trends in Cognitive Science* and *Nature Neuroscience* feature many such articles). One might safely say, then, that neuroeconomics is at present glaringly visible to most economists and neuroscientists (even if such visibility remains to some extent contested).

Economic geography does not, however, appear within the current purview of neuroeconomics. That is, although many sub-disciplines of economics have tried at (and variously succeeded in) etching their marks upon neuroeconomic theory, methodology, and experiments (a list of only a few examples includes: game theory, behavioral economics, financial economics, decision theory, and welfare economics), economic geography's influence on neuroeconomic work is conspicuous for its absence. Of note: one of the leading and most recent texts in the field opens by asserting that: "Each of the major threads of contemporary research is reviewed in these pages" (Glimcher and Fehr 2014, p. xxvii). Yet none of those pages bears any direct imprint from contemporary geographic research. Likewise, an extensive review (by this author) of several hundred neuroeconomics publications from the last decade failed to furnish any evidence that neuroeconomics has been influenced whatsoever by thinking and findings from economic geography. This absence does not suggest neuroeconomics has considered, although ultimately rejected, a role for geographic principles, but rather that geography just remains invisible to it. In its attempts to map minds, neuroeconomics appears to not be paying any mind to geography.

That invisibility is unfortunate because many strands of economic geography have, in principle, substantial familiarity with numerous forces at work within the dynamic brain. Moreover, just as economic geography is an internally-heterogeneous and hybrid field, so too is neuroeconomics. And that parallel seems germane to several disputes presently embroiling neuroeconomics.

For one thing, some neuroeconomists, as well as purebred neuroscientists, have expressed disenchantment with subservience by much of neuroeconomics to economic orthodoxy, including (e.g.) the rational-optimization paradigm. These qualms may appear strange, given how heterodox neuroeconomics approaches seem (for now at least) to many workaday economists.

But the gravity of such concerns becomes clearer if one understands them as scruples about attempts to shoehorn established economic theory into the workings of the brain, rather than (as Shiller suggests) using discoveries about how brains function on their own terms to elucidate economic phenomena (cf. Camerer (2013)). From these and others quarrels are emerging divergent, and sometimes contradictory, views on theory, methods, assumptions, and goals in neuroeconomics (Ross 2008; Sobel 2009; Fumagalli 2010). Some factions of neuroeconomists are urging a sincerer embrace of complex dynamics by emphasizing the multi-scalar interactivity between space and time within brains (van Rooij and Van Orden (2011)). That is, there are theorists and experimenters who are advocating much closer attention to interplay between 'the global' and 'the local' in the brain, across multiple scales of both space and time together (see Ramsey et al. (2010). (That geographers may welcome this dawning appreciation is this paper's core assumption, and seems to jibe with persuasions of many prominent economic geographers; see, e.g.: Clark, Feldman, and Gertler (2000); Martin and Sunley (2007); and Clark, Dixon, and Monk (2009).)

Moreover, the importance of multi-scalar and complex dynamics is widely accepted in the broader discipline of neuroscience (especially computational neuroscience; see, e.g.: Rabinovich et al. (2006); Hilgetag, Müller-Linow, and Hütt (2011)). Nevertheless, contrasting views by other, more economically-oriented neuroeconomists have sparked recent infighting that threatens to intensify (see, e.g., correspondence between Camerer et al. and van Rooij and Holden (2013)). How these disputes over treatment of multi-scale dynamics (cf. Berhneim (2009)) get resolved may shape the

trajectory of not only neuroeconomics, but also future thinking within its parent disciplines of economics and neuroscience, in light of the present traction and momentum that neuroeconomics enjoys within each. But this incorporation of multi-scalar and complex dynamics is no minor matter, for: "complexity is not something that just bolts on to or can be blended with an existing conceptual/theoretical framework" (Martin and Sunley 2007, p. 595).

This paper argues that some headway might be made in resolving concerns over complex, multi-scale dynamics through fostering of pragmatic mutualism between economic geography and neuroeconomics. It might be blindingly obvious that neuroeconomists and economic geographers traditionally emphasize different subject matters (and are diverse groups among themselves). But distinct historical foci should not distract either camp from the potential benefits from exchanging compatible principles and insights.

While any interdisciplinary exchange is potentially endangered by imperialism from one side or the other (as may be the ongoing case with much of neuroeconomics being heavily-dominated by economic precedent; see, e.g.: Glimcher (2003); Glimcher and Rustichini (2004); Glimcher, Dorris, and Bayer (2005); Fehr and Rangel (2011)), fields that share common conceptual forces may mutually gain from exchange, so long as the boundaries of the analogy upon which their compatibility is premised remain respected.[65] And two such compatible forces that transcend the disciplinary boundaries of both economic geography and neuroeconomics are: *multi-scale dynamics*; and *reflexivity*.[66]

[65] Note, e.g., the cross-disciplinary insights stemming from complexity studies (Page 2011)). Yet it is only ever in probing those boundaries that they come to be known.

[66] Reflexivity and multi-scalar dynamics across time and space are not the only forces and conceptual principles connecting economic geography and neuroeconomics. A surprisingly lengthy (but by no means exhaustive) list of neural regularities with non-superficial geographic correlates includes: path dependencies; backward and forward linkages; imperfect competition; inflecting returns to scale; agglomeration effects; circular and cumulative causation;

Economic geographers generally wield a comparative advantage over other social sciences in the depth of their appreciation for how time-space interactions across multiple interdependent scales of activity shape, and are reflexively reshaped by, the structure, function, and dynamics of complex adaptive systems (see: Martin and Sunley (2007); Schwanen and Kwan (2012)). This stock-in-trade of economic geography may help explain more clearly the structure, function, and dynamics of human brains (truly complex adaptive systems). This prospect should hopefully embolden economic geographers to engage with, and aim to inspire, neuroeconomics in geographically-familiar terms. That is, although non-negligible gaps may now exist between terminology, methods, and philosophies of economic geographers and neuroeconomics, some of the forces and phenomena with which economic geographers are most comfortable may have direct import for neuroeconomics' future, as this paper will endeavor to illustrate. And while surely not every economic geographer will wish to dive into the realm of brains (and even fewer may aspire to learn the *linguae francae* of neuroscience by familiarizing themselves with its mathematics, empirics, and models), neuroeconomics could, for its part, propel advances for economic geography.

While some geographers express skepticism about the value of importing neuroeconomic findings into economic geography (cf. Pykett (2013)), benefits from keener understanding on how human minds actually operate – especially if they share with geography forces and ideals already well-known to it – should be obvious indeed. There is not space here to cover all of the paradigms and theories in economic geography that might potentially be influenced by improved understandings of how human minds work, so I focus on one of seeming interest to many geographers: *neglect*.

nonlinear and chaotic dynamics; spontaneous organization; transportation costs; specialization; spill-overs; lock-in effects; and hierarchies. These connections are touched on throughout this paper.

That people pay attention to some features of their environments, but neglect others, is a core message of Clark's (2011) work on time-space myopia. And his message that such myopia can be terribly detrimental at both the individual and societal levels inspires the question: might a sharper understanding of why human minds seem predisposed to neglectful cognition help economic geographers to better comprehend key phenomena and behaviour in economic and social systems? The response to that question derived from within this paper is, encouragingly, affirmative.

In this paper I also describe how a more geographically-attuned neuroeconomics may help economic geographers to garner insights on neglectful human behavior and judgments as reflexive outcomes of multi-scalar interactions between environments and brains (a goal that Clark (2010) advocates in citing Simon's (1990) metaphor of behavior being like scissors because it cuts through time at the intersection of the twin blades of environment and the human mind). I do so by introducing a new metaphor for neglect, a simple toy die, and show how such cubes can represent not only time-space myopia, but also another form of neglect anchored in visual metaphor: *scotamatous neglect*. Together, myopia and scotomata may explain better why individuals and societies can be blind to many large-but-rare (tail) events, like the recent global financial crisis, as well as Black Swan events famously treated in the work of Taleb (2007; 2009).

Furthermore, a geographically-sensitive neuroeconomics might help illuminate a surprising, but as yet inexplicable, finding about neglect of tail events: it is characteristically unstable, as conveyed by the work of Kahneman and Tversky (1979) and Tversky and Kahneman (1992) (among many others). Resolving this characteristic instability of neglect would harbor serious implications for behavioral economics. Moreover, as interest in general matters of neglectful thinking and tail events (separately and combined) has spiked among economists of late (see: Sims (2003); Barro (2006); Weitzman (2007); Gennaioli and Shleifer (2010); Masatlioglu, Nakajima,

and Ozbay (2012); Gabaix (2012) Bordalo, Gennaioli, and Shleifer (2012); Barberis (2013a); Wachter (2013); Gennaioli, Shleifer, and Vishny (2013; 2014); and Gabaix (2014)), it would appear that a geographically-attuned neuroeconomics may aid in projecting some core principles and ideas from economic geography well beyond the confines of the nascent neuroeconomic sub-discipline, and help sow more economic geography concepts into mainstream economics itself.

Yet, contrasted against the larger potential that a pragmatic neuroeconomics-economic geography mutualism might realize, this paper's ambition stands somewhat modest. Its aims to demonstrate to economic geographers that many analogous concepts in brain dynamics are familiar from discourses common within economic geography. The hope is that this initial exercise of translation will encourage others to seek further, as-yet-unnoticed parallels between neural and geographic forces, and attempt to reinterpret them as viable ideas for advancing neuroeconomic investigation.

The remainder of this paper proceeds as follows. Section 3.2 covers the compatible roles for analogy in economic geography and neuroeconomics. It also discusses the nature of the reflexive and dynamic brain, and defuses some errant concerns about 'biological determinism'. Section 3.3 delves into the problems of excessive concentration on brain structure and function – at the expense of dynamics and multi-scalar inquiry – and the ensuing accentuation of 'localized' views about how human brains work. Section 3.4 discusses the powerful (and portable) concept of near-decomposability as it applies to both economic geography and neuroeconomics. Section 3.5 demonstrates how neuroeconomics might help geographers rethink notions of neglect, plus how it may give fresh geographic influence to wider economic research. Section 3.6 concludes.

## **3.2**\\ Crossovers, Reflexivity, and Flux

### **3.2.1**\\ Quantitative and Qualitative Analogies

As mentioned earlier, because any pragmatic connections between economic geography and neuroeconomics must initially hang upon shared principles rather that common subjects of inquiry (or at least levels *within* subjects, as it is arguable that brains are needed to sustain economic activity and – at least in modern, globalized society – *vice versa*), the links between them are necessarily those of conceptual analogy. But neuroscience is generally both familiar with, and a beneficiary of, analogical and metaphorical approaches. Consider Minsky's (1986) likening of minds to societies, and the enduring insights that comparison has generated for computational neuroscience in (to brandish but one example) the neural-network models that Minsky helped develop (see, e.g.: Minsky and Papert (1969); Dayan and Abbott (2001)). Even more broadly, it should be noted that *all* models, including formal mathematical and economic models, are themselves analogies (Martin and Sunley 1996; Gilboa et al., forthcoming). And, while some strands of economic geography, for their own parts, lean considerably on qualitative metaphors and evocative imagery (cf. Clark (2005); Clark and Wójcik (2007)), others (e.g., the 'new economic geography') rely much on formal, quantitative models.

While numerous contemporary economic geographers seem to favor one phylum of model – qualitative versus quantitative – over the other (with some opting to outright "abandon formal models" (Martin and Sunley 1996, p. 261; see also Martin (1999b)), coexistence of qualitative and quantitative approaches in geography is advantageous (even if at times disharmonious) for engaging neuroeconomics, as innovative study of brain dynamics seemingly demands them both.

Nonetheless, if economic geographers are to engage with neuroeconomics, it may seem at first tempting to do so by invoking mathematical models from the new economic geography (NEG), given its quantitative emphasis on endogenous (i.e., semi-reflexive) and nonlinear dynamics that can give rise to structural changes. Furthermore, NEG's attentiveness to interactivity among components, its use of "imaginary histories" (Krugman 1998a, p. 12), and its concentration on self-organization (Krugman 1996; Fujita, Krugman, and Venables 1999) are focal principles which permeate wider neuroscience but have not fully-percolated into neuroeconomics. And NEG's bundling of these ideas into tidy formalisms may be warmly-received by some neuroeconomists, given how many such researchers have welcomed the rigor of axiomatic approaches to testing neuroscience (see, e.g.: Caplin and Dean (2008); Rutledge et al. (2010)).[67]

But NEG tools and theories alone will not suffice; they stress principles of general equilibrium and individual maximization (Krugman 1998b). The brain, however, surely is not a closed system (it continually receives energy inputs, both trophic and sensory, from the extra-cranial world, and also projects signals back into that outside space). So general equilibrium is at best limitedly-applicable in neuroeconomics. And whether the brain ever reaches any recognizable equilibrium across multiples levels simultaneously seems much open to enormous doubt. Moreover, the myopic, binary, and threshold-based nature of neuronal activity (simply: neurons either do or do not fire; see, e.g.: Kandel, Schwartz, and Jessell (2000)) makes maximization questionable as a

[67] Furthermore, NEG's acceptance of computer-based simulations (see Krugman (1998b)) also may be a point of interest to some neuroeconomists, given the ubiquity of such simulations in broader (especially computational) neuroscience (see Trappenberg (2002)). Although, for the most part, neuroeconomics seems to have shown a general preference for pencil-and-paper tractability in its formal analysis (a priority it shares with traditional economics) (peruse, e.g.: Glimcher and Fehr (2014)). This preference for tractability may be behind the degree to which multi-scalar dynamics are often downplayed in some of the more popular neuroeconomic work done so far. Interestingly, however, much of Paul Krugman's work on geographical economics makes abundant references to the similarities of complexity and self-organization between economies and brains; see: Krugman (1994, 1998a).

foundational operating principle.[68] Additionally, while some very elegant qualitative concepts underpin NEG models, caution must always attend rich mathematics that conceivably lack real-world manifestations (Karni 2010). Thus, while NEG tools could sizably contribute to neuroeconomics, they cannot prove the sole footing for geographic-neuroeconomic mutualism.

Joint necessity of both 'the qualitative' and 'the quantitative' should mark good news for critical geographers who defend mixed-methods approaches. It also makes economic geography relevant and applicable to exploring brains and their contextual sensitivities and dependencies. Certain qualitative and hybrid techniques of inquiry within economic geography could vitally enhance neuroeconomic understanding, especially given the roles that some neuroscientists are coming to apprehend for culture and identity in the functioning of the human brain (see Cacioppo, Visser, and Pickett (2012)). In this regard, ethnographic techniques of interview and holistic observation, combined with some of the monitoring techniques of neuroscience, may breathe fresh perspective into neuroeconomic experiments and positive theory. Somewhat ironically, these sorts of approaches might invert the traditional role of close dialogue (Clark 1998): rather than examining individuals to learn about larger-scale interactive dynamics, a gainful reversal may stem from neuroeconomic experiments that leverage close dialogue with individual subjects (combined with neural monitoring) to reveal the roles of reflexive dynamics in brains.

### **3.2.2**\\ Reflexive Dynamics

Emphasis placed on reflexive dynamics should be contrasted against alarm-sounding about 'biological determinism' (among geographers, see Pykett (2013)). Warnings of biological determinism, while popular with some authors as ammunition against neuroscience as a source of

[68] Instead, the threshold-based nature of neuronal operation has led some to suggest that the fundamental principles of neuron-based processing is nearer to Simon's (1956) notion of 'satisficing' than to conventional optimization.

understanding for human behavior, are akin to advice against sailing over the edge of the world: well-intentioned, but stemming from naïve misunderstandings of the true nature of the arguments to which they pertain. On the wider topic of the fictitious beast that is biological determinism, realize that (healthy) genes only ever become expressed in response to environmental conditions, either present or past: biological influences on behavior are both context- and path-dependent (consult, e.g., Part I in Kandel, Schwartz, and Jessell (2000)). Nor is any behavioral or cognitive action, whether reactive or proactive, ever possible for any human without some initial genetic expression having given rise to the necessary anatomical structures and physiological (pre)conditions essential for executing that action; genes and environmental context are *inseparable* (Purves et al. 2008; Squire et al. 2013). On the hollowness of biological determinism, the neuroscientist Robert Sapolsky (2014) chides: "The problem with a 'gene-environment interaction' is the same as asking what height has to do with the area of a rectangle, and being told that in this particular case, there is a height/length interaction".[69]

Furthermore, if there were anywhere in the human body that biological determinism would carry the least sway it would assuredly be in the brain. It has been demonstrated widely that the quantity of information carried by the human genome is insufficient to dictate the entire neural blueprint (see Arbib, Érdi, and Szentágothai (1999) and Trappenberg (2002)). Formation of only a minor fraction of the billions-upon-billions of neuronal interconnections (synapses) in the adult human brain is guided genetically (and even those synapses which are genetically-determined are generated during prenatal development, and many weaken or dissolve without adequate, environmentally-driven stimulation in postnatal life). The vast majority of synapses self-organize

[69] Economists also seem to occasionally make mistakes on the biological fixedness of genetic contributions; see especially: Bernheim (2009).

and adjust (in synaptic strength) due to environmental interactions. Szentágothai (1993) even asserts self-organization as *the* core force which allows for neural functioning (see: Bi and Poo (1999); Zhang and Poo (2001); Kandel (2006); Lubenov and Siapas (2008)).

Hence, human brains at any point in time are therefore plastic products of experience as much as (and plausibly more than) genetic inheritance.[70] While bound by biochemical and physiological strictures, both minds and their neural substrates are affected by the environments to which they are exposed (presently and historically), in addition to reciprocally being able to affect those environments. A neural substrate is thus a necessary, but by itself insufficient, precondition for a mind. A brain is the mind's interface with the environment, and it is inextricably embedded within both *simultaneously*, and is thus a product of their various scales.

A complex web of endogenous interdependencies enmeshes collectively the brain, mind, behavior, and the environment by dynamically binding them together. Dynamics are integral to brains – as well as their interactions with mind, behavior, and the environment – because brain function is characterized by change. Transient changes in polarity (charge) among neurons are needed to support thought (or any other activity). Other, more-durable changes (e.g., in synaptic connectivity) are further required to support learning and memory, which in turn influence and support other subsequent and/or consequent changes, which themselves are varyingly transient or durable in nature. Hence, neural change itself stays persistently changing.

While change characterizes the brain, there seems to not necessarily be a characteristic scale for such changes that remains constant through time. Brain dynamics are enduringly multi-scalar in both space and time, with changes occurring at the intracellular, intercellular, regional, and global

[70] But let us be clear: human brain function is a joint product of both forces of contextualized experience and genetics *in tandem*, and never results from one or the other in isolation, as biological-determinism fallacies imply.

levels of the brain at rates ranging from fractions of milliseconds to multiple decades (Purves et al. 2008).[71] These changes are vital. As Padoa-Schioppa (2008) notes: "the mind *is* brain activity" (emphasis his). Furthermore, these multi-scalar changes are not only fuelled by the environment and thoughts about it; they also drive thinking and behavior that reciprocally impacts the environment. To retool MacKenzie's (2006) metaphor: brains are dynamically adaptive engines, not catalytic cameras that go unchanged by their own activities.

Moreover, brains do not merely capture the environment in high-fidelity for the mind's sake. Rather, brains transform environmental signals and are, in turn, transformed by those signals, both before and after passing them along to 'the mind' (and they need not ever pass all such signals along to the *conscious* mind). Nor is the mind always just a passive recipient of environmental signals, as has been recognized since at least the path-breaking work of Magoun (1958). Contrariwise, the human mind regularly engages and alters, and is reflexively engaged and altered by, the brain whilst actively operating in, or upon, elements from the environment.

This reflexivity couples with the self-referential nature of the brain such that the structures, functions, and dynamics of brains are always in flux. To what degree these properties of structure, function, and dynamics are always changing (and whether stably so) is relative to the neural level and components one is examining. It is important to not understate that (at least for levels below that of the emergent, conscious mind) the self-referencing of '*the* brain' is not a centralized, unitary self, but rather multiple (often parallel) neur(on)al entities that both cooperate and compete, but to none of which is given perfect information of the simultaneous operation of the activities or states of the whole (this decentralized nature of the brain has unsurprisingly inspired some

[71] Notably, these changes may be not only in neurons, which (ostensibly) do the bulk of 'communicating' in brains, but also the supporting glial cells responsible for homeostasis and other maintenance activities necessary for function.

neuroeconomists to make comparisons with markets). It is this particular brand of decentralized self-referencing which seems regularly to trouble some social scientists from traditions that cling to the vestiges of logical positivism, because: "Following [Bertrand] Russell…logical positivism…banned self-referential statements" (Soros 2013, p. 312).[72] Through ignoring self-references and the reflexive interdependencies linking the brain, mind, behavior, and the environment, one cannot comfortably capture the continual volley between exogenous and endogenous changes that transpire within any one of them over a given timeframe (see: Mandler (1985); Simon (1979; 1996)).

Understanding the dynamic brain is thus crucial to resolving questions that behavioral and cognitive approaches alone (and even together) have failed to satisfactorily answer (such as the problem of unstable neglect mentioned in Section 3.1); nor are such questions reasonably answerable from cultural, socio-political, or other higher-level vantages (the instability occurs when individuals are both in isolation and when they interact). For deepened inquiry, an improved holism in understanding is needed of multi-scalar interactions between brain dynamics and all of such high-level products of minds (i.e., cultures, societies, and economies would not be possible without human learning and imagination, which necessarily rely on brains), and the contexts of which they are parts.

## **3.3**\\ Transcending Local Maps

Any understanding of the human brain (except the most partial and simplistic) must integrate three perspectives on neural organization: structural, functional, and dynamic (Arbib, Érdi, and Szentágothai 1999). The stimulus-response construct typically utilized by neuroeconomics

[72] Recall the lasting hold that logical positivism retains on economics from efforts by Friedman (1953) to promote it.

experiments has been the impetus behind neuroeconomic theorizing that attempts to link environment with brain, and brain with behavior (Harrison and Ross 2010). Yet this paradigm has resulted in preferential treatment for structural and functional understandings over those of dynamics. That is, neuroeconomics has focused heavily on mapping behavioral and cognitive functions to specific neural regions (structures) by looking for correlations between changes in activity in those regions among experimental subjects, and in externally-observable variables.

But, as Bernheim (2009) observes, it would seem that neural functions (processes) are more important for the social sciences than are neural structures in determining the foundations of higher-level behavioral and cognitive outcomes because functions may be more useful in discriminating among the enormous number of models (formal or otherwise) of human judgment and choice that are presently in circulation.[73] That is, the time-course and possible alternative paths of a neural process may be of first-order interest to social scientists, whereas the neural location(s) of such a process is only likely to be of second-order interest (at best), even though knowledge of the latter may often be needed to establish deep details about the former.

### **3.3.1**\\ Multi-Scalar Understandings

It is perhaps unsurprising, then, given the aforementioned structural-functional prioritization, that structural-functional mappings in neuroeconomics remain pervasively incomplete in several ways. For one, they typically focus upon only a single scale of space and time in the brain at once. Regarding time, they tend to focus on temporal resolutions at the level near (within an order of magnitude of) seconds because that timescale is easiest to measure with current technologies

[73] Preferential treatment of structure over process in economics is not specific to neuroeconomics; as Smith observes: "The fundamental error has been the universal tendency to think about the world in terms of equilibrium models and the consequent welfare theorems that are devoid of process" (2008, p. 14). See also McFadden (1999).

(Squire et al. 2013). And such technologies perform best at spatial resolutions (usually one cubic millimeter, or 'voxel') that naturally make brain regions the chief targets of structural-functional investigations. While pragmatic, any such single-scaled confinement of inquiry inside the brain may mask diverse neural dynamics that bear vitally upon human judgments and choices (even if such single scales are then mapped to larger external scales).[74]

Like their geographic counterparts, brain regions are not internally-homogenous structurally, functionally, or dynamically. While some regions are composed of neurons that are physiologically or anatomically similar enough to be considered homogenous, the connectivity patterns that these neurons share with other neurons in the same region, as well as neurons in other regions, are such that their function and dynamics may not be comfortably deemed homogenous (see Squire and Kandel (2009)). Sibling concepts have gained currency in economic geography as studies of "relational economy" (Bathelt and Glückler 2011), which find evidence that heterogeneous activity might be a – if not *the* – dominant force in economic dynamics across multiple scales (see also: Bathelt and Glückler (2003; 2005)).

Within neuroscience the idea that some functions are structurally 'localizable' has become popular and indeed has been confirmed for many brain regions and functions that manifest at the level of organismic behavior. Nevertheless, as pointed out long ago by the pioneering neuropsychologist Aleksandr Luria (1973): localization represents only a fraction of the truth.

[74] Recordings of activity in single neurons within non-human species, however, takes place at timescales several orders of magnitude less than one second. Such neural recordings are nonetheless highly invasive and expensive, and remain rare in experiments conducted on humans. New methods (e.g., diffusion spectrum imaging) may help overcome these hurdles, but have yet to become popular in neuroeconomics. Curiously, it seems that some geographers may occasionally have better field-vision regarding these new technologies than do neuroeconomists (see, for instance, the article on these new approaches by Zimmer (2014)).

Firstly, most of the functional localizations that have been successfully (i.e., repeatedly and robustly) identified are physically situated not in single regions, but instead across *systems* of regions. Secondly, exactness of regional (or even system) boundaries is far from unquestionable (Daw and O'Doherty 2014). (It seems that academic disputes over the precise locations of *de facto*, rather than *de jure*, borders may too be cited among neuroeconomic parallels with geography.)

Further shortcomings of localization-based depictions of human brains can be identified in phenomena that are similar to backward- and forward-linkages in economic geography (see: Li et al. (2004); Dan and Poo (2006)). The ability of an economic entity to generate value comes as much from relationships it shares with suppliers and customers as from the activities it performs; its abilities to generate value are meaningless without its linkages and have no absolute determination independent from them. Activity and communication among neurons operates similarly. Any neuron operates not in isolation, but as part of a unidirectional (often branching) relay of communication; neurons need connections for their firings to be useful. And neurons both upstream and downstream to a target neuron can impact the ultimate signal that gets transmitted; and the joint interactions among them (in relation to other competitive or cooperative chains within the brain) are what eventually give rise to behavior and judgment at the level of the organism. (These cooperative-competitive dynamics are well-established in neuroscience, but have been only thinly-treated in neuroeconomics; Caplin and Dean (2008) and Caplin et al. (2010) are exceptions).

Nor does every neuron connect equally with every other within a brain. Communication between neurons is metabolically costly, both for neurons themselves and the larger organism (i.e., there are transportation costs for signaling in the brain; Trappenberg (2002) observes that about one-fifth of a person's average caloric intake is needed to support brain function). Particular

connectivity configurations among neurons seem therefore to mitigate incremental costs – to neurons and the organism alike – of wasteful activity.

These specific connectivity patterns give rise to both structural and functional limitations on cognition that ultimately change though time as the connectivity patterns also change. That is, connectivity is dynamic and shaped by both cooperation and competition among neurons. Usually, the activity of no single neuron alone suffices to trigger activity in a downstream neuron to which it connects (Purves et al. 2008). If such synchronized activity is performed well (in terms of eventuating in accuracy at some higher scale of organization), then upstream neurons are 'rewarded' with a trophic signal by their downstream partners (Squire et al. 2013).

The limited size and divisibility of such a reward, which is typically distributed *pro rata* when delivered, induces specialization and competition among upstream neurons. In turn, these forces of competitive-cooperative tension give rise to agglomeration effects linked to specialization, and prompt varying degrees of functional specialization at the regional scale. Yet, the fact that backward- and forward-linkages necessarily extend outside of specific regions makes such localizations imperfect, and in various cases substantially so. Moreover, the principle of inflecting (and not simply increasing) returns to scale inherent in this competitive-cooperative tension may make regional functions either more or less homogenous over time, and thereby somewhat unstable at local-enough scales (which may or may not carry on up for larger ones).

### **3.3.2**\\ Roots of Localized Thinking

The prominence (and pitfalls) of localization thinking in neuroeconomics comes from three sources primarily: one theoretical-historical; one methodological; and one structural. On the theoretical-historical front, modest (purported) successes have been achieved in identifying

specific (localizable) physical structures in brains that correspond to expressions of utility and preference. Many neuroeconomists remain loyal to variants of expected utility theory (EUT) wherein preferences and subjective value are seen as separable from expectations, beliefs, and probabilities. Further, within neuroeconomics far more work has been done, and more understanding (supposedly) attained, on the structural substrates of preferences than beliefs (Sanfey et al. 2006) (see also Kable and Glimcher (2009); Caplin et al. (2010) marks an exception). This imbalance may well be owed to the fact that functional expression of preferences in the brain *appears* more localized structurally than is so for beliefs (although whether beliefs and preferences are fully-separable in practice remains debatable, as in psychology and economics; see, e.g.: Heath and Tversky (1991); Barseghyan et al. (2013a)).

Even if preferences and utility representations in the brain may be structurally localizable, beliefs do not appear to be (Glimcher 2014). Any form of non-stochastic neural change (including any learning) entails memory, and beliefs always involve memory (or lack thereof). (Pithily: cognition requires *some* form of recognition (Simon 1979).) Memory functions are known to be highly-distributed throughout the brain both within and across spatial and temporal scales (Kandel, Schwartz, and Jessell 2000; Purves et al. 2008). To be helpful to the organism, structural changes involved in memory use must be somehow synchronized both within and across these various spatial and temporal scales. The exact mechanism(s) behind this coordination remains mysterious, but some neuroscientists believe that changing patterns of dynamical activity within the brain itself allows this synchronicity across scales to be self-organized, and to persistently remain as self-reorganizing (Arbib, Érdi, and Szentágothai 1999).

In light of this self-(re)organization and decentralization it is perhaps unsurprising that any central role for memory in neuroeconomics typically remains only implicit (of note, among the numerous

review articles by neuroeconomists over the past decade, most omit any meaningful or explicit mention of memory's role in neuroeconomics; cf.: Glimcher and Rustichini (2004); Camerer, Loewenstein, and Prelec (2004; 2005); Sanfey et al. (2006); Loewenstein, Rick, and Cohen (2008)). Worryingly, if (in violation of EUT) preferences and beliefs are not perfectly-separable, then the distributed and multi-scalar nature of memory may mean that those brain functions underlying preferences are more widely-distributed and multi-scalar in the brain than is commonly presumed by neuroeconomic research, which could nullify (very) many findings.

Methodologically, localization has been claimed as a prominent finding in neuroscience partly because the most common statistical techniques in neuroscience and neuroeconomics tend to focus on measuring activity differences according to principles of maxima or thresholds. Yet, multiple brain regions will often respond to any one stimulus, particularly if it is multidimensional in its attributes (as nearly all stimuli are) (see van Rooij and Holden (2013)). In typical experimental analyses, however, only that one brain region (whose borders may or may not be disputed) which responds most strongly within the specified temporal window is counted as mattering (or possibly any number of regions exhibiting activity 'significantly' different from some, possibly semi-arbitrary, threshold of significance chosen by the experimenter(s), e.g.: 5.0%). This approach instigates several complications that carry various troubling implications.

For one, the approach of mapping functions locally to the brain area of maximum response encourages a 'winner-take-all' effect in analysis and theorizing. Stephen and Van Orden lament:

> We see an endless variety of newly discovered effects yielding an indefinite catalogue of special-purpose mechanisms. We worry that empirical cognitive science may begin to resemble a parody of reductionism because mechanisms accrue pretty much one to one with empirical effects – as though each empirical effect is visibly transparent to its underlying mechanism – a problem that has been called the *effect = structure fallacy*…and the *module mistake*. The effect = structure fallacy originates in a physical metaphor of effect = cause. (2012, p. 95)

This problem can be restated as follows. A region (or system of regions) of maximal response might not be (and generally is not) the only one active or involved in responding to a stimulus. Nevertheless, the 'winning' regions are often treated so because they register greatest *measured* changes, and thus the highest degrees of correlation, to an experimental stimulus. That winning region's causal influence on ultimate cognition and behavior may not be stronger than that of other regions which were less active *on average* over the time window examined experimentally. Given the speeds of neuronal transmissions, other regions that might have been more (or at least meaningfully) influential often get disregarded (see Bernheim (2009)). The winning region may thus, in some instances, simply be the limiting step in a larger procedural chain that visits, and is distributed across, more of the brain than one single region alone (for a computational justification, see Maass, Natschläger, and Markram (2002)). The relatively small timescales involved for contributions from these non-winning regions may be to blame.

The spatial scales involved in functional magnetic resonance imaging (fMRI, the most common empirical monitoring technique for brain activity in use within contemporary neuroeconomics) may also be impeached in the downgrading of connectivity as a core focus in the sub-discipline. Neuroeconomics seems to emphasize *gray matter* over *white matter*; whereas localized connections can be understood to represent the bulk of so-called gray matter in the brain, non-local and long-ranged projections of neurons constitute much of what is known as white matter. In each of the two cerebral hemispheres in humans, white matter comprises only around 2% of the volume of physical connections between regions in the cortex (a large network of regions typically associated with high-level thinking and judgment). The remainder is mostly made up of local, intra-regional connections (Braitenberg and Schüz 1998). The long-range (global) connections of white matter are exceedingly important functionally, but, given their lower volume relative to local

physical connections, most non-local activity of white matter becomes far less-noticeable in fMRI scanning methodologies than is more-localized activity due to gray matter.

Other authors have levelled various criticisms about the current (ab)use of fMRI data in terms of hasty and/or overstated conclusions. Particular concerns circle around problems from reverse-inference techniques (i.e., confusing data and estimates) and treatment of fMRI data as first-order observations rather than products of chains of statistical inferences (Harrison and Ross 2010) (see as well: Poldrack and Wagner (2004); Poldrack (2006). Another, perhaps less-enunciated concern is that of specificity. It is still a contested matter as to whether specific neural responses become (with, e.g., repeated exposure) more highly-localized. If specificity operates through extremely-high levels of localization at the sub-regional level in the brain, then worries should arise surrounding fMRI evidence from chains of neural responses where some links are far more localized than are others. In such instances, highly-localized (but sub-regional) responses may in fact matter strongly but be subverted in inferred importance by the greater measured activity of less-important steps that are not as strongly-concentrated in neural space.

Finally, the structural roots underneath the prominence of localization seem to emanate from a property of all *near-decomposable systems* (i.e., systems that cannot be completely described as simple sums of their parts, but that exhibit substantial amounts of activity within any of their constituent parts relative to interactivity levels between such parts). Near-decomposable systems have small-world-like connection patterns: intra-unit connections are dense, whereas inter-unit connections are sparse for any given level of the system (Simon and Ando 1961; Simon 1973; Simon 1996). This property appears to hold for brains which, like other near-decomposable systems, are complex and hierarchical (Squire et al. 2013). (Note as well the geographic parallels with models of hierarchical, but interconnected urban systems; see, for example: Fujita, Krugman,

and Mori (1999)). That brains appear to be decomposable into distinct structural and functional regions, with seemingly tight correspondences between these structural and functional regions, enables ideas of brain modularity, and thus localizability. Ostensibly, it is this supposed modularity of neural composition which perpetuates the popular notion that certain functions occur in specific structures in some isolated or hermetic fashion, such that they may (sometimes wrongly) seem like discrete steps within some input-output process.

While structural localization may be practicable for some (very) well-defined neural functions, it does not seem an applicable concept for all, especially at the level of such general, but vital, functions as memory, which tend to be highly-distributed throughout the brain (Kandel, Schwartz, and Jessell 2000). Moreover, mistaking near-decomposable systems for decomposable systems is a dangerous confusion to commit. Decomposable systems are authentically modular: in them any one module performs it function(s) independently of other modules. In near-decomposable systems there is a temporal dependence between modules; functions of specific modules may be affected by other modules and strengths of such dependencies may vary according to timescale.

## **3.4**\\ Near-Decomposability and Multi-Scalarity

Much work on inter-modular dependence in near-decomposable systems in social science has emphasized the point that dependencies tend to weaken as timescales compress. For example, note how standard thinking on regional economic activity in geography treats some regions as practically independent over very short periods (and see also Simon (1996)). In the brain, however, there may be substantial inter-modular dependencies on both very short and very long timescales, with higher proportional intra-modular dependencies along mid-range timescales.

Not only the coexistence, but also the interactivity, of multiple timescales in the brain is beginning to be understood within neuroscience as indispensable for cognitive function. For example, the interactions between so-called theta and gamma rhythms in the brain are believed to underpin (and constrain) working memory capacity. Importantly, both theta and gamma rhythms occur along relatively short timescales, (although these are still immensely long in comparison with, e.g., the all-important activities of ionic channels within neurons, which permit neuronal firing and synaptic transmissions). The significance of activity over short time horizons for cognition seems problematic for some neuroeconomic work founded on results from relatively mid-range timeframes of activity. Of further concern is the fact that working memory involves a large number of regions and sub-regions across the brain, and that these intra- and interregional connections are all essential for the proper functioning of working memory. That such a crucial mental function as working memory has both significant distributed and short-timescale components may have implications for the theories and experiments of neuroeconomics precisely *because* of the near-decomposability property within human brains.

For a near-decomposable system with some crucial interactions between its modules occurring over very short timescales (in addition to long timescales), there is a strong possibility of high sensitivity to initial conditions. A very small change early in the time-course of neural activity (either in response to some external stimulus, or else endogenous in origin) might generate vast differences in subsequent activity: the system may exhibit chaotic dynamics (Riley, Shockley, and Van Orden 2012). Indeed, much of neuroscience outside of neuroeconomics has embraced approaches toward neural modelling that rely on substantial chaotic determinism. Nonlinear dynamics can often seem like random patterns of behavior when examined along inappropriate timescales (Wolfram 2002). Yet, much of neuroeconomic theory includes a sizable role for

stochastic brain activity (Glimcher and Fehr 2014). (And many economists coming to the subject celebrate the purported roots of random utility models in what they feel to be the inherently stochastic behavior of the brain.) Perhaps examination should be undertaken in neuroeconomics of how much of this supposed randomness may truly be chaotic determinism masquerading (or misinterpreted) as randomness due to improper selections of relevant scales.

Any ability to separate deterministic from (genuinely or spuriously) stochastic components (or, to judge whether such separations are even merited), however, relies at least partly on the ability to establish the existence of and classify any *attractors* bearing on the system (some conditions on the existence of such attractors are famously discussed in Cohen and Grossberg (1983), and attractors are also common constructs in the new economic geography; see: Krugman (1994) and Fujita, Krugman, and Venables (1999)).[75] Such taxonomy is also important for the realization that specific forms of attractors, and attractor dynamics, may provide important foundations for some essential mental abilities in humans, such as learning.

In the brain, attractors may have both functional and dynamic manifestations (combinations of which may significantly decouple function from structure, or introduce nontrivial obfuscations that might hinder attempts to map neural functions onto discrete, circumscribable structures). Moreover, shifting interactions between separate attractors may spawn still further difficulties. This possibility may be one of the primary obstacles to what philosophers (and those toiling in artificial intelligence research) label *computational theories of mind*. Computational theories of mind generally maintain that thinking can be described by functional algorithms (read as: 'procedures'), and such theories typically allow for compositions of multiple algorithms (see:

[75] As phenomena, attractors are similar to the idea of agglomeration points in geography, and also tend to be self-reinforcing when they grow large enough, and difficult to break after sufficient thresholds of stability are achieved.

Koch (1999); Trappenburg (2002)). Yet, what they consider less often are situations wherein functional algorithms not only interact with one another, but also can dynamically rewrite one anther owing to competition and cooperation.

While neuroeconomists pervasively support this algorithmic prospective, they (by-and-large) significantly downplay the dynamic competitive-cooperative interdependencies of neural functions. They fail, in the main, to give sufficient attention to how these intra-cranial competitive-cooperative interdependencies evolve continuously along multiple timescales due to the ongoing exchanges between the brain, behavior, and external environment owing to feedback (whether positive or negative). Further, when competitive-cooperative dynamics have been explored elsewhere in neuroscience, they have widely been found to be nonlinear (Caporale and Dan 2008). Neuroeconomics, on the other hand, has (like economics) displayed an especial favoritism for linearity, despite longstanding non-linearity (evident even at the behavioral level; for example, non-linear functions pervade behavioral economics, and are prominent in mainstays like prospect theory (Kahneman and Tversky 1979; Tversky and Kahneman 1992)).

Thus, although neuroeconomics has done well in embracing notions of *difference*, it has not acquired (as economic geography has) due "appreciation of the process of *differentiation* – a realization that difference is actually the product of ongoing…processes that sustain long-term…variation" (Clark, Feldman, and Gertler 2000, p. 4). The competitive-cooperative processes occurring at multiple scales of space and time in the brain, coupled with persistent environmental feedback/interaction, eventuate in the brain being a continually-evolving entity in three ways: structurally, functionally, and dynamically.

The synthetic perspective common among geographers stands well-equipped to understand and make sense of this multi-scalar evolution. Whereas neuroeconomists generally still tend to

construe space and time as singular (homogenous) notions in the way economists do, "geographers see landscapes as continually evolving with the very meaning of space and time varying historically and geographically" (Sheppard 2000, p. 102). Sheppard also notes that geographers "conceive of distance as a product of…processes, not as a simple metric. They give importance to geographic scale; to the complex ways in which local, regional, national, and global processes articulate with one another. But space cannot be separated from time" (2000, p. 102).

Thus, geography – and particularly economic geography – may know far more about the dynamic, non-monolithic brain than might initially be suspected. Geography's affinity for, and intimate experience with, multi-scalar study of time-space interactions might prove exactly what neuroeconomics needs to progress in not only its growth and internal ambitions, but likewise in deepening its enhancement of other disciplines. Specifically, an important, and conceivably portable, transformation seen recently within geography may be invaluable to neuroeconomics:

> There is now a growing recognition that industries are organized in places between sets of places rather than being exclusively determined by national spaces…One of the imperatives underlying the intellectual transformation of the field [of economic geography] has been the decline of the nation-state as the dominant locus of economic activity. By focusing on the decline of the nation-sate we do not mean to suggest that it is somehow irrelevant or profoundly compromised in the face of the twin forces of globalization and localization…Firms occupy various spaces and various levels in the economic hierarchy. A better understanding of the allocation of economic activity requires a better appreciation of the functional and competitive structures of firms in relation to the inherited different system of places and spaces. (Clark, Feldman, and Gertler 2000, p. 8-13)

If one identifies the parallels among the pride-of-place that structural brain regions now hold within neuroeconomics with the former (but since downgraded – although not defunct (Rodrik 2013)) emphasis by economic geographers upon nation-states as determinants of economic activity, then a hopeful blueprint emerges regarding one possible future for neuroeconomics. Moving away from structural fixations might accomplish for neuroeconomics what the decline of nation-states as foci achieved in economic geography. Such progression for neuroeconomics need not be the Herculean (or even Sisyphean) task that it may sound in light of the complexity of neural

dynamics. Economic geographers have achieved a comparable intellectual migration despite their subject matter (humans) posing a far less-simple and generalizable focus than are neurons.

Geographers therefore both can and should help facilitate the advance of neuroeconomics by translating relevant concepts and paradigms into neuroeconomics. Meanwhile, economic geography's relationship with neuroeconomics need not be one of unilateral contribution. It could instead be one of pragmatic mutualism. But mutualism must nevertheless be more than idealistic and symbolic to be worthwhile. The next section hints that its value surpasses its cost.

## **3.5**\\ Dicing Neglect, and Socially-Oriented Cognition

Clark's (2011) notion of time-space myopia stands as a potent description of a common fact in social (not just economic) life: people tend to pay most mind to those experiences and possibilities which are most familiar (and thus pervasively near) to them spatially and temporally. Time-space myopia is thereby an invaluable encapsulation of some key ideas about neglectful cognition. Nonetheless, it demands extension in two key regards to capture neglect more comprehensively.

For one, people can and do think (on occasion) about events distant in both time and space; they are not unswervingly myopic. Yet, after having thought about distant events, people sometimes nevertheless then march on as if they never had paused to consider them at all. They neglect these distant events not because they cannot (fore)see them, but because they, for some reason(s), reject their immediate pertinence. Yet, at other times, these distant events become points of fixation, and people might worry excessively about (i.e., overweight) them in their judgments. This point is perhaps best understood with respect to the concept of *tail events*, which are infrequent but high-impact episodes that tend to have strong socio-economic consequences for individuals or groups (Taleb's (2007) notion of Black Swans is a sub-class of tail events, given his subtle definition for

Black Swans). Tail events are, due to their low probabilities, distant with respect to what should be expected. A view of neglect rooted solely in myopia would have people primarily disregarding such tail events; yet reality finds that people sometimes fret excessively over them. This instability of the threshold for neglect is well-known in behavioral economics (see, e.g.: Kahneman and Tversky (1979); Quiggin (1982); Tversky and Kahneman (1992)): when it comes to minute-probability events with big consequences, people tend to behave in ways that can seem erratic. And this finding applies to tail events about which people are aware (i.e., where myopia does *not* apply; such erratic treatment might become even wilder for aggregate levels when some people consider tail events, and either neglect or over-accentuate them, while other people fail to even consider tail possibilities in the first place).

And there is even more to the neglect story. People can also exhibit unstable neglect in their thinking when they ponder events near to them in time and space: people can ignore events within what might be called their 'immediate field of vision'. I thus suggest extending Clark's metaphor of myopia to more-fully encompass the concept of neglectful cognition by incorporating another species of visual deficiency: *scotomata*.

### **3.5.1**\\ Augmenting Myopia

Scotomata are blind-spots (or otherwise degraded acuity) within the immediate fields of all mammalian vision (even healthy vision). We humans are generally not aware of them because the brain 'fills in' what it expects us to see there. I suggest that, together, myopic and scotomatous forms of neglect might paint a fuller picture on instabilities of neglectful thinking.

To supply intuition, we can think of differences between myopic and scotomatous neglect by envisioning a simple gambling die, with six equally-sized surfaces (i.e., a cubic structure) with

patterns of colored dots dimpling each surface. If one holds the die such that one of its surfaces is perfectly perpendicular to the center of one's field of vision, then only one surface is visible and the other five are hidden from view; and only the dots on the visible surface can be counted at that time. The invisibility of the dots on the five non-visible surfaces from this positioning analogizes myopia: these dots and the surfaces on which they are imprinted are distant from the viewer; but changing perspective slightly (i.e., thinking more 'globally') can bring them into view. Yet, simply because surfaces come into view does not mean that now-visible dots will be correctly counted. A careless glance or 'intuitive' tally could easily mistake the true number of dots visible. That is, some dots may be neglected even for being in full view; this type of neglect encapsulates scotomata. Translating these ideas into real-world applications, we might readily see that the processes driving neglect (or uncertainty) from myopically not considering a distant event (i.e., viewing a particular die surface) differ in meaningful ways from considering, but scotomatously neglecting (or otherwise mistaking), the impacts of the event (i.e., its count of dots).

Moreover (as mentioned), behavioral economics and cognitive psychology have demonstrated that processes driving scotomatous neglect appears unstable, but have not succeeded in describing this instability fully for tail events (prospect theory (Kahneman and Tversky 1979) and anticipated-utility theory (Quiggin 1982) both suggest that they may not pertain only to tail events).

Tail events, however, are some of the more consequential in socio-economic systems, as the recent spate of research among economists on rare events suggests. Consider as well which of the following seems more plausible regarding the recent global financial crisis (GFC):

1) **Myopia**: no one anticipated it; or

2) **Scotoma**: its possibility (or potential for similar crises) was pondered by some, but a sufficient number neglected it so that it was treated as not a genuine threat and thus ignored.

In truth, the genuine roots of neglect in the GFC are likely some mixture of both, but it seems unduly hard to fathom that the latter possibility had no role whatsoever to play. What is interesting about this second, scotomatous possibility is the social element it engenders: it implies that if an adequate number of people had *not* neglected the GFC's possibility, then the outcome might have differed (again, this social dimension suggests the second possibility's plausibility – the infamous quotation by Chuck Prince (the former Chief Executive Officer) of Citigroup about continuing to participate in providing cheap credit as long as others continued to do so provides anecdotal support for this plausibility). But this social explanation for neglect stands as a chicken-and-egg problem prompting yet another, deeper question: why would an 'adequate number' of people have neglected the possible risks *ab initio*? One reasonable answer seems to bend back upon the instability of the neglect threshold, but with deeper consideration for its social origins.

### **3.5.2**\\ Social (Neuro)Economics

*Social neuroeconomics* seems to promise some deeper insights into the roots of this instability. It embraces an others-as-context view of how judgments progress in the brain, and thus is sensitive to (some) environmental-cognitive interactions that drive choices.[76] This sensitivity should encourage geographers, because of its prioritization of the (socially-) embedded nature of brains. But, unless it corrects for the flaws of ignoring multi-scalar dynamics in the brain, it may remain

[76] Works on this topic, however, still emphasize localization arguments (cf. Fehr and Camerer (2007)). Further, Cacioppo, Visser, and Pickett (2006) is an excellent reference in general on the wider field of social neuroscience.

shackled to the same problems that have beleaguered other strands of neuroeconomics: insufficient regard for reflexivity and multi-scalar dynamics across time and space together within the brain.

Yet, it is in social neuroeconomics perhaps more than any other branch of neuroeconomics that economic geographers might succeed, or wish, to affect sea-changes in. Given the nature of predominant scanning techniques in neuroeconomics (mainly fMRI), individual decision-making has been the logical first target for study, and therefore judgments influenced by social contexts have had to await the (backlog of) work to 'establish' foundations for individual choice. Social neuroeconomics, therefore, is younger than the more individually-focused core of the rest of the sub-discipline, and may be ripe for stronger infiltration by the geographic principles variously mentioned throughout this paper. But potentially improving explanations for neglectful cognition (and the instability of the neglect threshold) might not be the only motivations for economic geographers to aim for closer dialogue with social neuroeconomics.

A subtle (and, until recently, glacial) change is detectable in mainstream economic research (both theoretical and empirical): theories, models, and experiments rooted in context-dependencies (i.e., sensitive to the role of environmental influence on people's judgments) are on the uptick. This uptick is also coinciding, however, with slippage in the pervasiveness of total rationality.[77] It is almost tautological to say that many branches of economics are concerned with 'the social': sub-disciplines such as game theory and financial economics have long been concerned with what people think about other thinkers. And it has been a quarter-century since thinkers with realistic limits to their cognitive abilities were introduced into economic models wherein other thinkers made up part of 'the environment' in which decisions and judgments had to be made. But a

[77] Along with works listed in Section 3.1 (and references therein), see: Gabaix et al. (2006); Rubinstein and Salant (2006); Rubinstein (2007; 2013); Caplin, Dean, and Martin (2011); Shleifer (2012); Eliaz and Rubinstein (2014).

longstanding feature of those models is eroding: the efforts of 'fully rational' thinkers no longer are seen to necessarily offset the 'bounded rationality' thinkers in the majority of cases.[78] Indeed, more and more economists are exploring models wherein people can rely neither on the perfectness of their own cognitive abilities, nor those of others.[79] This movement still has not found a common rallying-point (although the number of preeminent personas involved with it suggests its momentum and potential sway); foundations are still being sought, and in the contest to embrace both context-dependencies and limits to cognition, no clear winners have emerged. Hence, many now appear to be turning to neuroeconomics for inspiration.

And it is through this deepening interest in the foundations of human cognition that geographic principles might become more widely dispersed throughout the economic mainstream. The fitness of social neuroeconomics with the needs of the aforementioned zeitgeist (with its curiosity about context and cognitive limits) seems apparent, and the path for economic geographers to mount the stage via the staircase of social neuroeconomics appears the clearest route toward center-stage.

Undoubtedly, some economic geographers may still desire to steer clear of the economic mainstream. But for those who may appreciate usefulness in a reengagement, making an alliance with social neuroeconomics (and neuroeconomics at large) poses one pragmatic course of action.

### **3.6**\\ Summary Remarks

This paper has discussed some possibilities and potential motivations for both parties in a pragmatic mutualism for economic geography and neuroeconomics. It was demonstrated how the geographies of economies and neural dynamics actually share (at an analogical level) several

[78] For example, see: Abreu and Brunnermeier (2003). Such economic models of occasional overwhelming of 'rational' agents by 'boundedly-rational' ones is, however, not new; notice, e.g.: De Long et al. (1990a; 1990b).
[79] See, e.g., the recent review article by Crawford, Costa-Gomes, and Iriberri (2013).

conceptual forces, despite their differing primary subject matters, and that these shared forces are not merely superficialities. The main area wherein economic geography could assist neuroeconomics was suggested as being in deepening the appreciation and understanding for the multiple levels of reflexive dynamics in time-space interactions within both the brain and the environment within which the brain is necessarily embedded. For their part in such a mutualistic relationship, economic geographers might harvest from neuroeconomics (and especially one freshly-reoriented towards accentuating reflexive, multi-scalar dynamics) deeper understanding of the mechanics underlying neglectful thinking (among many other topics unmentioned here, such as the foundations of identity, along with redefinitions of rationality). Furthermore, attention was directed to current shifts in interests within economic research and the suggestion was made that economic geography may be able to sizably influence a sea-change in economic thought that seems to be gravitating toward context-dependence and cognitively-limited, socially-oriented judgment and choice. It was proposed that one of the clearest routes to such influence lies through pragmatic mutualism with social neuroeconomics.

*In toto*, this article may be viewed as a call-to-arms inviting economic geographers to critically reconsider how they intend to interact with neuroscience. As this article suggests, economic geographers may be more familiar with the dynamics of the brain than they might realize initially. As a consequence, this paper has endeavored to illuminate some of these conceptual overlaps by situating the discussion in more recognizable terms and clarifying some key ideas. Clarification and translation, however, can only constitute an inaugural step, and truly pragmatic mutualism can only be achieved through action. Equipped with this fledgling awareness of what economic geography and neuroeconomics *could* do together and for one another, perhaps more economic geographers will become moved to fully capitalize on that potential.

And while the remaining space cannot fully articulate the myriad paths along which that potential might be pursued, I suggest three promising routes (aside from the topic of neglect) along which geographers may find benefit in engaging more fully with neuroeconomics. First, there has recently arisen spiking interest by human geographers in the notion of what constitutes *context* for purposes of decision-making, and how people both derive well-being from and adapt to changes in context (see, among the voluminous literature: Strauss (2008); Clark (2010; 2011); Clark, Strauss, and Knox-Hayes (2012); Schwanen and Wang (2014)). As this article has aimed to demonstrate, the dynamic construction of perception and experience of context, as well as derivation of well-being from it and adaptation to it, are necessarily mediated by both mind and brain. A deeper understanding of the temporal and spatial natures of context, therefore, might be gleaned from exploring how the reflexive dynamics of the brain help to selectively construct context from broader elements of a wider environment (both past and present).

Secondly, ideas of the performance of identity (rather than identity as fixed and exogenously endowed) are particularly magnetic for many geographers (cf. McDowell (2004; 2006)). Neuroscience is shedding new light, however, on more fundamental questions in the construction of identity, and how the mind is capable of constructing and reconciling multiple perspectives on 'self' simultaneously (see Pickel (2012) on multiple neural selves, and Kandell, Schwartz, and Jessell (2000) on some neural substrates of individuality). And while identity as a driver of economic phenomena has not seemingly enjoyed great momentum (despite its initial promise; see Akerlof and Kranton (2000)), there seems to be no established party line (or even front lines) that has emerged yet on the topic of identity within neuroeconomics. Indeed, geographers may fruitfully assert the identity of their discipline in neuroeconomics by forming the vanguard of

researchers who formulate and pursue questions about how brains, identity, and economic situations interact through time in order to generate the sorts of judgments and choices we see.

Finally, unlike economics (or, for that matter, psychology), economic geography has not been a field that has deeply interrogated the role of beliefs from a behavioral perspective. Given refreshed interest in behavioral vantages within some corners of economic geography, however, a critical reconsideration of the dynamics of beliefs (i.e., how people form and change their subjective expectations about their worlds) seems warranted. What this paper might contribute to that critical rethinking of the standard economic approach to beliefs is another challenge to the conventional supposition that people behave *as if* they formulate a comprehensive and exhaustive state space encompassing all contingencies (i.e., one that "resolves all uncertainty" (Savage 1954)), and then use all of the information they have available to them to formulate, and continually update, a set of beliefs over such a gargantuan, complicated space in ways that are not only consistent, but *consistently consistent*. Indeed, arguments here would instead suggest people selectively neglect much information at their disposals. Yet, and even more importantly, it also suggests that the ways in which people learn are themselves exposed to evolving dynamics, and that the characteristic Bayesian take on learning is not sufficient to capture the dynamic reflexivity that the ultimate neural substrate for beliefs imposes. If nothing else, it is hoped that this paper might inspire some economic geographers to tackle new paradigms for belief generation and change that are more sensitive to these (generally systematic) instabilities that are induced by the human brain itself as it continually interacts with the (limited subsets of the) environment in which it stands embedded.

CHAPTER 4

# COGNITIVE LIMITATIONS ON BELIEF APPROXIMATIONS: A THEORY OF EVIDENTIAL QUANTIZATION

## **4.1**\\ Introduction

A longstanding tenet of decision theory asserts that people can and should quantify their beliefs. Hence, normatively, they should express expectations about the world as subjective probabilities. Further, normative convention generally maintains that beliefs should be based on evidence. Generating numerical beliefs as subjective probabilities thus seems to require that evidence used in forming such probabilities also must be numerically represented within decision-makers' minds.

Yet recent work in numerical cognition demonstrates that people's mental representations of quantities are imprecise and approximate, and that this imprecision can be due to cognitive constraints on working memory and perceptual discrimination. These limitations in turn cause people to use *approximate number systems* for mental representations of quantities (see, e.g.: Whalen, Gallistel, and Gelman (1999); Gallistel and Gelman (2000); Feigenson, Dehaene, and Spelke (2004); Leslie, Gelman, and Gallistel (2008); Cantlon, Platt, and Brannon (2009)).[80] But, if a decision-maker's mental representation of evidence is imprecise and approximate, then so too may be her beliefs. A pressing question for the psychology of deliberative judgment then becomes: how does such evidential approximation impact the confidence which a decision-maker feels regarding her beliefs? And also: how might such impacts from confidence distort her probabilities?

These questions are addressed in this paper by exploiting similarities between: 1) how people deal with mental quantities; and 2) *quantization* methods taken from research on signal-processing. The approach proposed here takes as its starting point support theory (Tversky and Koehler 1994; Rottenstreich and Tversky 1997), which is a psychological approach to how decision-makers

[80] Wagner and Johnson (2011) describe these approximate representations as exhibiting "confusability", which they cite as being a key force that underlies the "protracted course of counting acquisition in early childhood" (p. 19). On approximate number systems, see also the pioneering psychophysical research done by Moyer and Landauer (1967).

(DMs) generate subjective probabilities from evidence under conditions of uncertainty.[81] Unlike many other popular theories of probability-based judgment, support theory is particularly attentive to how DMs form beliefs on the basis of evidence (although it does not say much at all about preferences or utility; the same agnosticism on preferences is maintained here).[82] Yet the current approach, called quantized evidence theory (QET), departs from support theory in several meaningful ways. First, although support theory suggests that decision-makers have limited memory and attentional capacities in summoning evidence to mind when forming probability estimates, it fails to make those cognitive limits explicit. Support theory does demonstrate that the level of detail in the description of an event can alter the subjective probability that a DM assigns to that event, because increased detail tends to: evoke more evidence from memory; or else calls to a DM's attention different aspects of each description (Tversky and Koehler 1994, p. 548-549). But it is silent about what the limits on memory and attention might be, and thus fails to make predictions on when attending to additional detail might cease to further alter judged probabilities.

QET improves upon this deficiency by designating a specific role for cognitive limits to both working memory and attentional resources that translate into crisply-defined constraints on how much evidence can be accessed from memory at once (i.e., how large of a 'sample' can be held in memory at one time). Imposing these explicit cognitive limits permits a second improvement by QET over original versions of support theory: specifying a role for confidence in terms of distinctions between evidential strength and weight (as identified by Griffin and Tversky (1992)). Essentially, original renditions of support theory assert that subjective probabilities are generated

[81] Uncertain situations here are defined as those wherein a decision-maker is not exogenously supplied with explicit probabilities. Alternatively, conditions of risk are those where probabilities are exogenously given to decision-makers.
[82] On advances in support theory, see: Brenner, Koehler, and Rottenstreich (2002); Brenner (2003); Koehler, White, and Grondin (2003). Unless otherwise stated, however, references to "support theory" made in this paper refer to its form as originally exposited by Tversky and Koehler (1994), and generalized by Rottenstreich and Tversky (1997).

from ratios between levels of personally-perceived evidence ("support") for and against a given description of an event (a "hypothesis"). Nevertheless, support theory does not distinguish between equivalent ratios based upon distinct evidential volumes. For example (and adopting Griffin and Tversky's (1992) terminology), the "strengths" of the ratios 1/5, 2/10, and 4/20 are all equal, even though their "weights" differ. Support theory does not recognize this inequality of weights as having a defined impact on probability judgment, even though the two evidential volumes may inspire distinct levels of confidence in the ensuing estimates of probability within any single DM. QET, however, does discriminate between the strength and weight of evidence, and, following Griffin and Tversky's findings, assigns a primary role to strength, with a secondary role for weight.

And QET defines weight of evidence relatively, by linking it to a DM's cognitive limits. This link between weight of accessible evidence and DMs' cognitive constraints captures properties of approximate number systems (ANSs). Existing work on ANSs suggests that many aspects of thinking involving numbers obeys a discrete, albeit imprecise, incrementing process, which characteristically exhibits *scalar variability*. Additionally, ANS research finds that people's capacities to discriminate between magnitudes is subject to ratio dependencies according to Weber's law.[83] Together, these two cognitive regularities of ANSs might help to explain Griffin and Tversky's findings that DMs often commit errors in accounting for the weight of evidence.[84] QET shows how such regularities can impact the confidence a DM has in her approximate beliefs.

The yoking of confidence to relative weight of mentally-accessible evidence allows QET to depart from standard support theory in a third way: bidirectional, systematic probability 'slanting' as an expression of confidence. That is, QET predicts when DMs might either systematically inflate or

[83] Cantlon, Platt, and Brannon (2009, p. 83) claim that the "primary behavioral signature" of ANSs is Weber's law.
[84] Support theory cannot account for either of these regularities.

deflate (i.e., slant) judged probabilities on the basis of confidence in the accessible evidence. Support theory, meanwhile, only accommodates upward slanting (inflation), by dictating that, when DMs distort their probability judgments, such judgments may tend to be subadditive (i.e., total allocation of probability across all considered alternatives may often exceed unity after summation), but will practically never be superadditive (i.e., total probability assigned by a DM will be less than one when summed). Contrariwise, a number of esteemed theories on probability judgment instead posit superadditive probability (cf.: Keynes (1921); Fellner (1961); Shafer (1976)). And, although subadditivity appears to be more commonplace, empirical work does show that superadditivity can occur across a range of different judgment situations and people (see particularly: Macchi, Ocherson, and Krantz (1999); Brenner and Rottenstreich (1999); Idson et al. (2001); Hadjichistidis, Sloman, and Wisniewski (2001); Sloman et al. (2004)). QET therefore yields a powerful enrichment of support theory: the ability to predict when a decision-maker's assessments of subjective probability may be either subadditive or superadditive on the basis of systematic slanting, driven by her own (quantifiable) confidence in mentally-accessible evidence.

The predictive ability of QET on probability slanting also bears upon a psychological debate that is yet to achieve convincing resolution: the so-called 'description-experience gap' (DEG) over small probabilities. Succinctly, the DEG pertains to findings that small-probability (rare) events are generally over-accentuated in judgments wherein likelihoods of such events are described (e.g., given exogenously through a summary probability) to a DM, but under-accentuated when such likelihoods are not given, and instead must be constructed from the DM's experiences of the event. A general (secondary) finding of research on the DEG is that when people must learn about rare-events likelihoods through experience (e.g., by 'sampling' from one or more outcome

distributions) they usually are content to make inferences based on samples that are unduly small.[85] And while Fox and Hadar (2006) (among others) have suggested that the DEG is primarily an artefact of sampling bias, other research shows that elements of the DEG persist even when corrections are made for under-sampling of rare outcomes (see, e.g., Barron and Ursino (2013) and references therein). The present paper argues that these residual elements may be explained by systematic slanting of subjective probabilities, which is driven by both: DMs' reactions to the experimental paradigms within which the DEG is typically tested; as well as their cognitive limits.

Moreover, the QET approach to capturing probability generation aligns with Wakker's (2004) insistence that, under conditions of uncertainty, decision weights (which, in many non-additive theories of probability, serve to generalize additive probabilities, and are often interpreted as distortions of DMs' probability estimates) might involve components describable as cognitive limitations, as well as preferences. Wakker's interpretation of this former potential component of decision weights is as reflecting the influence of belief misestimates on decision weights. Following Wakker's interpretation, then, QET can contribute a third class of probability-distorting 'operations' of which decision weights may be capable. Aside from *elevation* (generally understood as due to preferences) and *curvature* (typically construed as due to perceptual interferences), QET permits *shifting* of probabilities along the probability scale via slanting, which is interpreted here as a reaction to confidence in the evidence, and related to 'precision' of a DM's beliefs. Yet whether one classifies the impact of QET on decision-making as distorting of probability judgments due to weighting or misestimates *per se* becomes something of a semantic

[85] That is, if an experimental subject is permitted to engage in costless sampling from a fixed set of distributions, then a general finding is that she will tend to halt her attempts to learn the structure of those distributions after making fewer observations than would be statistically required for satisfactory inference of their actual parametric structures.

determination. Details of that distinction, as well as the place of probability slanting in the 'order of operations' of subjective decision-making, are covered further within later sections of this paper.

Importantly, as a process, quantization is straightforward; it simply involves compressive mapping of a (relatively) large range of input values into a smaller set of output values.[86] Quantization thus constitutes a many-to-few compression algorithm which typically distorts the original input value as it becomes converted into an output value. But in quantization this 'lossy' distortion carries benefits, as it entails resource-saving reductions in the sizes of outputted information (which may need to then be stored, transferred, or further manipulated). A main assertion of this paper is that quantization methods present a convenient and flexible approach as devices for modeling human thinking. Still, readers unfamiliar with them may benefit from highlighting of similarities between quantization and yet another, perhaps more mainstream, approximation tool: Riemann integration.

#### **4.1.1**\\ Analogy with Riemann Integration, and Process-Overview

Riemann integration is a basic mathematical tool for approximating areas under curves (i.e., integrals of functions) over linear intervals. As with Riemann approximations, the inaccuracy of approximation by quantization (for purposes in this paper, as well as standard approaches in signal-processing applications) is proportional to the widths of the subintervals into which the target interval is partitioned; in the limit of infinitesimally small widths, inaccuracy disappears as the discrete converges to the continuous.[87] In the case of QET, the relevant interval to be partitioned

[86] Quantization processes often entail analog-to-digital conversions and are common in image-processing applications.
[87] This reliance on subintervals marks another point of departure by QET from support theory, as well as many other so-called 'non-expected' utility theories. Although QET ultimately makes point predictions of subjective probabilities, its treatment of ranges of evidential inputs as equal to single values of probability (i.e., it violates strict monotonicity) distinguishes it from many rank-dependent models of utility. This property of QET (i.e., general equivalency among the quantized probability values for ranges of input values) makes it similar in many ways to the 'multiple priors' classes of models from economics (e.g.: max-min expected utility theory (Gilboa and Schmeidler 1989); robust beliefs (Hansen and Sargent 2001; 2008)). QET differs from most multiple-priors models, however, in that it strives to describe cognitively-plausible mechanisms for probability generation, and also imposes limits on cognitive resources.

is the unit interval of standard probability scales. Furthermore, as in Riemann integration, for coarse partitions (i.e., wherein widths of the subintervals are considerably 'large') it becomes a matter of concern *where* inside a given subinterval the approximation is made (i.e., whether one of the endpoints of the subinterval is used, or some point in between). Slightly more technically: if one looks at QET as approximating the probability density function of a DM on the basis of her evidence about some hypothesis, then QET is like a Riemann integration wherein only the height of the rectangle containing the mean (maximum height) of the curve over the relevant interval is of concern, and heights of all other rectangles are then zero-valued. These ideas will be formalized within the following sections, but a further, narrative description of QET may assist some readers.

QET intends to capture judgment processes under uncertainty wherein a DM assesses the likelihood of truth for some hypothesis about her world, for which she has some evidence both (or either) for and against. This relevant evidence under QET may be understood as a collection of discrete 'instances' (to be detailed later), which may be historically-observed cases, or else imagined or counterfactual ones which the DM simulates from memory. Like support theory, QET does not treat directly the selection process for evidence, but instead concerns what the DM does with evidence accessible to her. QET presumes that the DM engages in a mental tallying process on accessible evidence, which yields three counts: evidence favoring the hypothesis; evidence disfavoring the hypothesis; and the total collection of accessible evidence. Owing to how human minds are understood to mentally represent numerical quantities, these counts are typically only approximations; consequently, probability judgments made via QET embed cognitive imprecision.

This imprecision is mainly captured in how the third count's size (total accessible evidence) is limited by DMs' cognitive resources which (for QET's purposes) involves a joint limit on working memory and attention. Limited attention also serves another purpose in dividing the probability

scale into disjoint subintervals. QET thus depicts how a DM categorizes her beliefs and then selects an approximating probability from that chosen category to represent her beliefs with probabilities.

*Figure 4.1: Process-Flow Sketch of Evidential Quantization*

A) **Assemble**, **discretize**, and **tally** accessible evidence to produce an 'intermediate evidence ratio' (IER) →

B) **Categorize** the IER on a (potentially nonlinearly-distorted) probability scale →

C) **Adjust** the categorized value based upon comparison of the evidence against current cognitive limits

The ensuing adjustment stage of the categorize-and-adjust procedure of QET (which resembles quantization processes from signal-processing applications) is governed by how much confidence a DM has in her judgment based upon relative properties of the evidence accessible by her.[88] These evidential properties stand in relation to her cognitive limitations on working memory and attention, which may fluctuate over time and contexts. Thus, QET functions as a context-specific judgment theory, in which 'context' is a plastic mental capacity of the DM, and may vary in both available attention and working memory at any one time. This context-variability is proposed to partly explain variability in decision weighting that has been shown to occur in empirical findings.

### **4.1.2**\\ Structure

Briefly, the remainder of this article proceeds as follows. Section 4.2 discusses the motivations for treating the cognitive processes of evidential assessment as oftentimes operating on discrete collections of evidence. Justifications are offered from both research on approximate number systems and other work throughout psychology and cognitive science. Sections 4.3 and 4.4 present

[88] This process of categorization-and-adjustment driven jointly by properties of accessible evidence and cognitive limits serve an important purpose identified from the cognitive sampling literature by Fiedler (2000): cognitive judgment processes oftentimes involve two stages: 1) an estimation based on accessible evidence; and 2) a "meta-cognitive control process" that adjusts the estimates based upon specific properties of: the evidence itself; the evidential conversion process; or the context within which the gathering or conversion of the evidence transpires.

the formal model of QET. Section 4.3 focuses mainly on the roles of ratio-dependencies in human cognition, and relates these directly to cognitive limitations on working memory. Section 4.4 stresses links between: human probability judgment based upon evidence; and quantization processes. To facilitate exposition, that section temporarily presumes a linear probability scale. But Section 4.5 discards that assumption and demonstrates the possibilities and benefits of embedding nonlinear distortions of the probability scale directly into QET. Section 4.6 covers how QET may offer insights on the description-experience gap phenomenon. Section 4.7 summarizes.

## **4.2**\\ Discretizing Evidence

This section provides a description and defense of (part of) the first step in evidential quantization: discretizing accessible evidence. While not perfectly realistic across all circumstances, the presumption that DMs often operate by using discrete assessments of evidence in casting their beliefs as probabilities has much to recommend it, both empirically and theoretically. Both theoretical and empirical bodies of support for treating evidence as discretized are discussed here. First, however, it becomes important to more formally stipulate the format this discretization takes.

Following the original versions of support theory (Tversky and Koehler 1994; Rottenstreich and Tversky 1997), let a DM consider a *focal hypothesis A*, which is a description of some event. In QET (as with support theory) DMs are considered to formulate beliefs and subjective probabilities about hypotheses, rather than about events (or states of the world) directly. In considering a focal hypothesis, a DM's aim is to evaluate its truth on the basis of evidence which she is able to gather in mind. This evidence may be taken from memory, or drawn from her immediate decision environment; but in either case no probabilities of the focal hypothesis being true are assumed to be directly (exogenously) supplied to her: she must formulate them herself. To do so, she tallies the evidence $\boldsymbol{d}_A$ in favor of the focal hypothesis $A$, along with that evidence $\boldsymbol{d}_{\neg A}$ which instead

favors some alternative hypothesis $\neg A$ (which represents the negation of the focal hypothesis, wherein the focal and alternative hypotheses are mutually exclusive and collectively exhaustive propositions: one of them must be true; let these properties of mutual exclusivity and collective exhaustiveness be apparent to, and respected by, the DM).[89] Furthermore, assume that $\boldsymbol{d}_A \cup \boldsymbol{d}_{\neg A} \neq \emptyset$ and $\boldsymbol{d}_A \cap \boldsymbol{d}_{\neg A} = \emptyset$ both hold (intuitively: a DM is always able to summon some evidence to support a hypothesis; and evidence cannot support both focal and alternative hypotheses at once).

It also merits mentioning that $\boldsymbol{d}_A$ and $\boldsymbol{d}_{\neg A}$ represent the *accessible* evidence from the perspective of the DM; that is, other evidence may exist beyond $\boldsymbol{d}_A$ and $\boldsymbol{d}_{\neg A}$, but is (for purposes here) inaccessible to the DM due to (e.g.) her unawareness of it, or cognitive limitations that prevent her from gathering it in mind. As with support theory, a rephrasing of the focal or alternative hypotheses might spark different evidence to come to mind through shifts in memory or attention.

Moreover, let it be true that both $\boldsymbol{d}_A, \boldsymbol{d}_{\neg A} \in \mathbb{N}$, where $\mathbb{N}$ is the set of non-negative integers (i.e., natural numbers). This restriction to the set of natural numbers causes assessments of support for both focal and alternative hypotheses to effectively amount to 'tallying' respective increments of evidence.[90] An underlying assumption, thereby, is that these evidential increments are essentially undifferentiated in size (akin to the unit of the 'bit' in communications theory advanced by Shannon (1948); recall that such size-differentiated bits are central to Miller's (1956) seminal

[89] Some details implied by this setup are worth noting. First, the bolded $\boldsymbol{d}_A$ and $\boldsymbol{d}_{\neg A}$ signify that evidence here is taken to be set-based. Furthermore, either or both of $\boldsymbol{d}_A$ and $\boldsymbol{d}_{\neg A}$ could be time-variable and will generally differ across distinct DMs; but as the treatment here only concerns judgments made at single points in time by isolated DMs there is no need for time or individual-specific subscripts. Notice further that a lack of evidence for $A$ may sometimes (but not always) be interpreted by a DM as evidence in support of $\neg A$; whether or not it does so depends upon the specifics of the hypotheses which the DM is considering.

[90] It should be noted that, in correspondence with findings on ANS representations, no conscious counting is necessarily at work. The results merely take counting as a result of judging cardinalities, which has been proposed to reflect an *accumulation* process. See Dehaene (1997) for an extended description of the mental-accumulator metaphor.

analysis of working memory capacities). What then matters for probability assessment is the numerousness of these evidential increments, which now must be discrete, rather than continuous.

This discretization of evidence is intended to capture the scalar variability phenomenon from approximate number systems. It is important to note, however, that the discrete increments of evidence need not each represent separate 'cases' of evidence (although they well may; Koehler, White, and Grondin (2003) present an "evidential support accumulation" version of support theory that relies on evidence as pure frequency counts, but their insistence on error-free counts by DMs stands at odds with findings from research on ANSs and ratio dependencies). Discretization of evidence may instead cause one particularly strong case of evidence to be spread across a number of evidential units (or, conversely, one evidential unit to represent a number of weaker cases packed together). The main points are: subjective assessments of evidence are usually comparative (between focal and alternative hypotheses), and subject to 'rounding errors' during discretization.[91]

Rounding errors appear to arise naturally (and pervasively) as a by-product of approximate number systems, which can function cognitively as systems "for representing sets and comparing their approximate cardinal values" (Feigenson, Dehaene, and Spelke 2004, p. 311). To reiterate: a fundamental point about ANSs is that such comparisons of cardinal values (i.e., counts) are only approximate, and subject to imprecision due to ratio-dependencies. That is, ANSs appear to obey Weber's law, wherein discrimination between two magnitudes becomes more difficult as the differences between those magnitudes shrinks in proportion to the aggregate magnitudes themselves. This 'distance effect' is a robust finding within cognitive research (Purves et al. 2008).

[91] Because $\boldsymbol{d}_A$ and $\boldsymbol{d}_{\neg A}$ both concern the cardinalities of (sub)sets of evidence, it may be worth reminding the reader that all appraisals of cardinality are *by definition* comparative (as they involve comparisons with an incremental scale).

Additionally, a core assumption of support theory is that not all evidence initially comes to a DM's mind due to constraints facing her working memory. And Gallistel and Gelman (2000, p. 62) remark that scalar variability in mental representations of quantity seems to "originate within memory itself or in the reading of memory, rather than in the process for measuring duration or counting number…memory signals…are noisy". Explicitly restricting working memory and asserting discrete mental representations of evidence allows the extension of support theory proposed here – quantized evidence theory (QET) – to directly reflect this noisiness in remembered evidence. Yet QET does not rely on random noise to implement the imprecision that it reflects (as with, for instance: Brenner, Griffin, and Koehler (2005); Beard, Wallsten, and Fox (2007)). Instead, it is a deterministic theory of imprecision, and avoids complications (and potential modeling biases) which follow from the need to specify particular distributional forms for noise.[92]

Returning to the topic of discretizing evidence: some readers might reasonably protest that such discretization appears to overly-restrict the range of potential cognitive outcomes. Yet confining evidential assessments to being based on discrete rather than continuous values (set cardinalities) is by itself not nearly as restrictive as it might first appear. It is true that, by restricting $\boldsymbol{d}_A$ and $\boldsymbol{d}_{\neg A}$ to values in the set of natural numbers $\mathbb{N}$, judgments of probability based on ratios between $\boldsymbol{d}_A$ and $\boldsymbol{d}_{\neg A}$ (to be discussed in the next section) are confined to the set of rational numbers $\mathbb{Q}$; and although the set of real numbers $\mathbb{R}$ is uncountable, $\mathbb{Q} \subset \mathbb{R}$ is countable. But realise that $\mathbb{Q}$, while countable, is still infinite. Thus, sans further restrictions on $\boldsymbol{d}_A$ and $\boldsymbol{d}_{\neg A}$, judgments of probability which are based on ratios between $\boldsymbol{d}_A$ and $\boldsymbol{d}_{\neg A}$ can assume any of an infinite number of (nonnegative) values.

[92] It is notable that scalar variability has also been observed for remembered time (Gibbon 1992); while time's role is not directly considered here, 'accumulator' approaches to scalar variability commonly used to describe ANSs carry natural temporal components, and therefore links with time perception and memory seem to merit future investigation.

Nevertheless, it seems unrealistic that either $\boldsymbol{d}_A$ or $\boldsymbol{d}_{\neg A}$ might become arbitrarily large. Thus, a cognitively-grounded upper boundary for their sizes is proposed in the next section. But first some more needs to be said regarding the plausibility and psychological realism of discretized evidence.

First, discretizing evidence by using only cardinal values for accessible evidence ($\boldsymbol{d}_A$ and $\boldsymbol{d}_{\neg A}$) both fits with and further justifies a core intuition which motivated the original exposition of support theory: making direct reference to the constituent components of a hypothesis as an explicit disjunction often makes some evidence that was not previously accounted for come to count. That is, evidence packed into one single evidential unit as an implicit disjunction may become itself unpacked by splitting into multiple (equally-sized) evidential units. Such unpacking has two consequences: it can increase the total 'size' of evidence brought to bear on a judgment problem; and decreasing the size of any single unit of evidence relative to the size of total evidence. The latter consequence appears less clear under standard support theory (owing to continuous values for evidence) than does the former consequence; discretization, however, increases clarity.

Secondly, to claim that any form of human perception is continuously-valued is to essentially assert infinite discriminability for that sensory mode. While discriminability may sometimes prove quite fine, it is empirically well-established that perceptual discrimination obeys finite limits. Thus, if perceptions and judgments of probability follow other cognitive abilities, then it seems more likely that evidence is discrete, and incremented in intervals, rather than continuous, as it is with, e.g., standard Bayesian approaches (that is, even though stimulating phenomena may be continuous, mental interpretation is ultimately discrete).[93] Discrete, interval-valued evidential

[93] That finite discrimination should hold at least as much for *generative* judgments – such as subjective probability assessments – as it does for receptive judgment tasks seems reasonable, and further suggests evidential discretization. Moreover, note well the point made by Gigerenzer et al. (1989) that the eponymous rule formulated by Reverend Bayes (1763) for updating beliefs was original cast in terms of discrete frequencies, and not continuous probabilities.

bases for subjective probability judgments also conform neatly to natural frequency-based (see, e.g.: Gigerenzer and Hoffrage (1995); Gigerenzer and Todd (1999), and sample-based (see, e.g.: Stewart, Chater, and Brown (2006); Fiedler and Juslin (2006) perspectives on adaptive cognition.

Thirdly, this discrete, interval-based view of evidence may also be suitably realistic in light of the fact that evidence in judgments must often be integrated from several disparate sources, and comparisons of evidence from those various sources may pose a cognitively daunting task. A discretized re-coding of evidence would seemingly invite easier conversions between different evidential formats, even if rounding errors might arise during the course of some such conversions.

Fourthly, discretization of evidence seems to fit with psychological theories on unitization of memory; and memory often supplies evidence. Chunking theories (see, e.g.: Chase and Simon (1973); Gobet and Simon (1998); Gobet and Clarkson (2004)) may be included as one such view of the packaging of declarative memory into discrete units (and likewise into sub- and super-units).

Fifthly, within research on the mental representation of set properties, confirming evidence exists for the hypothesis that: "when presented with a set of four or more similar items, the visual system creates a representation of the set, and discards information about the individual items in the set" (Ariely 2001, p. 160). Likewise, Feigenson, Dehaene, and Spelke (2004, p. 312) suggest that "humans are attuned to the cardinal values of arrays from the beginning of life". Furthermore, episodic experiences may also become mentally represented by "summary values" (Ariely and Carmon 2000).[94] (If individual experiences are summarized, then the set-summarizing tendencies of cognition may well create meta-summaries of sets of experiences.) If, like episodic experiences,

[94] Norms have also been proposed as represented by 'summary statistics' of elements (Kahneman and Miller 1986).

evidential sets also become mentally summarized, then it would seem that set cardinality might serve as an appropriately-realistic basis for representing the judged weights behind evidential sets.

With these justifications for discretizing evidence thus exposited, we now wield some license to proceed with describing the process for quantizing subjective evaluations of evidence and beliefs.

## **4.3**\\ Ratio-Dependencies and Cognitive Limitations

As earlier stated, one of the main arguments for utilizing quantization concepts in a model of probability generation resides in the ratio-dependence of perceptual judgments, which gives rise to psychophysical phenomena such as Weber's law, along with so-called distance and size effects. As this ratio-dependence has been shown to be a pervasive feature of cognition (for both humans and many non-human animals), it seems that it deserves accounting for within theories on beliefs.

To begin, realize that the very heart of support theory itself relies on a ratio: the support (i.e., continuous-valued assessment of evidence) for a focal hypothesis, divided by the sum of the support for both the focal and alternative hypotheses being considered together by any given DM. As a DM starts to accumulate more evidence for *both* the focal and alternative hypotheses, the ratio may remain roughly constant, despite absolute growth of both its numerator and denominator. Then, as both focal and alternative support grow, so too do propensities for error in correctly discriminating evidence (i.e., discernment of the scale of differences between the numerator and denominator). As the denominator in support theory accounts for the evidence supporting both the focal and alternative hypotheses, effects due to its relative size (beyond those due to its size as multiples of the numerator) may distort probability judgments made by a cognitively-limited DM.

Pointedly, the size of the denominator (union of the support for both the focal and alternative hypotheses) in the ratio of support may induce biases in probability estimation, and these potential

biases are not readily captured in the standard formulation of support theory. Yet introducing an intermediate step accommodates them. Whereas support theory takes appraisals of evidence as translating directly into subjective probability by way of support evaluation, the quantization approach proposed here perturbs that step which translates subjective support into subjective probability. It assumes that probability is not evaluated directly as a ratio of supports, but rather that the ratio of supports may be transformed via quantization. Given the (sub)sets of evidences accessible by the DM for the focal ($A$) and alternative ($\neg A$) hypotheses, $\boldsymbol{d}_A$ and $\boldsymbol{d}_{\neg A}$, respectively, this ratio of evidential cardinalities is here called the *intermediate evidence ratio* (IER), and equals:

**(1)** $$\sigma(A, \neg A) \equiv \frac{\boldsymbol{d}_A}{\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}}$$

Further rationale for not immediately taking the IER, $\sigma(A, \neg A)$, to be equal to the DM's judged probability of $A$ being true (although this equality may ultimately hold; a key departure from ordinary support theory is that this equality can, but generally will not, hold) lies in the *ratio-bias* phenomenon. The ratio-bias phenomenon (unsurprisingly) appears rooted in Weber's law, and seems to have direct implications for probability judgment. An example of this phenomenon follows: chances of 200 in 1000 are often judged as more likely than 20 in 100, which in turn are routinely judged as more likely than 2 in 10. But the reverse does not hold; practically never is 2 in 10 seen as likelier than 20 in 100, or 200 in 1000 (see Denes-Raj, Epstein, and Cole (1995)).[95]

Instead it is proposed that the IER must be mapped into the probability scale by the DM. This mapping is, however, taken to be governed by the cognitive capacities of the DM, and some of these must be explicitly-defined for any mapping to be made specific. Limitations on these

[95] This discrepancy suggests not taking the naked ratio $\sigma(A, \neg A)$ directly to be equal to the DM's probability judgment. Interestingly, this effect also seems to be embodied in the "enhancement effect" described by Tversky and Koehler (1994); as the support for all hypotheses increases, so does subadditivity.

capacities are what is presumed to give rise to quantization of $\sigma(A, \neg A)$ (which, before quantization, may yield any rational-valued number in the unit interval, at least hypothetically; realistically, however, not all of the rational-valued points in the unit interval will be achievable). Formally, such limits restrict the maximum size $\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}$ might take according to the inequality:

**(2)** $(\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \leq l(\theta, \delta)$

Where:

**(3)** $l(\theta, \delta) \equiv \theta(1 - \delta)^{-1}$

Here $\theta \in \mathbb{Z}_+$ is a *discrimination parameter*, and $\delta \in (0,1] \subset \mathbb{R}_+$ is a *focusing parameter* ($\delta$ is akin to the "local thinking" parameter introduced by Gennaioli and Shleifer (2010)).[96] The joint necessity of $\theta$ and $\delta$ in describing 'effective' working memory capacity can be understood as follows. A high (alternatively: low) value for $\delta$ – i.e., at or near one (alternatively: near zero) – indicates that a DM is able to focus on many (few) items of evidence simultaneously. But a high $\delta$ would not necessarily amount to a genuinely high effective working memory capacity if a DM were unable to finely distinguish between these evidential items (such would demand a relatively high value for $\theta$). Conversely, a high capacity to discriminate among evidential items becomes effectively devalued if a DM lacks the capacity to bring to and hold in mind a large number of such items simultaneously. (It merits emphasis that $\theta$ does not reflect a DM's ability to distinguish between exogenous stimuli, but captures how fine her mental representations of evidence can be.)

[96] Gennaioli and Shleifer (2010) do not justify the assumption of a strictly-positive $\delta$, but presumably it corresponds to the requirement that a DM will always have *some* evidence in mind. That $\theta$ must be a positive integer, rather than a natural number (which includes zero) comes from the simple requirement in the original treatment of support theory that the aggregate support for both the focal and alternative hypotheses together must be positive. That requirement precludes $\theta$ here from being zero-valued, because $\delta$ is allowed to move arbitrarily close to (though never reach) zero.

Thus, $\theta$ and $\delta$ interact to dictate effective working memory limits (which might differ from absolute memory limits, or working memory span; those limits better reflect, e.g., the number of random digits a DM can commit to memory) in the same manner that the psychophysics of Weber's law is governed by both a stimulus' size *and* changes to it. (This interaction seems implicitly anticipated by Miller's (1956) celebrated study of working memory.) A conjecture not pursued here is that $\theta$ and $\delta$ may vary (even semi-independently) over different decision contexts.

It should also be understood that $\theta$ and $\delta$, as well as $l(\theta, \delta)$, amount to 'potentials': simply having high values for these cognitive capacities does not necessarily mean that a DM will be able to summon large amounts of evidence to mind for a given judgment problem. The amount of evidence $(\boldsymbol{d}_A + \boldsymbol{d}_{\neg A})$ accessible by a DM may fall substantially short of the limit $l(\theta, \delta)$ because (e.g.) little evidence exists ecologically, regardless of cognitive limits. Moreover, $\theta$ and $\delta$ may be suggested to a DM by a problem's framing, and initially anchor the cognitive intensity visited on a problem.

Nonetheless, if one considers $(\boldsymbol{d}_A + \boldsymbol{d}_{\neg A})$ to be possibly comprised of subjective as well as objective evidence (as Tversky and Koehler (1994) appear to do), then one might expect a higher value for $l(\theta, \delta)$ to generally be accompanied by a higher value for $(\boldsymbol{d}_A + \boldsymbol{d}_{\neg A})$ (otherwise the discrepant gap between the sizes of $(\boldsymbol{d}_A + \boldsymbol{d}_{\neg A})$ and $l(\theta, \delta)$ might be felt by the DM as 'surprising', or otherwise abnormal; potential consequences of this possibility are discussed later).

One can understand this composition of $l(\theta, \delta)$ metaphorically by thinking about the clarity of visual images rendered on a digital screen. Generally, image clarity on digital screens is a product of: 1) screen-size; and 2) number of subdivisions of the screen into pixels. Large screens with few pixels generate coarse, low-resolution images, whereas very small screens endowed with many pixels are typically unclear due to the tininess of overall pictures rendered (relative, e.g., to the

scale of normal human vision), irrespective of pixel volume. Just as both minimum screen-sizes and pixel counts are needed to render satisfactorily clear digital images, so too might the cognitive limit governing processes of constructing subjective probabilities from accessible evidence be affected by both: the depth of evidence that can be considered at once (controlled by $\delta$); and the capacity to finely discriminate between units of evidence (controlled by $\theta$). Interplay between $\theta$ and $\delta$ seems naturally capable of capturing substantial degrees of richness in bounding cognition.[97]

Additionally, the value of $\theta$ may contribute not only to the cognitive limit $l(\theta, \delta)$ in restricting evidence; it might also control how fine partitions of the subjective probability scale can become for DMs. In line with quantization concepts, let $\theta$ convert the probability scale (i.e., the unit interval, represented by: $C \equiv [0,1] \subset \mathbb{R}$) from being a continuous interval into a partition of half-open, disjoint subintervals. For expositional clarity, first consider the fictitious case with all subintervals of uniform length: $\theta^{-1}$.[98] Thus, $\theta$ is an index of a DM's 'sophistication' in using the

[97] At first glance it might seem anticipatable that some proportional correspondence should hold between $\theta$ and $\delta$, such that an increase in 'generic' cognitive capacity can be described by a strictly monotonic increase in both. If this were the case, however, there would be little need for two distinct parameters. Particularly for (most) cognitive domains where deep familiarity may harbor meaningful advantages to the DM, it is here posited that $\theta$ and $\delta$ may exhibit 'bounded independence' (i.e., mutual lack of dependencies when inside some limit of proportionality). Chase and Simon's (1973) theory of "chunking", e.g., presents one mechanism by which domain experts might have access to far more evidential units than would novices. Thereby such experts can enjoy far higher *effective* limits to their values for $\delta$ because the units with which they are operating permit a scaling effect, even though at their highest level of operational unit they obey 'normal' limitations for $\delta$ (such as those classically described by Miller (1956), and studied also by Kareev (2000); outside of their domains of expertise, however, experts can regularly exhibit unremarkable working memory capacities (Simon (1979)). While the exact correspondence between $\theta$ and $\delta$ cannot be resolved here, it is left as a conjecture (for later research) that they can depart from strict proportionality. Proportionality between $\theta$ and $\delta$ aside, it seems also plausible that they each exhibit some degree of *plasticity*, or variability, which might stem from (e.g.) motivational concerns (i.e., constraints or competing claims on a DM's limited attentional resources). For example, it seems wholly realistic that both $\theta$ and $\delta$ might be smaller for spontaneous, off-the-cuff estimates than for more reasoned judgments emanating from studied deliberation. $\theta$ and $\delta$ may therefore exhibit some intra-subjective variability resulting from personal or situational circumstances. In fact, variable expandability for $\theta$ may be why incentives work to improve the performance of some experimental subjects in judgments tasks but not others. Furthermore, it would seem to present a cognitive foundation for the probability elicitation methods famously advocated by, e.g., Ramsey (1926) and Savage (1954) that are premised on willingness to bet: incentives may push $\theta$ nearer to its natural upper limits.

[98] Equi-partitioning of $C$ into uniform subintervals may be reasonable from a normative standpoint (see Kreps (1988) on the importance of this assumption to Savage's (1954) theory of subjective expected utility). From a descriptive standpoint, however, this assumption is dubious and can be remedied with a probability-distortion function that

probability scale, and corresponds with how much of it is usable by her in judging probabilities. Division of the probability scale into discrete subintervals makes quantization processes important.

## **4.4**\\ Quantizing Evidence

Within the field of signal-processing, quantization involves dividing an output space or scale into discrete subintervals. If these subintervals are equal, as they are (for the present section) taken to be here, then the process entails *uniform quantization* (again: subdividing the probability scale into uniformly-sized subintervals is used only for honing initial intuitions, and will be jettisoned as an assumption in the next section). The process of *forward quantization* (i.e., converting an input value taken from a continuous domain into an output value from a limited, discrete range) involves four key components and three main steps; and each of these finds an analogous component or step in QET. In the standard, signal-processing version of quantization, the quartet of components is: an *input value*; multiple *decision intervals*; a *quantization index*; and a *quantized value* (output). And the three steps (chronologically) are: *decision interval lookup*; *encoder mapping*; and *index transmission*. Woven together, these components and steps work schematically according to the following. First, an input value is used to identify an appropriate decision interval (which is a particular subinterval of the output scale or space). To each decision interval is associated a unique quantization index value (typically integers); identification of the appropriate decision interval thus maps the input value into a quantization index through a process of encoder mapping. Finally, the resulting quantization index value is then transformed into the quantized (output) value in a process of index transmission. The flow chart below summarizes the overall process schematically.

compresses some subintervals while expanding others; this route is taken in a later section, but for now we consider the partition of $C$ as involving uniform subintervals, which implies linear treatment of the probability scale by a DM.

*Figure 4.2: Schematic Flow Chart of Generic (Signal-Processing) Quantization*

| | |
|---|---|
| **Component sequence**: | input value → decision interval → quantization index value → quantized (output) value |
| **Step sequence**: | decision interval lookup → encoder mapping → index transmission |

This schematic sequence of steps and components has an analogous sequence in how DMs are understood to perform a mapping $\sigma(A, \neg A) \rightarrow C$ under QET when facing constraints on: $\theta$ and $\delta$.

$\theta$ partitions the unit interval $C$ into disjoint subintervals (decision intervals) $c_v \subset C$, each of length $\theta^{-1}$. And the judged probability is only able to assume specific values within any $c_v$, which can induce some error. The endpoints of each decision interval are determined by the quantization index $v \in \boldsymbol{V} = \{1, \dots, \theta\} \subseteq \mathbb{Z}_+$. The lower endpoint of each $c_v$ is therefore the product $(v-1)\theta^{-1}$, and the upper endpoint is $v\theta^{-1}$ (where $\theta^{-1}$ serves as an incrementing scalar). For each $c_v$, only one of its endpoints is closed (inclusive); the other is open. Whether this inclusive endpoint is the upper or lower endpoint proves to be only important at the extreme: when $\delta$ equals its maximum value of one. Endpoint inclusion and exclusion for decision intervals will be discussed very soon. First, however, we must stipulate an index transmission function for completing the mapping: $\sigma(A, \neg A) \rightarrow C$. Such a function yields a quantized probability $Q\big(\sigma(A, \neg A)\big)$ and can be written as:

(**4**) $$Q\big(\sigma(A, \neg A)\big) = \theta^{-1}\left(\sum_{v=1}^{\theta} 1_{c_v}\big(\sigma(A, \neg A)\big) S(v, \delta)\right)$$

(The version of the transmission function above is essentially a probability mass function.) In the above equation, $1_{c_v}(\cdot)$ is the standard indicator function, where $1_{c_v}\big(\sigma(A, \neg A)\big) = 1$ if it is true that $\sigma(A, \neg A) \in c_v$; but $1_{c_v}\big(\sigma(A, \neg A)\big) = 0$ if $\sigma(A, \neg A) \notin c_v$. This indicator function identifies which decision interval $c_v$ corresponds appropriately to the received input value of the IER: $\sigma(A, \neg A)$. $1_{c_v}\big(\sigma(A, \neg A)\big)$ thus serves the role of a 'decision interval lookup' function. (The

summation notation is only required to ensure that the DM considers the correct decision interval.) The fact that $1_{c_v}\big(\sigma(A,\neg A)\big) = 0$ if $\sigma(A,\neg A) \notin c_v$ means the DM neglects all other subintervals. In essence, when judging $Q\big(\sigma(A,\neg A)\big)$, the DM at first only considers the strength of evidence. (The intuition that DMs focus on the strength of evidence and then make (generally inadequate) adjustments for other properties of the evidence allows QET to follow Griffin and Tversky (1992).)

Once the appropriate decision interval is identified, encoder mapping must take place to identify the correct value $v$ from the quantization index which corresponds to that decision interval. The behavioral reality reflected here is that DMs typically express probability judgments as point values, even though these point estimates can actually reflect ranges of probability values that the DM would find acceptable for representing her accessible evidence for the judgment problem faced. This reality should be captured by theory. Yet, encoder mapping does not seem entirely satisfying for doing so. That is, encoder mapping simply identifies the value $v$ from the quantization index $\boldsymbol{V}$ which is the upper endpoint of the decision interval $c_v$ that is selected as containing $\sigma(A,\neg A)$. The value from multiplying $v$ by the scalar $\theta^{-1}$ might not, however, be the most 'representative' value for $c_v$ within the DM's mind. As with signal-processing versions of quantization, this value must be converted by index transmission into a more representative value. The inside function, $S(v,\delta)$, within the transmission index function accomplishes such conversion. This embedded function serves as a bidirectional 'slanting' function that is conditionally equal to:

**(5)** $$S(v,\delta) = s_{up} = (v-1)\delta + v(1-\delta) = v - \delta, \qquad \text{if } \big((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta,\delta)\big) \geq \tau$$
$$S(v,\delta) = s_{down} = (v-1)(1-\delta) + v\delta = v + \delta - 1, \qquad \text{if } \big((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta,\delta)\big) < \tau$$

And, in turn:

**(6)** $$c_v = [(v-1)\theta^{-1}, v\theta^{-1}), \qquad \text{for } s_{up}$$
$$c_v = \big((v-1)\theta^{-1}, v\theta^{-1}\big], \qquad \text{for } s_{down}$$

Given how their associated decision intervals $c_v$ are bounded, $s_{up}$ and $s_{down}$ respectively correspond to the two major types of uniform quantizer: *mid-tread* and *mid-riser* quantizers. (These also correspond, as will be shortly shown, to notions of 'upper' and 'lower' probabilities.)

Mid-tread quantizers (like $s_{up}$) can contain zero as an endpoint, whereas mid-riser quantizers (like $s_{down}$) cannot. It can be seen, however, that which endpoint is open and which is closed only matters for $\delta = 1$, because $\delta \in (0,1]$ (i.e., it cannot equal zero, and thus fails to reach either endpoint even though it approaches zero). Whether $s_{up}$ or $s_{down}$ gets selected is therefore more a matter of which endpoint of a decision interval receives greater weight in generating a quantized output. Thus, the step of index transmission can be understood as an interpolation within the selected decision interval. This interpolation is comparable to a piecewise-linear process (among the simplest forms of interpolation), whereby the quantized probability value $Q\big(\sigma(A, \neg A)\big)$ becomes a weighted average of the selected decision interval's endpoints: $(v-1)\theta^{-1}$ and $v\theta^{-1}$. The focusing parameter $\delta$ allocates relative weights to both of these lower and upper endpoints. DMs with higher values for $\delta$ (i.e., near or equal to one) should, *ceteris paribus*, exhibit different interpolative judgments (i.e., different quantized probabilities) than those with low values for $\delta$.[99]

[99] It should be stressed, before moving on, that $Q\big(\sigma(A, \neg A)\big)$ may not be the probability that a decision-maker would necessarily *report* (e.g., verbally) if directly asked for her belief about $A$. Instead, $Q\big(\sigma(A, \neg A)\big)$ is a distorted probability upon which the DM would act (or further distort, based on interferences from preferences; see the following section) when making a decision. This distinction follows similar logic to that offered in prospect (Kahneman and Tversky 1979) and cumulative prospect (Tversky and Kahneman 1992) theories, wherein a DM may be aware that the probabilities which she professes to hold are not necessarily those upon which she takes actions. Thus, distortion of probabilities by quantization makes $Q\big(\sigma(A, \neg A)\big)$ a form of decision-weighted probability influenced by the DM's own particular perceptual and cognitive limitations, following suggestions by Wakker (2004).

Two particular instances exist wherein $s_{up} = s_{down}$. First, that equality holds when $\delta = 0.5$. Second, $s_{up}$ and $s_{down}$ will converge as $\theta \to \infty$ (and $\theta^{-1}$ shrinks to become infinitesimally small). Essentially, that infinite limit of $\theta$ describes a fictional agent with perfect discriminatory powers.

Yet for all other combinations of parameter values, it will be true (all else equal) that $s_{up} \neq s_{down}$. In the original exposition of support theory, Tversky and Koehler (1994) relate ideas of upper $p^*(A, \neg A)$ and lower $p_*(A, \neg A)$ probabilities to support theory by modelling upper and lower support functions (wherein their measures of support are taken as positive, and real-valued; furthermore, realize that such probabilities do not undergo the quantization process described here). Respectively, $p_*(A, \neg A)$ and $p^*(A, \neg A)$ designate the smallest and largest values which a DM's subjective probability judgment *could* take under uncertainty, and still have that DM accept it as (reasonably) reflecting her appraisals of evidence accessible at the time of formulating a judgment. The difference $p^*(A, \neg A) - p_*(A, \neg A)$ can proxy the DM's uncertainty. Analogously, $p_*(A, \neg A) = s_{down}$ and $p^*(A, \neg A) = s_{up}$ can reflect the respective lower and upper limits that $Q\big(\sigma(A, \neg A)\big)$ might take over the full range of $\delta \in (0,1]$ for a given pair of: $\sigma(A, \neg A)$ and $\theta$. Notably, for such a situation, $p^*(A, \neg A) - p_*(A, \neg A) = \theta^{-1}$, which aligns with the fact that the difference between upper and lower probabilities $p^*(A, \neg A) - p_*(A, \neg A)$ disappears with $\theta \to \infty$.

It should also be apparent that the degree of error emanating from the quantization process itself is simply the difference: $\Big(Q\big(\sigma(A, \neg A)\big) - \sigma(A, \neg A)\Big)$ (although this value might be more fittingly called the *quantization distortion*). And the inequality $\Big(Q\big(\sigma(A, \neg A)\big) - \sigma(A, \neg A)\Big) < \theta^{-1}$ describes a bound on possible quantization error. Thus, $\theta$ both helps to induce and bound degree of error in probability generation that DMs may make in QET (again, the idea is reinforced that $\theta = \infty$ represents the discriminatory powers of a DM who would fail to make quantization errors).

Before going further in interpreting $s_{up}$ and $s_{down}$, some more treatment must be given to the conditions under which either one or the other becomes selected. Whether $S = s_{up}$ or $S = s_{down}$ is determined by the size of the value $\left((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta, \delta)\right)$ relative to a specific threshold value $\tau \in (0,1] \subset \mathbb{R}_+$. The value $\left((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta, \delta)\right)$ essentially determines how close to her cognitive limits a DM is operating. When $\left((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta, \delta)\right)$ is high, the subject has nearly as much total evidence $(\boldsymbol{d}_A + \boldsymbol{d}_{\neg A})$ gathered in mind as her cognitive capacity (which is dictated by $l(\theta, \delta)$) can accommodate. But when $\left((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta, \delta)\right)$ is low, then the DM will have 'spare' capacity. The threshold $\tau$ demarcates when $\left((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta, \delta)\right)$ is relatively high or low, and this highness or lowness should hold differing impact for DMs with distinct values of $\delta$.

The threshold $\tau$ is thus an important driver of judgments under QET because it directs the slanting function $S(v, \delta)$. But before $\tau$ is discussed more thoroughly, further examination is needed of the value $\left((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta, \delta)\right)$. Considerable debate among psychologists has arisen surrounding both *when* (or even *whether*) and *how* DMs take sample sizes into account in judging probabilities (see, e.g.: Sedlmeier and Gigerenzer (1997)). While it stands plausible that subjects might sometimes take sample size into account, but other times neglect it, when making probability estimates, a firmer theory is needed on when such adjustments occur and when they do not occur.

Furthermore, if one insists that adjustments are made when sample sizes are small, then one needs both a mechanism for relative determination of whether indeed a sample qualifies as 'small', and also an additional mechanism to govern the size and direction of adjustment. The combined apparatus presented here roughly fulfils those desiderata, primarily through its reliance on the size of $\left((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta, \delta)\right)$ relative to the threshold $\tau$ (as opposed to some notion of absolute sample size or reliance upon references points which are otherwise external to the judgment process itself).

Importantly, the usage of $\left((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta, \delta)\right)$ relative to the threshold $\tau$ allows for the empirically-driven insight by Macchi, Osherson, and Krantz (1999) that, in constructing a theory wherein DMs are capable of exhibiting either superadditive or subadditive probability judgments, such a theory cannot simply map low evidence directly into low probability. (More will be said shortly on how the present theory can alternately generate superadditive and subadditive probabilities.) Hence, we follow here a suggestion (by Heath and Tversky (1991)) that the value $\left((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta, \delta)\right)$ relative to the threshold $\tau$ engenders the feeling of "competence" by the DM as a relativized proportion of what is "known", $(\boldsymbol{d}_A + \boldsymbol{d}_{\neg A})$, benchmarked against what "can be known" (if only hypothetically): $l(\theta, \delta)$.[100] Notably, whereas $l(\theta, \delta)$ restricts the total *weight* of evidence that can be brought to bear in a judgment problem, the fact that $l(\theta, \delta)$ serves as the divisor in $\left((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta, \delta)\right)$ means that $l(\theta, \delta)^{-1}$ impacts how the assessed *balance* of evidence (i.e., its strength) can become shifted due to its specific relation with the threshold value.

Again, the threshold $\tau$ is a vital component of the quantization process (but can become neutralized when: $\delta = 0.5$; or $\theta \to \infty$), and its role in reflecting the effects of ratio bias (among other psychophysical phenomena) seems substantial. While still conjectural, it does not stand altogether unreasonable that $\tau$ might actually be a function of the evidence itself (whether from memory or the external, immediate environment). Speculatively, one plausible functional form for $\tau$ could be:

$$(7) \qquad \tau(\boldsymbol{d}_A, \boldsymbol{d}_{\neg A}, \varsigma) = 1 - \frac{|\boldsymbol{d}_A - \boldsymbol{d}_{\neg A}|}{\boldsymbol{d}_A + \boldsymbol{d}_{\neg A} + \varsigma}$$

The above functional form is similar to the "salience" function proposed by Bordalo, Gennaioli, and Shleifer (2012) in their salience theory of risky decision-making. This function embeds both

[100] Heath and Tversky gave that suggestion in qualitative terms, and used neither the present mechanisms nor notation.

size and distance effects that might arise from a difference in the discretized values of evidential strength: $\boldsymbol{d}_A$ and $\boldsymbol{d}_{\neg A}$. ($\varsigma$ is a context-specific 'normativity' parameter that will be explained shortly.) This salience functional form for $\tau$ captures Rottenstreich and Tversky's (1997) intuitions that, generally, "judgments are likely to be superadditive when there is little evidence for each of the component hypotheses, and they are likely to be subadditive…when the evidence strongly favors one of the components" (p. 413). As can be seen, for values of $\delta < 0.5$, superadditivity is more likely if $|\boldsymbol{d}_A - \boldsymbol{d}_{\neg A}|$ and $(\boldsymbol{d}_A + \boldsymbol{d}_{\neg A})$ are both small (thus making $\tau$ smaller) and therefore $s_{down} = \nu + \delta - 1$ more likely, *ceteris paribus*. Meanwhile, subadditivity is more likely if $|\boldsymbol{d}_A - \boldsymbol{d}_{\neg A}|$ is large relative to $(\boldsymbol{d}_A + \boldsymbol{d}_{\neg A})$ (which makes $\tau$ larger and, all else being equal, $s_{up} = \nu - \delta$ more likely). Realize that the opposite situation from the foregoing holds whenever $\delta > 0.5$.

Furthermore, notice that the degree of either super- or subadditivity may generally be determined by the difference: $Q\big(\sigma(A, \neg A)\big) - \sigma(A, \neg A)$. A brief series of examples might solidify intuitions. First, consider the following situation (*Example 1*): $\boldsymbol{d}_A = 2$; $\boldsymbol{d}_{\neg A} = 2$; $\theta = 5$; $\delta = 0.2$; $\varsigma = 0$ (i.e., for the time being disregard the impact of $\varsigma$; it will be discussed in detail shortly). Under this arrangement, $\tau(\boldsymbol{d}_A, \boldsymbol{d}_{\neg A}, \varsigma) = 1$, $\big((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta, \delta)\big) = 0.64$; the DM in this example has not been able to gather enough evidence in mind to be relatively near to her cognitive limit, and she therefore slants her quantized probability downward such that $Q\big(\sigma(A, \neg A)\big) = 0.44$, even though $\sigma(A, \neg A) = 0.5$.[101] If we assume that the DM has access to this same set of evidence in making the complementary judgment (i.e., with the focal and alternative hypotheses reversed), then

[101] Recognize that this difference, $Q\big(\sigma(A, \neg A)\big) - \sigma(A, \neg A)$, is equal to the signed quantity $\delta\theta^{-1}$. That effective working memory capacity drives the degree of slanting of quantized probabilities in the present theory thus aligns with findings by Dougherty and Hunter (2003a; 2003b) that working memory strongly influences the degree of subadditivity in DMs' subjective probability judgments (along with processes of judging probabilities more generally).

$Q\big(\sigma(\neg A, A)\big) = 0.44$ and $\sigma(\neg A, A) = 0.5$. In total, $Q\big(\sigma(A, \neg A)\big) + Q\big(\sigma(\neg A, A)\big) = 0.88$, which causes this quantized probability judgment to be superadditive (i.e. sum to less than one) by $-0.12$.

Next, consider a slightly altered version of the above setup (*Example 2*) which is identical to Example 1, except that now $\boldsymbol{d}_{\neg A} = 4$. In Example 2, then, the DM's cognitive limits remain unchanged, but she now has access to more evidence for the alternative hypothesis ($\neg A$) on both a relative (i.e., as comparted to $\boldsymbol{d}_A$) and absolute basis; both the strength and weight of evidence have changed for $\neg A$ vis-à-vis Example 1. With this new evidence for $\neg A$, the DM now slants her quantized probabilities upward (rather than downward), as: $\tau(\boldsymbol{d}_A, \boldsymbol{d}_{\neg A}, \varsigma) \approx 0.67$ and $\big((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta, \delta)\big) = 0.96$. Furthermore: $Q\big(\sigma(A, \neg A)\big) = 0.36$; and $Q\big(\sigma(\neg A, A)\big) = 0.76$. As $Q\big(\sigma(A, \neg A)\big) + Q\big(\sigma(\neg A, A)\big) = 1.12$ the overall judgment proves subadditive to degree of: 0.12.

Cases where $\delta > 0.5$ prove interesting, and will be treated in a later section in discussing the description-experience gap. For now, recognize that a peculiarity exists (*Example 3*) when $\boldsymbol{d}_A = 3$; $\boldsymbol{d}_{\neg A} = 3$; $\theta = 5$; $\delta = 0.2$; $\varsigma = 0$. Here, the weight of total evidence $(\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) = 6$ is the same as in Example 2, but the quantized probabilities remain the same as from Example 1. Then for Example 3, $\tau(\boldsymbol{d}_A, \boldsymbol{d}_{\neg A}, \varsigma) = 1$, $\big((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta, \delta)\big) = 0.96$, so quantized probabilities are slanted downward, and there is a degree of superadditivity equal to $-0.12$. Intuitively this result feels uncomfortable: we should anticipate, as per empirical findings elsewhere, more distortion for probabilities falling at the extremes of the probability scale than in the center (i.e., Examples 2 and 3 should exhibit different degrees of deviation from summation to unity). Indeed, this result stems from the simplistic assumption of the probability scale being split into a number $\theta$ of uniform (and disjoint) decision intervals, each of equal length. This linearity assumption will be discarded in the

next section to restore greater realism. Before implementing that alteration, however, let us return to a consideration of the role played by the normativity parameter $\varsigma$ in generating threshold values.

The threshold function $\tau(\boldsymbol{d}_A, \boldsymbol{d}_{\neg A}, \varsigma)$ is basically a mental distillation of how symmetric the accessible evidence is (i.e., how large is $|\boldsymbol{d}_A - \boldsymbol{d}_{\neg A}|$) once it has been adjusted for its total cardinality, $(\boldsymbol{d}_A + \boldsymbol{d}_{\neg A})$, and then again by a normativity parameter $\varsigma$. An intuition is that, the more asymmetric is the cardinality-weighted balance of evidence, the more overconfident DMs will tend to be in judgments. Conversely, with greater symmetry, confidence erodes. As the salience functional (used above) clearly measures degrees of asymmetry, it thus appears fairly well-suited.

But what about the normativity parameter $\varsigma$? For technical purposes it is required here that $0 \leq \varsigma$; but otherwise no restrictions are placed on the value of $\varsigma$. One interpretation of $\varsigma$ follows from norm theory (Kahneman and Miller 1986), and reflects how 'surprising' or 'atypical' is the evidence that ultimately becomes accessible to the DM. (Sloman et al. (2004) find that changes to probability judgments from unpacking typical evidence tends to increase subadditivity, whereas atypical evidence increases superadditivity; but their results are stronger for the effects of atypical evidence on the degree of superadditivity.) While we suggest that normativity may pertain to the overall amount of evidence that becomes accessible to the DM, it could equally pertain to aspects of the individual items of evidence which become accessible to the DM (the latter situation would mean that, although individual evidential units may not differ in size, they still might retain enough of their individual characters to have distinguishable identities based on how surprising they are).

One aspect of this perspective on the normativity parameter embeds a claim by Kahneman and Miller that psychological norms are usually constructed *ex post facto*. Here such would mean that the normativity of the recruited evidence which becomes accessible to a DM is judged only *after*

it becomes accessible, which jibes with Griffin and Tversky's (1992) assertion that adjustments for properties of evidence other than its strength occur subsequent to initial estimations of strength.

This interpretation naturally aggregates several cognitive phenomena which have been cited as influencing probability judgments, such as familiarity (including habituation) and competency effects (see, e.g.: Heath and Tversky (1991); Fox and Tversky (1995); Fox and Levav (2000)). For example, if we assume (plausibly) that evidence comes to mind sequentially, and not all at once, then a DM making a decision in a familiar domain may tend to recruit to mind incremental evidence that does not violate the norm constructed from evidence that preceded it; whereas a DM operating in an unfamiliar domain may have evidence come to mind that is surprising relative to the norm constructed for it.[102] Etchart-Vincent (2004) states: "Ability to discriminate is also connected to the decision-maker's familiarity and competence for the given gamble situation. High competence may counterbalance or even prevent diminishing sensitivity" (p. 228). Thus, $\varsigma$ should be inversely related to normativity, and hence should tend to decrease when normativity increases.

Yet $\varsigma$ may also reflect a powerful circularity in human judgment: often only the most typical or representative (in a subjective sense) instances come to mind, and therefore typicality may usually keep the parameter $\varsigma$ low, which, *ceteris paribus*, would keep $\tau$ lower and induce pervasive subadditivity for $\delta < 0.5$. Still, whenever $\varsigma$ is high, $\delta < 0.5$ might then increase superadditivity. (Such circumstances will be detailed later in the section covering the description-experience gap.)

In sum, it should be apparent that substantial descriptive richness may derive from the normativity parameter $\varsigma$, and identification of proper values for it, along with empirical investigations about

[102] It seems plausible that such an explanation may help to sketch the boundaries of when confirmatory bias may affect a DM; for one model of confirmatory bias, see Rabin and Schrag (1999).

the validity of the proposed functional form for $\tau(\boldsymbol{d}_A, \boldsymbol{d}_{\neg A}, \varsigma)$, are topics begging future research. But, more urgently, let us jettison the assumption of uniform partitioning of the probability scale.

## **4.5**\\ Non-Uniform Decision Intervals

Perhaps the most contentious assumption in the foregoing elaboration of QET has been the usage of a linear probability scale, such that the decision intervals are all of uniform lengths. This assumption seems patently false for standing at loggerheads with an enormous amount of empirical demonstrations on both: the nonlinearity of human probability judgments; as well as compressions of mental representations of quantities. As mentioned previously, this linearity assumption was introduced to ease exposition, and to facilitate concentration on introducing the quantization process into a model of subjective probability judgment during decision-making under uncertainty. It would be self-contradictory, however, to insist on its continued retention as a necessary assumption. This section depicts how QET can permit nonlinear distortion of the probability scale.

Among the first temptations in trying to incorporate nonlinear distortion is to simply suppose that the quantized value $Q\big(\sigma(A, \neg A)\big)$ is transformed by a probability weighting function which induces nonlinear compressions of the probability scale to yield 'decision weights' (on such weighting functions see, e.g.: Prelec (1998); Gonzalez and Wu (1999)). Yet this approach to nonlinearity seems artificial for QET: it would appear more appropriate to have it embedded into quantization. Hence, it seems natural to insist that the decision intervals $c_v$ are themselves of non-uniform lengths from the viewpoint of the DM. Rather than distorting $Q\big(\sigma(A, \neg A)\big)$ *ex post* in order to account for probability nonlinearity, then, we might map $\sigma(A, \neg A)$ into a probability scale that is partitioned into disjoint, but unequally-sized, decision intervals, and then is adjusted via $\delta$.

A main question for that approach, however, is how to go about determining the specific lengths of each decision interval if they are not all of equal lengths. Worryingly, within the empirical literature on probability weighting functions, there exists substantial disagreement about the particular functional form for generating decision weights. A selection of some forms is listed here.

In Table 4.1, $p$ denotes the input probability which gets distorted, and the other parameters, $\gamma$ and $\beta$, control the extent and direction of distortions. Like $\theta$, $\gamma$ is often considered a 'discrimination' parameter that dictates the amount of curvature exhibited by the decision weighting function. Meanwhile, $\beta$ is generally deemed an 'elevation' parameter that is commonly construed as being influenced by a DM's attitude toward risk, uncertainty, or ambiguity (depending on the decision setting). In light of these interpretations, QET can be seen as adding a new form of operation to the distortion of probabilities by decision weights: *slanting* along the probability scale owing to $\delta$.

Many proposed decision weighting functions seem to harbor components which reflect both preferences and cognitive limits (Wakker 2004). A common empirical finding, however, is that such weighting functions appear considerably variable (see, e.g.: Kahneman and Tversky (1979); Gonzalez and Wu (1999); Bruhin, Fehr-Duda, and Epper (2010)). It might be useful, therefore, to separate the weighting that is due to cognitive limitations from that of preference interference. That desideratum motivates demonstrating QET here using the functional specification suggested by Tversky and Kahneman (1992) (see below), as its lone parameter is the discrimination parameter: $\gamma$. As the general claims made in the following apply equally to the other decision weighting functions listed below (given their similarities to Tversky and Kahneman's), the particular choice to utilize one specification over the others here is made from convenience. (Future empirical work could well help to establish which form(s) align best with ideas from QET.)

*Table 4.1: Examples of Functional Forms for Decisions Weights*

| Study Authors | Functional Specification |
|---|---|
| Gonzalez and Wu (1999)[103] | $\frac{\beta p^{\gamma}}{\beta p^{\gamma} + (1-p)^{\gamma}}$ |
| Prelec (1998) | $\exp(-\beta(-\ln(p))^{\gamma})$ |
| Tversky and Kahneman (1992) | $\frac{p^{\gamma}}{(p^{\gamma} + (1-p)^{\gamma})^{\gamma^{-1}}}$ |

It is worth noting that the measured values of $\gamma$ from experimental work seem to agree with the ranges proposed for $\theta$ in the paper here, if one takes the two parameters as related by the following:

**(8)** $\gamma = 1 - \theta^{-1}$

This relation bounds $\gamma$ between zero and one (the latter boundary amounts to no curvature of the probability scale, i.e., linearity in probabilities) if $\theta$ is taken to be a strictly-positive integer, as in the present theory. Pointedly, most empirical estimates (including both those derived under risk and uncertainty) place the value of $\gamma$ somewhere between 0.5 and 0.9, which would situate $\theta$ at between 2 and 5, and would imply that, in formulating probabilities under uncertainty, people would use very limited portions of the probability scale by dividing it into relatively few subintervals (on these empirical values, see, e.g., the studies in Table 4.1, as well as those by: Etchart-Vincent (2009); Bruhin, Fehr-Duda, and Epper (2010)). And if the relation in Equation 8 does hold, then nonlinear mapping of evidence into the probability scale becomes straightforward.

[103] The functional specification used by these authors appears to be due originally to Goldstein and Einhorn (1987).

But to achieve such a nonlinear mapping we must revise our formal definition of decision intervals. Rather than speaking of uniform decision intervals $c_v$, then, let us concern ourselves with non-uniform decision intervals: $\hat{c}_v \subset C$. These non-uniform subintervals still form a disjoint partitioning of the unit interval $C = [0,1] \subset \mathbb{R}$, as well as retain the original quantization index $v \in \boldsymbol{V} = \{1, \ldots, \theta\} \subseteq \mathbb{Z}_+$ used earlier. Yet the lengths of these $\hat{c}_v$ are no longer taken as equal to $\theta^{-1}$ (except at the infinitesimal, and fictitious, limit wherein $\theta \to \infty$) because the scalar size now varies with the location of the interval along the overall probability scale $C$. Thus, a new mechanism for determining the endpoints for each $\hat{c}_v$ is needed, which should be the task of the decision weighting function (or at least its perceptually-related components). As before, whether these endpoints are inclusive or not will be dictated by whether $S = s_{up}$ or $S = s_{down}$ proves true. Likewise, the lower and upper endpoints are still indexed by $v - 1$ and $v$, respectively. The main point of departure from the earlier presentation, however, is that such endpoints need transforming.

Let $k$ equal the index value of the endpoint being transformed (i.e., either $k = v - 1$, or else $k = v$). Then one endpoint transformation function $\lambda(\cdot)$ takes as inputs $k\theta^{-1}$ and $\gamma$, and may be written:

(**9**) $$\lambda(k\theta^{-1}, \gamma) = \frac{(k\theta^{-1})^{\gamma}}{[(k\theta^{-1})^{\gamma} + (1 - k\theta^{-1})^{\gamma}]^{\gamma^{-1}}}$$

With $\gamma = 1 - \theta^{-1}$, the above function compresses probability intervals in the middle of the probability scale, but stretches those near the endpoints. Such compressions and stretching becomes more pronounced with small values for $\theta$. And to generate any particular $\hat{c}_v$ using $\lambda(k\theta^{-1}, \gamma)$, one would apply Equation 9 for the two values $k = v - 1$ and $k = v$ (i.e., the lower and upper values from the quantization index, respectively; for ease of notation, let $k_* = v - 1$ and $k^* = v$ in the following, where both $k_*$ and $k^*$ can be substituted for $k$ in Equation 9 above).

To assist readers, Table 4.2 below provide examples of decision intervals that arise when $\theta = 5$.

*Table 4.2: Nonlinear Decision Intervals for $\theta = 5$*

| Index value ($v$) | $v = 1$ | $v = 2$ | $v = 3$ | $v = 4$ | $v = 5$ |
|---|---|---|---|---|---|
| Upper endpoint | 0.242 | 0.406 | 0.561 | 0.732 | 1.000 |
| Length of decision interval | 0.242 | 0.164 | 0.155 | 0.171 | 0.268 |
| Distortion from average | $+0.042$ | $-0.036$ | $-0.045$ | $-0.029$ | $+0.068$ |
| Midpoint of decision interval | 0.121 | 0.324 | 0.483 | 0.647 | 0.866 |

Note: Figures above are rounded to three decimal places.

It is immediately noticeable that decision intervals distorted in this way are lengthier when near the endpoints of the probability scale, and shorter when near its middle. Furthermore, this effect is not symmetric: decision intervals nearer to one are relatively more distorted than those nearer zero.

Additionally, the above method of distorting decision intervals suggests a novel interpretation of a widespread observation from empirical work: the commonplace overweighting of small probabilities and underweighting of large probabilities. Quite simply, if distortions are determined by the value of $\delta$ scaled by the length of the decision interval, and $\delta = 0.5$ is a reasonable 'middle value' for the focusing parameter $\delta$, then, generally speaking, the midpoints of the distorted decision intervals may tend to (averaged across DMs and time) approximate DMs' probabilities at the levels of the probability scale covered by those decision intervals. Put differently, if $\delta$ can fluctuate over time and across people, and DMs sometimes slant probabilities upward, but at other times slant them downward, then if $\delta = 0.5$ is a fairly typical value for $\delta$, then midpoints of decision intervals should roughly approximate DMs' probability judgment when distorted by perceptually-based decision weights. If such is true, then figures in Table 4.2 suggest perhaps one (partial) reason why overweighting of small probabilities and underweighting of large probabilities

transpires: the midpoints of the associated decision intervals move nearer to the middle of the probability scale when distorted (as by Equation 9). The general similarity between leading specifications for DMs' decision weighting functions cause this finding to hold generally across many popular specifications of the functional form: their similar curvatures prompt the same result.

In the next section a related argument will detail how, in particular circumstances, the generality of the phenomenon of overweighting small probabilities and underweighting of large probabilities may be reversed, which may fuel the description-experience gap. For now, however, attention returns to the other modifications of QET needed in order to allow non-uniform decision intervals.

The reworking of decision interval sizes to be non-uniform requires some slight retooling of a few other relationships and specifications taken to underpin QET. First, redefine $S(v, \delta)$ as $\hat{S}(v, \delta)$ by:

(**10**) $\hat{S}(v,\delta) = \hat{s}_{up} = [\lambda(k_*\theta^{-1},\gamma)\delta] + [\lambda(k^*\theta^{-1},\gamma)(1-\delta)]$, if $\left((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta,\delta)\right) \geq \tau$

$\hat{S}(v,\delta) = \hat{s}_{down} = [\lambda(k_*\theta^{-1},\gamma)(1-\delta)] + [\lambda(k^*\theta^{-1},\gamma)\delta]$, if $\left((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta,\delta)\right) < \tau$

In turn, this modification requires another slight rejiggering; let:

(**11**) $\hat{c}_v = \left[\lambda\left(k_*\theta^{-1},\gamma\right), \lambda\left(k^*\theta^{-1},\gamma\right)\right)$, for $\hat{s}_{up}$

$\hat{c}_v = \left(\lambda\left(k_*\theta^{-1},\gamma\right), \lambda\left(k^*\theta^{-1},\gamma\right)\right]$, for $\hat{s}_{down}$

Correspondingly, there is no longer a need for $\theta^{-1}$ directly in the index transmission function. That function can now be more simply specified as:

(**12**) $$\hat{Q}\left(\sigma(A,\neg A)\right) = \left(\sum\nolimits_{v=1}^{\theta} 1_{\hat{c}_v}\left(\sigma(A,\neg A)\right)\hat{S}(v,\delta)\right)$$

Apart from the removal of $\theta^{-1}$ and the slight alterations in notation relative to Equation 4, the index transmission function remains otherwise unchanged. Usefully, no other changes are required to QET in order to move from a linear probability scale to one with non-uniform decision intervals.

It should nevertheless be remarked that $\hat{Q}\big(\sigma(A, \neg A)\big)$ from Equation 12 may still undergo further distortion (due, e.g., to the influences of preferences on decision weighting, such as might be enacted through an elevation parameter, like $\beta$ discussed earlier).[104] But such preference-driven distortions are not treated here. What is seemingly more pertinent is both the inter-subjective and intra-subjective variability that has been observed for the decision weighting function in empirical work. That is, in empirical findings, parameter values (particularly for the discrimination parameter) can vary widely between different DMs, as well as the same DM across time and judgment contexts. This variability has led in part to the proliferation of diverse specifications for suggested decision weighting functions, of which a few were mentioned earlier within this section.

But is it more plausible that people act *as if* judging probability by using different relationships between cognitive limitations and preference across time and context (as well as across individuals)? Or is it simply that their effective cognitive limitations, such as are characterized by $\theta$ and $\delta$, exhibit some (possibly joint) plasticity across DMs, time, and contexts? The suspicion here (particularly given the remarkable structural similarities among many of the functional specifications for those decision weighting functions which demonstrate the strongest and most robust empirical fits) is that variability of the parameters presents by-and-large the more acceptable explanation. In the next section, a case will be made for how parametric variability across contexts may help to explain peculiar experimental findings about the so-called *description-experience gap*.

[104] It is therefore posited here that the order of operations for distorting subjective probabilities proceeds in the following order: first, compression of the probability scale (curvature); second, slanting of probabilities due to effects of quantization; and, lastly, elevation effects as resulting from the interference of preferences upon DMs' probabilities.

## **4.6**\\ Exploring the DEG with QET

The description-experience gap (DEG) was first noted by Hertwig et al. (2004), and involves the following divergence: when DMs have the likelihoods of rare events (i.e., outcomes with small 'objective' probabilities) conveyed explicitly to them (e.g., provided as stated probabilities by an experimenter), then they tend to act *as if* they overweight the probabilities of those rare events; but if DMs are instead required to make judgments under conditions where rare events are possible, but neither the existence nor exact likelihood of such rare events is described to the DMs in advance (and instead the DMs must make judgments based on their own experiences), then, generally, DMs appear to act *as if* underweighting the probabilities of the rare events in question. The DEG has generated substantial interest within psychology over the last decade, and a selection of perspectives on the topic include the following (see also references therein): Fox and Hadar (2006); Rakow, Demes, and Newell (2008); Barron and Yechiam (2009); Ungemach, Chater, and Stewart (2009); Hadar and Fox (2009); Hertwig and Erev (2009); Rakow and Newell (2010); Hertwig and Pleskac (2010); Gonzalez and Dutt (2011); Camilleri and Newell (2009; 2011a; 2011b; 2013; 2013b); Abdellaoui, L'Haridon, and Paraschiv (2011); Lejarraga, Hertwig, and Gonzalez (2012); Barron and Ursino (2013). The topic has also seen heated debate among psychologists as to whether separate theories of judgment are needed to describe DMs' divergent responses under situations when rare-event probabilities are given, versus those when they are not.

A general sentiment among those researchers studying the DEG has been that calls for two separate theories are premature, and that there may be ways of reconciling the two phenomena through a better understanding of the processes by which assessments of probability are actually being made. For example, Fox and Hadar (2006) suggest that the purported underweighting of rare events under judgments from experience was significantly a sampling effect: subjects in experiments were

simply not encountering the small-probability events in question with sufficient frequency to allow them to make accurate assessments of the true likelihoods for those events. Yet other researchers have shown that, even after correcting for sampling-bias influences, there remains a residual tendency for DMs to underweight rare outcomes in experience-based judgments (Barron and Ursino 2013). It is conjectured here that this residual underweighting might be resolvable through considering both the decision settings in which DEG experiments are typically conducted, as well as the interactions between such settings and subject DMs' perceptions of evidence they encounter. That is, a combination of QET and the context of judgment might be able to cast light on the DEG.

To begin, first reflect upon the typical experience-based judgment experiment that is used in DEG research (consult the references mentioned above for various detailed examples of arrangements). Perhaps the most popular setup involves what is comparable to a multi-armed bandit machine from casinos: subjects are able to 'sample' from one or more fixed distributions (the compositions of which are not revealed to the subjects) by, e.g., pressing a button (or one among several buttons) to reveal a payoff drawn from such a distribution. Depending upon the particular setup, subjects may be allowed to sample from such distributions an unlimited number of times (so-called 'free-sampling' paradigms); and oftentimes not only the individual outcomes of such sampling attempts are conveyed to the individual, but also some aggregate statistic describing the total 'earnings' of the DM. (But a complete record of the subject's performance on each trial is not usually furnished.) After the sampling period concludes (in the free-sampling paradigm, such conclusion is often at the subject's discretion), the subject is then prompted to undertake a gamble 'for real' based upon what she has learned across her experience in sampling (in such a way that the experimenter is able to estimate her expectations based upon the succession of choices she makes among gambles).

Experiments with this setup generally find that subjects engage in particularly frugal sampling, even when there is no non-cognitive cost to them in collecting larger samples from the experimental distributions. That is, subjects typically content themselves with small samples, and (as criticized by Fox and Hadar (2006)) routinely suffer from small-sample biases in ensuing judgments. Yet, even when compelled to fully sample from experimental distributions in such a way that the frequencies with which they experience outcomes accurately reflects the underlying distributions, subjects have been shown to persistently exhibit tendencies to underweight the probabilities of rare outcomes. Why might this tendency persist? While the matter cannot be settled fully here, it seems plausible that the explanation lies partly in how the problem is presented to subjects. When faced with such a sparse decision setting as a panel of (very) few buttons, a subject DM may not feel compelled to finely discriminate between outcomes. Drawing upon mechanisms from QET, the initial value of $\theta$ utilized by such subjects (as a consequence of their decision setting) may be quite low. As an illustrative example, say that a subject DM might use a value of $\theta = 4$. Yet, if the DM is tasked with recalling the frequencies of individual outcomes, her working memory resources might be taxed in maintaining many outcomes in memory (Fox and Hadar (2006) present findings that suggest subjects' estimates of the relevant probabilities are accurately calibrated, which suggests that the primary influence upon the observed underweighting should be from decision weights). Especially if the subjects are aware that they must maintain mental tallies of the outcomes which they sample, then their effective values for $\delta$ may be relatively large (e.g. $\delta > 0.5$).[105] Assuming that such a subject draws a modest sample of values (e.g., between 10 and 20, which is commensurate with sample sizes subjects often use when allowed to sample freely),

[105] Meanwhile, it also stands possible that $\delta$ 'expands' over the course of experiments to meet the memory requirements of the subject; this possibility suggests the need to track changes in subjects' expectations and memories of experienced outcomes through the course of sampling, which is not regularly done in DEG experiments.

and that she encounters the 'rare' outcome roughly proportionally to its true likelihood, then QET predicts that she will underweight such rare outcomes. To see this fact, consider a specific example.

As stated, let $\theta = 4$ and $\delta = 0.8$ (a relatively low and high value, respectively), which makes $l(\theta, \delta) = 20$. Assume that the subject DM samples near the maximum of her working memory capacity, and encounters the rare outcome twice in eighteen attempts. In assessing the hypothesis $A$ of the likelihood of encountering the rare event when gambling 'for real', then, the DM should muster as evidence: $\boldsymbol{d}_A = 2$ and $\boldsymbol{d}_{\neg A} = 16$. For simplicity, further presume that the DM encounters no 'surprising' evidence in her sampling activities, such that $\varsigma = 0$.[106] This setup amounts to $\tau \approx 0.22$, and $\left((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta, \delta)\right) = 0.9$, such that $\left((\boldsymbol{d}_A + \boldsymbol{d}_{\neg A}) \,/\, l(\theta, \delta)\right) > \tau$, and therefore $\hat{S}(v, \delta) = \hat{s}_{up}$. Because $\delta$ is relatively large, the experienced probability of $\sigma(A, \neg A) = \sigma(2,16) \approx 0.111$ is slanted downward to $\hat{Q}\big(\sigma(A, \neg A)\big) \approx 0.058$ (or just over half of the objectively-experienced probability). (This example presumes a nonlinear probability scale conforming with Equation 9 of Section 4.5; a linear scale would actually produce slightly more pronounced underweighting, with $\mathrm{Q}\big(\sigma(A, \neg A)\big) = 0.050$.) This simple example therefore suggests that QET may have something to offer to the DEG debate. Yet to be truly convincing as an explanation of the DEG (and potentially other phenomena), QET may need some enhancement.

First, QET is a theory of uncertainty, and it is not yet clear what role it could play as a theory of probability distortion when likelihoods are exogenously given (i.e., under risk). It is plausible, however, that DMs do not take described probabilities at their face values, and instead engage in some form of evidential substitution (as in, e.g.: Viscusi (1989)). Whether or not DMs 'convert'

[106] Notice that, for this particular example, assuming $\varsigma = 0$ is highly innocuous because, for the other values assumed here, a (rather enormous) value of $\varsigma > 122$ would be required to alter the outcome.

given probabilities into some form of evidential equivalents, or whether QET might be extended to risky domains is beyond the scope of this paper, but that possibility invites future research effort.

Second, QET requires empirical validation. In particular, the task of separating its parameters (especially $\theta$ and $\delta$) may require novel experimental techniques. But, like all theories embedded in formal languages, QET can have value beyond strict veracity. That is, all models are necessarily false in the sense that they neglect some features of reality by ignoring or reducing them in order to more deeply explore other features. The main purpose of QET is to accentuate the possible interactions between the nature and mental representation of accessible evidence in the judgment of subjective probabilities under uncertainty. It is hoped that even if it ultimately proves unsatisfactory as a model of such interactions QET at least helps draw attention to their importance.

### **4.7**\\ Summary and Extensions

This paper has presented a descriptive theory, called quantized evidence theory (QET), about how people make subjective judgments of probability based upon uncertain evidence when they face limitations on their attentional and memory capacities. A primary departure by QET from other descriptive theories of probability generation is that it seeks to directly embed a robust regularity, the imprecise mental representation of quantities by approximate number systems (ANSs), from cognitive science into a formal model of how decision-makers express their approximate beliefs as distorted probabilities. Another key advance made by QET is to tie these distortions directly to explicit values for the cognitive limitations that decision-makers face on their working memories. Taken together, these features allow QET to make precise predictions about the ways in which decision-makers will make distorted judgments on the bases of the severity of constraints they face on memory and attention, and the volume of evidence which becomes mentally accessible to them.

And although QET shares features with support theory (Tversky and Koehler 1994; Rottenstreich and Tversky 1997), it improves upon it in several ways. Firstly, it makes explicit those cognitive limitations which support theory relies on for motivating its key predictions. Second, unlike support theory, which can only predict distortions that amount to subadditive probability judgments, QET makes crisp predictions about when decision-makers will exhibit either subadditive or superadditive assessments of subjective probability. Third, QET connects distortions of probability with a decision-maker's relative confidence in the evidence accessible to her (which is benchmarked to her personal cognitive capacity for working memory). Thus, unlike support theory, QET can reflect not only a decision-maker's appraisal of the "strength" of evidence, but also the adjustments which she makes to that appraisal based upon her impressions about the evidence's "weight" (see Griffin and Tversky (1992)). This capability of QET allows it to incorporate confidence effects that do not emerge in many other descriptive judgment theories.

The paper also discussed how QET might be able to offer deeper insights into multiple problems on belief and probability judgments of interest to social scientists. Particularly, it demonstrated how QET may be able to explain not only the typical overweighting of small probabilities in many instances, but also occasional underweighting of rare-event likelihoods in specific circumstances.

Yet, despite its promise in these early demonstrations, much work remains to be done with respect to developing QET into a fully-fledged theory for describing probability judgments. Like all young theories, QET demands empirical verification. This paper has presented a theoretical justification for QET, but, as a primarily descriptive construct, QET should be experimentally tested. In particular, validation of the conjectured values for its free parameters, $\theta$, $\delta$, and $\varsigma$, should be prioritized, and testing of appropriate nonlinear functions for its distortion of the probability scale should be conducted experimentally. But even if QET ultimately proves an untenable model of

human probability judgment, its further study may be worthwhile for its emphasis on deterministic forms of error in human belief generation. That is, QET stands against a growing trend of psychological (and economic) models that seem to posit an outsized explanatory role for stochastic processes in human cognition. While, assuredly, there is often a proper place for error in the modeling of people's thought processes (particularly for general models), it seems in many cases inadvisable to insist upon randomness as the core driver for people's subjective judgment-making.

Thus, QET represents a deterministic class of alternatives that arguably merit deeper investigation. Furthermore, QET can allow a novel approach to the variability in human judgment, and context-dependency of it, that is becoming a main topic for exploration within the social sciences. Specifically, QET permits imprecision and distortion in human judgment not as the consequence of errors drawn from some stable distribution, but instead as deterministic deviations due to temporal and situational (as well as inter-personal) fluctuations among its chief parameters which describe the effective capacity of working memory along two dimensions. And it is further exploration of circumstances under which these two parameters might vary across time, context, and individuals that would seem to warrant the most attention and inquiry in future work on QET.

# CHAPTER 5

# ENTROPY-BASED SOCIAL LEARNING

## **5.1**\\ Introduction

How much do people trust their own beliefs relative to beliefs that they ascribe to others? This question underpins a broad spectrum of contemporary research in economics, particularly when the 'people' concerned are not agents with unlimited cognitive resources, but thinkers with only finite memory and attention at their disposals. The question's answer often depends on various, and variously-shifting, contextual factors (i.e., properties of the specific system to which the beliefs pertain, and a system of which thinkers are themselves components), in addition to what might be considered the 'relative resources' of the thinker herself: her information set and cognitive abilities, compared against information and cognitive abilities she feels that other thinkers possess.

Simon (1955; 1956) long ago identified this dependence by boundedly-rational thinking upon interactions between context and cognitive resources. Nevertheless, the issues of which elements of context matter, and how much they matter, still linger as unresolved for behavioral economics.

Sometimes economic agents' beliefs can meaningfully reshape the contexts which they perceive themselves as facing; then system-level properties of the particular beliefs which an agent believes that other agents hold may become some of the primary elements of context driving how much trust that agent places in her own beliefs, relative to the beliefs she assigns to others. This paper argues that two such systemic properties of beliefs, *agglomeration* and *visibility*, can usefully aid in describing how cognitively-constrained agents dynamically reallocate the amount of trust they place in their own beliefs relative to the amount of trust they place in their beliefs on others' beliefs.

### **5.1.1**\\ Entropy-Based Social Learning

Agglomeration is perhaps more familiarly identified as a property of physical systems (e.g., locations of production, or infrastructure in the real economy), and a conceptual tool used by urban

economists and economic geographers. But this paper demonstrates how agglomeration can also be beneficially repurposed as a force capable of disciplining belief dynamics when agents learn socially. To do so, it develops a model of "circumscribed" social learning by cognitively-constrained agents that is governed by the evolution of *doxastic agglomeration* (i.e., agglomerative dynamics of agents' collective beliefs; Chakrabarti and Roll (1999) employ the term "circumscribed learning" to describe situations wherein agents struggle to infer one another's complete information sets). One novel aspect of the model presented herein is its representation of agglomeration through a decomposition of Shannon's (1948) measure of informational entropy.[107]

Representing agglomeration mechanics with (decomposed) entropy highlights how agglomeration and diversity can reciprocally connect via agents' beliefs. As Hirshleifer and Teoh (2003, p. 27) observe, most economic models of social learning (especially herding) emphasize convergence of agents within an action space. Such models regularly presume identical prior beliefs for agents, with those agents then receiving a diversity of private signals. This paper, however, inverts that setup by presenting a model of entropy-based social learning (EBSL) that permits (extensive) initial belief diversity and learning in the wake of public signals (as in Harris and Raviv (1993), but without requiring that each agent trust exclusively in her own beliefs, or that agents must "agree to disagree" in equilibrium; cf. Hong and Stein (2007)).[108] This initial belief diversity can encompass many circumstances interesting for economists.[109] A useful novelty of the EBSL model resides in how it allows hidden private evidence (or else diverse interpretations of already-public

[107] The extensive reliance by this paper on entropy concepts follows a recent and growing trend in usage of entropy measures within economic and financial research. See especially: Weitzman (2000); Hansen and Sargent (2001; 2008); Sims (2003); Epstein and Schneider (2003); Backus, Chernov, and Zin (2014); Caplin and Dean (2013).
[108] The setup here could, however, also be easily adapted to have agents receive diverse private signals. That possibility would permit further richness in dynamics, but is reserved for future study to streamline present focus.
[109] E.g.: incomplete aggregation of private information à la Diamond and Verrechia (1981) or Romer (1993).

evidence) to get revealed – partly, fully, or not at all – as a product of: the public signal's strength; the initial distribution of agents' beliefs; and visibility of that belief distribution across the agents.

Particularly, the mechanics of this revelation process under the EBSL model can cause endogenous belief changes that might serve as one of the "propagation mechanisms" which Cutler, Poterba, and Summers (1993) assert as deserving greater elaboration in order to explain some of the observably-large reactions by financial markets to minor (or even no) shocks from exogenous news (plus the more general finding that a substantial fraction of returns variation in financial markets cannot be attributed only to arriving of news (Fama 1981; Roll 1988; Shleifer and Summers 1990)).

Additionally, the explicit role which the EBSL model grants to the entropy of beliefs allows it to exhibit 'self-tuning' social learning, whereby relevant agglomeration parameters adjust endogenously.[110] This self-tuning allows the trust which agents place in their own beliefs (relative to what trust they place in the beliefs which they ascribe to others) to vary continuously between two poles: complete trust by each agent in her own beliefs when doxastic agglomeration is at a minimum; and complete trust in others' beliefs when doxastic agglomeration is maximal. Allowing this continuously-variable relative trust accommodates greater richness (and realism) than do those popular models of social learning that adopt more binary approaches, wherein agents trust exclusively in either their own beliefs or others' beliefs, but not in mixtures between both.[111] Indeed, rather than a dichotomy, Charkrabarti and Roll (1999) imply a continuum of trust, bounded on one end by the Keynesian situation of agents caring only about (but distrusting) others' beliefs, and on the other end by 'no-trade' situations (e.g.: Milgrom and Stokey (1982); Tirole (1982)), where agents' full *ex ante* trust in one another's beliefs perpetuates stalemates. The approach here

[110] For another take on models with self-tuning parameters, see the study done by Ho, Camerer, and Chong (2007).
[111] See specifically the classic herding models of Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992).

permits full traversal of that continuum. Also, the EBSL model's agglomeration parameters are not additional degrees of freedom available to the modeler, but are instead statistical properties of the distribution of beliefs itself, and can be pinned down without introducing more free parameters.

Moreover, the specific effects of the evolving agglomeration parameters upon learning within the EBSL model are also influenced by informational and cognitive resources of the agents themselves. (Cognitive constraints and context (agglomeration) therefore interact in shaping the model's learning dynamics, as advocated by Simon (1990).) The informational constraints introduced in the model stem from its *limited visibility* condition: agents' may not accurately know one another's beliefs, and find difficulty in correctly discerning what the "average opinion" of all agents in the system is (or will likely be – a point Keynes (1936) colorfully articulated).[112] This paper argues that limited visibility, when coupled with endogenously-evolving belief diversity, may underlie several curious phenomena within financial markets, including: excess volatility; and anomalous fluctuations which respond to public announcements and non-news events alike.[113]

Alongside informational constraints, the EBSL model also imposes cognitive limits upon agents, which are posited to follow quantized evidence theory (QET) (Rook 2014a). This paper is first to extend QET to settings beyond individual judgments. A core tenet of QET is that human thinkers face limits (due to scarcity of both memory and attention) in how finely they can discriminate between distinct units of evidence; hence, the subjective probabilities which people use to represent their beliefs do not (except in extreme and fictitious cases of cognitively unbounded

[112] Note that most models of social learning implicitly contain some visibility requirement (but whether such requirements demand more or less visibility depends upon the particular model), as might be guessed by the regularly synonymous intents for the terms 'social learning' and 'observational learning' within the literature (cf. Gale (1996)).
[113] The model takes seriously Rabin's (2013a; 2013b) calls for "portable extensions to existing models" (PEEMs), however, and likewise may be applicable to many non-financial problems that are of interest to various strands of social science research, such as, e.g., political lobbying or (the quintessential herding problem of) restaurant selection.

agents) enjoy infinite precision (i.e., they do not use the entire probability scale). Instead QET maintains that people use a 'quantized' scale which compresses the continuous range of probabilities into a finite number of disjoint subintervals. This quantization process usually injects systematic distortions into people's beliefs, and can cause subjective probabilities to be systematically either sub- or superadditive (that is, depart in either direction from summing to one).

Discretization of the probability scale under QET also confers cognitive foundations on the "indivisibilities" (Gale 1996) which have been demonstrated as essential to leading models of herding.[114] Various researchers have noted that herding essentially cannot occur among rational agents without either a sufficiently-discretized action space or signal space.[115] The model here does not require outright that either the action or signal spaces must be discrete, but instead discretizes the belief space (*doxastic space*), which indirectly may, with ample-enough coarsening, cause discretization of the action and signal spaces available to agents (because interpretations of signals, and ponderings of available actions, generally both become filtered through an agent's beliefs).[116]

Helpfully, discretization of the belief space also allows for models of herding (and other types of social learning) wherein the timing of agents' moves is determined endogenously, and any number of agents might move simultaneously.[117] This simultaneity rectifies Shiller's (1995) criticism that

[114] These indivisibilities are essentially what give rise to an effect that Hirshleifer and Teoh (2003) describe as "behavioral coarsening" within models of social-learning situations. The aforementioned cognitive foundations for discreteness in QET may also be a driver for coarse thinking (see, e.g.: Mullainathan, Schwartzstein, and Shleifer (2008)). Gul, Pesendorfer, and Strzalecki (2014) have also produced a 'coarse' model that connects to a competitive equilibrium and extreme prices; but their coarsening mechanism for agents' behaviors differs markedly from QET's.
[115] Eyster and Rabin (2010) nonetheless demonstrate that naïve agents who make biased inferences about the information and beliefs of other agents can herd even when the action and signal spaces are both fully continuous.
[116] Bikhchandani, Hirshleifer, and Welch (1998) have earlier conjectured that discretization in herding and social learning models might have a cognitive or perceptual basis, but they did not fully elaborate upon or explicitly formalize these cognitive mechanisms that give rise to discreteness within a model, as gets done within this paper.
[117] Other papers that utilize different approaches to make timing endogenous in situations of social learning include: Gul and Lundholm (1995); Chamley (2004); and Chari and Kehoe (2004).

(within financial markets) there are oftentimes too many first-movers for decision dynamics to be accurately captured by models of herding in which only a single agent can move in any one period.

Moreover, in many economic models of social learning, agents that have already moved are not allowed to change their minds thereafter and respond to subsequent moves taken by their peers (at least not without some specified process of re-joining a queue, which generates added difficulties in ordering rules). Learning is thus greatly-restricted within such models, which presents problems for findings that belief revisions may be behind several anomalous market phenomena (cf. Timmerman (1993)). EBSL removes some of these difficulties by allowing both simultaneity and revision in the updating of agents' beliefs as (some of) their peers' beliefs become publicly visible.

### **5.1.2**\\ Belief Echoing and Diversity

The potential for simultaneity, combined with the EBSL model's allowing of continuous shifts between the trust agents place in their own and others' beliefs, gives rise to a new notion in social learning: *belief echoing*. Belief echoing is taken to occur when agents alter their own beliefs in response to perceived updating of others' beliefs, which can spawn further updating, and so on (in what can become a back-and-forth process). And while belief echoing can share some features with informational cascades, it stands distinct enough to merit identification as a separate concept.

For one, unlike cascades (which are traditionally taken to occur only when agents completely abandon their own beliefs, and private information on which such beliefs are based, in favor of following a popular reaction), echoing can occur when agents merely decrease the weight on their own beliefs relative to what they feel is the average belief among other agents. Still, belief echoing and informational cascades can coincide under (what is termed here as) *total doxastic agglomeration*. Hence, echoing can be either as dramatic as, or more subtle than, conventional

informational cascades. To be more specific: belief echoing involves a (potentially incomplete) convergence process among agents' beliefs in a social-learning setting (similar to the convergence described by Lee (1993), but pertaining to convergence within a belief space, rather than in an action space). The final sections of this paper demonstrate the mechanics of belief echoing, and suggest echoing may also help to explain some curious empirical phenomena in financial markets.

Another redeeming feature of the entropy-based model of social learning introduced here is its straightforward accommodation of belief heterogeneity among agents. In fact, diversity serves as a general driver of learning in the model (as has been suggested in game-theoretic dynamics by Crawford (1995)). While the model can easily capture the benchmark single (representative) agent setup, it also expands naturally to fit any finite number of heterogeneous agents. For modest diversity levels, the model is highly tractable. Yet for more extensive diversity (e.g., dozens or hundreds of distinct belief 'types') solutions are better obtained via simulation. This paper undertakes both modes of analysis (as is conventionally done with many of the models in economic geography that inspired EBSL). After the model is introduced (in Sections 5.3 and 5.4) 'toy' examples with ten or less types of agents are presented. The final sections involve simulating settings of much richer diversity (over two-hundred types) via use of a novel, large-scale dataset.

### **5.1.3**\\ Fiscal Cliff Dataset

That dataset involves beliefs of institutional investors about the impact of Congressional action in connection with the US 'Fiscal Cliff' episode. Those beliefs involve the (potentially negative) impact of Congressional action on five macroeconomic variables over both the short and long terms. The belief data were collected during December 2012: *before* political uncertainty was (at least temporarily) resolved about the Cliff issue. The dataset thus captures the double-layer of uncertainty – political uncertainty and impact uncertainty – shown by Pástor and Veronesi (2012)

as generally important in describing financial market reactions to policy alterations. The dataset is novel for not only the timing of its collection (before an anticipated event of potentially significant financial and economic consequence), its size (over 800 investors), and its content (actual market professionals' own beliefs, as well as their beliefs about what the 'average' investor in the market believes), but also the unique event to which it is connected. The United States had not previously faced episodes in its history that were directly comparable to the Cliff situation, and the dataset essentially reflects the beliefs of agents that have not garnered extensive prior experience in the 'game' which they are about to engage in. Thus, not only are investors' own beliefs in the dataset diverse, so too are their beliefs about other investors' beliefs. The presence of limited visibility is also suggested by the data, which allows (with some mild assumptions) it to contain yet a third dimension of uncertainty (identified by Avery and Zemsky (1998)): "compositional uncertainty" (i.e., the distribution of agents' beliefs is not common knowledge; this notion seems equivalent to Crawford's (1995) idea of "strategic uncertainty" on beliefs). In total, the Fiscal Cliff beliefs dataset presents an ideal test-bed for demonstrating several of the EBSL model's main properties.

Findings from numerical analysis of the dataset are as follow. First, belief echoing is observable as a response to a wide range of exogenous signal strengths. Second, belief echoing can persist for a substantial duration, and the relationship between echo duration and absolute signal strength does not appear to be strictly monotonic. Third, echoing can eventuate in sizable and abrupt shifts in the distribution of agents' beliefs. Fourth, these shifts can appear well after an exogenous signal arrives, and with little advance warning, yet be pronounced when they finally occur. These results suggest that EBSL may be able to help explain several (otherwise anomalous) market phenomena.

### **5.1.4**\\ Structure

The remainder of this paper traces the following path. Section 5.2 introduces the theoretical foundations of the core model, which is presented in Sections 5.3 and 5.4. Section 5.4 also introduces and initiates investigation on the concept of belief echoing. Sections 5.5 and 5.6 introduce the Fiscal Cliff dataset, apply the model to it, and analyze and discuss results. Section 5.7 concludes by summarizing, and then proposing some directions for future extensions of EBSL.

## **5.2**\\ Theoretical Foundations

A primary psychological impetus behind the EBSL model involves double-application of a potent psychological insight studied by Griffin and Tversky (1992): people tend not to treat strengths and weights of evidence symmetrically. That is, people typically focus unduly on the balance of evidence for and against competing hypotheses (i.e., their strengths), and make improper adjustments for weights behind such bodies of evidence (e.g.: credibility of its sources; or overall volume of evidence). This asymmetry can cause systematic patterns of both over- and under-confidence in judgments of probability. This asymmetry between evidential strength and weight is embedded twice in the EBSL model: once in how agents formulate their own beliefs, and their beliefs about others' beliefs (according to QET); and a second time in how much weight agents assign to their own beliefs relative to the weight they place on beliefs which they ascribe to other agents. The remainder of this section first discusses the latter; it then delves deeper into the former.

### **5.2.1**\\ Weighting Evidence

Imbalance between evidential strength and weight has a natural correlate in Bayesian statistics. It is typically not straightforward (or even possible) to infer the exact weights of evidence reflected in a Bayesian posterior by observing only that posterior in isolation. In general, to make inferences

on the evidential weight supporting a posterior, one also needs two (or more) of: the prior; the likelihood ratio; and the data (i.e., incremental evidence) whose arrival prompted conversion of the prior into the observed posterior. The pith of the matter is: by observing agents' posteriors (or priors) one can witness the strengths of evidence which influence those beliefs; but it is only when beliefs change (or else via some supplemental investigation) that relevant weights of evidence get revealed.[118] In settings wherein experience with such belief dynamics is plentiful enough, agents may come to accurately infer the weights of evidence supporting each other's beliefs. Yet many settings can lack extensive histories of dynamics, and some evidence may remain "hidden" (to use Grossman's (1976) term). Or, at least, agents cannot be sure that no hidden evidence does exist.[119]

The potential for hidden evidence seems to be even greater for settings with *limited visibility*, wherein agents might encounter difficulties in trying to accurately identify one another's current beliefs in the first place, let alone the particular weights of evidence that underpin one another's beliefs.[120] Common knowledge may then be an inappropriate assumption in limited-visibility settings. Under limited visibility, modeling agents' reasonings about strengths and weights of others' beliefs may be beneficial, especially when some agents possess limited cognitive resources.

In describing how agents reason about others' beliefs in limited-visibility settings, people sometimes can obtain fairly accurate (if only intuitive) senses of how much diversity of opinion there is on a particular issue for a group of their peers without necessarily knowing accurately what

[118] This point is made by Aumann (1976, p. 1238) in an example on coin-flips, and he speculates about asymmetric weighting on the basis of evidential visibility. Furthermore, reliance on changes is reflected in claims by various economists that observation of price levels alone is usually insufficient for agents to draw accurate inferences on the informational contents of those prices; see, for example, work by Grossman (1976) and Milgrom and Stokey (1982).
[119] Note that many such interesting settings may not actually lack sufficient histories of belief dynamics outright, but just simply lack specific *relevant* histories (or components thereof). Consider, e.g., the Peso problem (Veronesi 2004).
[120] It is intriguing to find that Alan Turing viewed the idea of weight of evidence as a more intuitively-appealing concept than the amount of information (from the then-nascent field of information-processing); consult Good (1979).

the average opinion is among those peers.[121] In social settings, people may often succeed in detecting the degree of consensus for a group, even if they fail to identify the average belief for it.

As a plausible explanation for observed divergences in the accuracies among people's estimates of the means and dispersions of collective beliefs from social situations, Nisbett and Kunda (1985) suggest that people's "rough-and-ready estimates of dispersion are actually easier for them to compute and less prone to error than their rough-and-ready estimates of means" (p. 302). Furthermore, shared inferences about extents of consensus may lead people to agree on the degree of diversity in opinion for a situation, yet still hold non-equivalent beliefs about average opinion.[122] (Obviously, determining the average opinion becomes easier when agents can communicate directly; the relevant situations herein are those wherein direct communications do not transpire.)

Succinctly, despite shared and identical judgments about belief diversity among a collection of agents, there need not be agreement among those same agents about the particular distribution of beliefs present among themselves. For anything less than perfect consensus, various distinct distributions could permit the same level of diversity when diversity is appropriately measured.[123]

The present paper posits that entropy is one such appropriate tool for measuring belief diversity. Given a suitable space of beliefs (described in Subsection 5.3.2), considerably many belief distributions may share the same entropy value. Yet entropy is minimized (at zero) for perfect homogeneity among beliefs. Such entropy-minimization births a possibility: if an agent knows her own beliefs, and also knows that the entropy of the beliefs for all agents in the economic situation

[121] The two estimates can be psychologically disconnected, as Mullen et al. (1985) show.
[122] Some studies in social psychology find that people can, under some conditions, systematically overestimate the degree of consensus, which gives rise to a 'false consensus effect'. This bias would lead to an even more extreme version of the results reported in this paper, and might be easily modeled by applying constants to the entropy and doxastic agglomeration weights that are components of the mathematical model introduced later in the paper. Likewise, anti-imitation (see Eyster and Rabin (2012)) might also be modelled through such a parametric modification.
[123] Hence, unlike in Lucas (1975), the average belief does not need to become common knowledge among all agents.

she faces (of which her own beliefs are a part) is zero, then she might clearly figure out the average belief among all agents for that situation. In such zero-entropy situations, whether an agent places all weight on her own beliefs or the beliefs she ascribes to others is irrelevant: the two are identical.

Note that, at minimum entropy, the beliefs of all relevant agents become perfectly collocated (i.e., maximally-concentrated) in some space of beliefs. But, although many physical systems are in equilibrium at maximum entropy, it seems more reasonable for economic and social-learning applications that more 'stable' (in a way more concretely described in Section 5.4) configurations obtain when belief entropy is minimized. When minimum entropy among beliefs occurs, it then seems fitting to assert that there exists 'total agglomeration' of beliefs in a relevant doxastic space.

Just (as within standard urban economic thinking under conditions of significant transportation costs) as it becomes more difficult for firms to locate far away from an urban agglomeration unless they have access to special resources, so too might it become hard to resist joining a strong consensus in beliefs. And the stronger is a consensus in beliefs (or, equivalently: the degree of *doxastic agglomeration* in a doxastic space), the harder it may become for an agent to resist joining it. And, as with other types of agglomeration, doxastic agglomerations can become self-reinforcing (via ways comparable to positive-feedback noise mechanisms studied by De Long et al. (1990a)).

Nonetheless, when belief entropy is nonzero it may be that an agent's own belief differs from the average belief among all agents without her correctly recognizing this gap: it might not be 'visible' to her. But if such an agent manages to become aware of some gap between average opinion and her own, then she faces the thorny issue of how to reconcile that disparity.[124] If she were to know the evidence on which others' beliefs are based, then she might combine it with her own evidence

[124] Nevertheless, an agent's impression about any such gap may be inaccurate as a consequence of limited visibility.

and thereby enjoy a Bayesian reconciliation. Yet others' evidence may not be clearly visible either; it could then be imprudent for her to disregard that her own belief dissents from the average belief.

Entropy furnishes a possible resolution. It can indicate how much *potential* evidence is hidden across agents when visibility of one agent's beliefs by other agents in the system becomes limited. With relatively high entropy, considerable diversity of opinion exists, and suggests a substantial volume of evidence is potentially hidden among the beliefs of some agents in the limited-visibility system (regardless of that evidence which is not yet incorporated into the beliefs of *any* agent in the system, i.e., evidence of which *no* agent is then aware).[125] For such situations, an agent may be justified in weighting her own beliefs more strongly than those she feels that others hold. Such justifications may assume different forms (e.g., concerns over the quality of such hidden evidence).

Conversely, when belief entropy is relatively low in a system of agents' beliefs, then there may be little hidden evidence in it. The fact that, in such a system, beliefs have become fully-converged implies that any agent could perhaps improve her own evidential footing by trusting others' beliefs as being based upon well-vetted evidence. And, even if doubts linger about the quality of others' evidence in low-entropy situations, there may be various other deterrents (e.g., reputational considerations, such as Keynes' (1936) admonitions about unconventional failure; see also works by Scharfstein and Stein (1990) and Zwiebel (1995) on reputational concerns, as well as Bernheim (1994) on the possible disutility derived from non-conformity) that could motivate an agent to increase the weight which she allocates to the beliefs she estimates that others hold. As in the more conventional concept of agglomeration economies, agglomeration of beliefs in doxastic spaces can represent a plethora of benefits and/or costs, any of which may make them become self-reinforcing.

[125] Evidence of which none of the agents is aware (at some point in time) might be understood as hidden by 'nature'.

But this self-reinforcing property of doxastic agglomeration does not signify that highly-agglomerated systems of beliefs need be dynamically stable. Indeed, scarcity of hidden evidence in an agglomerated system of beliefs (wherein limited visibility precludes common knowledge among agents of each other's beliefs; note: accurate knowledge of beliefs differs from common knowledge, as Aumann (1976) astutely observes) means that agents collectively might move suddenly to pounce on evidence when it does become exposed (as stands a core tenet of the efficient markets hypothesis). Yet such sudden shifts under conditions of significant agglomeration can (as will be shown) cause processes of belief-updating which depart from Bayesian ideals, and may even lead to crashes (as suggested by Lee (1998)). This influence by agglomeration on belief-updating might explain why social-learning situations can often behave in seemingly wild ways. A proposal here is that this wildness might result from high, albeit imperfect, belief agglomeration.

From the foregoing prologue it should be evident that entropy and doxastic agglomeration might reflect various different motivational forces across economic domains, and thus could enjoy extensive scope as devices for improving models of social-learning phenomena.[126] But prior to treating social learning, some brief reflection on how agents initially form beliefs might be useful.

### **5.2.2**\\ Evidential Quantization

As stated in the introduction, the EBSL model is grounded in quantized evidence theory (QET) (Rook 2014a).[127] QET itself extends support theory (Tversky and Koehler 1994; Rottenstreich and Tversky 1997), which is a landmark psychological explanation of subjective belief formation that reflects many cognitive regularities like framing effects and non-additive judgments of probability.

[126] Indeed, Hirshleifer and Teoh (2003) assert that many such motivational forces and frictions might be present at any one time in market settings, which further hints at the importance of mechanisms (like entropy and doxastic agglomeration) which might be able to realistically capture a generous number of these influencing factors all at once.
[127] This section presents streamlined treatment of QET; see Rook (2014a) for more details and cognitive justifications.

Both QET and support theory assert that people form beliefs not on events, but on descriptions of events: hypotheses.[128] The impetus behind support theory is that not all evidence initially comes to mind for agents when they assess a hypothesis, and that differently-framed hypotheses can result in distinct probability evaluations, even though both describe the same set of states of the world (or events). Support theory is thus motivated by cognitive constraints on memory and attention; it nonetheless fails to model these cognitive limits as definite values. QET makes these values explicit (as described in Section 5.3), and additionally introduces systematic errors (which stem directly from the cognitive limits themselves) into the belief-generation process. This explicitness endows QET with more power as a testable model of belief-formation than support theory (as well as more recent variants of support theory that rely upon distributional assumptions for errors, and can rationalize further judgments just by altering the distributional form that is assumed to be used by the agents (see, e.g., the stochastic variant developed by Brenner, Griffin, and Koehler (2005)).

QET's specific, deterministic error structure causes it to make a novel prediction that distinguishes it from other non-additive probability theories: not only might some probability values be more accessible to agents than others, but some particular (and potentially wide) ranges of probability values will be completely inaccessible to most agents, except those agents without cognitive limits.

The non-stochastic error structure embedded in QET captures a pervasive feature of numerical cognition in humans: the scalar variability of *approximate number systems* (ANSs). It has been experimentally well-established that people's mental representations of quantities are inherently noisy, and that, although people may represent numerical values as continuous (i.e., analog) 'mental magnitudes', they can only compare such magnitudes using limited precision (thereby

[128] Gennaioli and Shleifer's (2010) "local thinking" model likewise utilizes hypotheses as the substrates for beliefs.

making them approximate), which is governed by a ratio-dependent scalar value. This precision-limiting scalar essentially generates an analog-to-digital compressive mapping used by people in comparing quantities; the scalar-variability property effectively discretizes numerical judgment.[129]

To be more concrete: the errors induced by such discretizing conversions tend to: 1) be inversely proportional to the scalar's size; 2) grow proportionally with the scale of the mental magnitudes being measured (the 'size effect'); and 3) grow in inverse proportion to the difference between the two mental magnitudes being compared (the 'distance effect'). These properties of errors due to ANSs seem to pervade human cognition about quantities and numerical judgments (Dehaene 1997; Purves et al. 2008). QET aims to reproduce these regularities via processes of quantization. Quantization is familiar within many signal-processing applications, and is a discretization method that naturally induces some degree of distortion when converting analog signals into digital ones.

Yet this 'lossy' distortion from quantization generally carries worthwhile benefits for memory: it involves resource-conserving reductions in the sizes of outputs, which is important for their subsequent transmission, storage, or further manipulation. And similar savings might drive quantized cognition, which could explain the observed connection between scalar variability and limits to working memory. For example, Gallistel and Gelman (2000) remark that scalar variability in mental representations of quantities appears to "originate within memory itself or in the reading of memory, rather than in the process for measuring duration or counting number…memory signals…are noisy" (p. 62). QET proposes portraying this mental noise via quantization. I now turn to some specific mechanics behind QET that allow its extension into the fuller EBSL model.

[129] The cognitive science literature on scalar variability and ANSs is vast; see Rook (2014a) and references therein.

## **5.3**\\ Core Model

The next two sections build a model of social learning based on doxastic agglomeration/entropy and quantized beliefs among agents. The first subsection (5.3.1) describes processes by which quantized beliefs are formed. It condenses much of the QET model (see Rook (2014a) for a more complete exposition). Subsection 5.3.2 describes the belief spaces (doxastic spaces) in which the learning model operates, while Section 5.4 introduces the dynamical aspects of the EBSL model.

### **5.3.1**\\ Quantized Beliefs

An agent faces the problem of judging the truthfulness of some focal hypothesis $A$ about her world. To do so, she considers the strength of evidence supporting $A$, relative to the strength of evidence in favor of another alternative, mutually-exclusive hypothesis: $B$ (i.e., evidence either supports $A$ or supports $B$; if it supports neither, then it fails to be evidence). Formally, let $\boldsymbol{D}_A$ denote the set of all relevant evidence which an agent without cognitive limitations would view as supporting some focal hypothesis $A$ (rather than its alternative hypothesis: $B$) about her world. Let $\boldsymbol{D}_B$ be the set of all relevant evidence that might support $B$ instead of $A$. Further assume that $\boldsymbol{D}_A \cap \boldsymbol{D}_B = \emptyset$ holds.[130]

A cognitively-unbounded agent should consider all evidence in the set $\boldsymbol{D}_A \cup \boldsymbol{D}_B$ (i.e., total relevant evidence) when judging whether hypothesis $A$ or $B$ is true. $\boldsymbol{D}_A$ and $\boldsymbol{D}_B$ are both assumed to be non-negative (which avoids double-counting evidence), with $\boldsymbol{D}_A \cup \boldsymbol{D}_B \neq \emptyset$. Agents with cognitive limits, however, may not consider all available evidence: only some subset of evidence may be mentally *accessible* by them at any one time, which could reflect either (or both) limits to attention

[130] Brenner, Koehler, and Rottenstreich (2002) remark: "Support theory specifies how support is translated into subjective probability, but is largely agnostic about the source of support for a hypothesis" (p. 492). The present formulation of QET due to Rook (2014a) likewise is generally silent about specific sources for all forms of evidence.

or memory for such agents. Bounded agents might thereby only consider subsets of total evidence. Let $\boldsymbol{d}_A \subseteq \boldsymbol{D}_A$ and $\boldsymbol{d}_B \subseteq \boldsymbol{D}_B$ denote evidential subsets used by agents on $A$ and $B$, respectively.[131]

Moreover, to implement cognitive limitations imposed by ANSs, take all agents (both bounded and unbounded) as making probability judgments not directly from subsets of overall evidence, but instead based upon conversion of those evidential subsets accessible by them. This evidential conversion, which emulates quantization from signal-processing theory, takes as its inputs the cardinalities of the subsets $\boldsymbol{d}_A$ and $\boldsymbol{d}_B$ (let $\widehat{\boldsymbol{d}}_A$ and $\widehat{\boldsymbol{d}}_B$ respectively denote these cardinal values). Assuming that $\boldsymbol{d}_A$ and $\boldsymbol{d}_B$ are both crisp (i.e., non-fuzzy) sets, the restriction $\widehat{\boldsymbol{d}}_A, \widehat{\boldsymbol{d}}_B \in \mathbb{N}$ then applies (and $\mathbb{N}$ is the set of non-negative integers).[132] An agent combines $\widehat{\boldsymbol{d}}_A$ and $\widehat{\boldsymbol{d}}_B$ according to:

**(1)** $$\sigma(A,B) = \frac{\widehat{\boldsymbol{d}}_A}{\widehat{\boldsymbol{d}}_A + \widehat{\boldsymbol{d}}_B}$$

Here $\sigma(A,B)$ can be interpreted as the *intermediate support* for the evident truth of hypothesis $A$ versus $B$. The resulting value of this intermediate support marks an agent's undistorted (subjective) belief in the truthfulness of $A$ (relative to $B$) on the basis of her accessible evidence: $\boldsymbol{d}_A \cup \boldsymbol{d}_B$. Using the cardinalities $\widehat{\boldsymbol{d}}_A$ and $\widehat{\boldsymbol{d}}_B$ (versus some other summary of evidential subsets $\boldsymbol{d}_A$ and $\boldsymbol{d}_B$) thereby causes evidence accessible by an agent who uses quantized thinking to become discretized.

Nevertheless, discretization of evidence by itself is not terribly restrictive, because as long as either or both of $\widehat{\boldsymbol{d}}_A$ and $\widehat{\boldsymbol{d}}_B$ can become sufficiently large in size, then the judgment $\sigma(A,B)$ may still take on any of an infinite number of values within the set of rational numbers $\mathbb{Q}$ along the unit

[131] Note that for cognitively unlimited agents the equalities $\boldsymbol{d}_A = \boldsymbol{D}_A$ and $\boldsymbol{d}_B = \boldsymbol{D}_B$ may hold, but for agents with cognitive limits the subsets $\boldsymbol{d}_A$ and $\boldsymbol{d}_B$ may be proper.

[132] The cardinalities of these subsets might (e.g.) represent distinct cases, or case frequencies, which the agent has in mind, as in Gilboa and Schmeidler (1995), although they need not; the theory here merely requires that these subset cardinalities represent counts of evidential units, which should make clear the connection between the apportioning of these units and the reference class problem from the philosophy of statistics (see Hájek (2007))

interval $C = [0,1] \subset \mathbb{R}$. Thus, while precision from limiting $\sigma(A,B)$ to the set of rational numbers is countable rather than uncountable (as it would be if $\sigma(A,B)$ could take any nonnegative real value), it is infinite. Because infinite precision for $\sigma(A,B)$ requires potentially large values for $\widehat{\boldsymbol{d}}_A$ and $\widehat{\boldsymbol{d}}_B$, we can suspect that bounded agents may suffer from imprecise judgments. (Indeed, even Savage (1964) voiced his skepticism about the realism of infinite-precision probability judgments.)

To help formalize this imprecision, one can set explicit upper bounds for $\widehat{\boldsymbol{d}}_A$ and $\widehat{\boldsymbol{d}}_B$ in order to create a joint expression of limited memory and attention. This joint cognitive limit is one significant source of psychophysical realism from QET, and it can be described via the inequality:

**(2)** $$\left(\widehat{\boldsymbol{d}}_A + \widehat{\boldsymbol{d}}_B\right) \leq l(\theta,\delta)$$

Whereby:

**(3)** $$l(\theta,\delta) \equiv \theta(1-\delta)^{-1}$$

The boundary function $l(\theta,\delta)$ is a composite cognitive constraint due to limits on both: attention, via a *discrimination parameter* $\theta \in \mathbb{Z}_+$; and memory, via a *local-thinking parameter* $\delta \in (0,1] \subset \mathbb{R}_+$.[133] The local-thinking parameter $\delta$ is the same as gets used by Gennaioli and Shleifer (2010), and Bordalo, Gennaioli, and Shleifer (2012). Meanwhile the discrimination parameter $\theta$ is familiar from many theories which involve nonlinear decision weights (e.g., prospect theory (Kahneman and Tversky 1979; Tversky and Kahneman 1992)) wherein the discrimination parameter $\gamma$ often enters as an exponent of the form $\gamma = 1 - \theta^{-1}$. $\theta$ controls how finely agents distinguish between items of evidence in *generating* probability judgments (note: agents may judge exogenous stimuli

[133] It is possible – and likely – that θ and δ may vary across individuals and time; yet to keep exposition and notation clean they are assumed here to be constant across both time and agents.

more finely; i.e., they may be able to more finely differentiate between given evidence than mental evidence, which could create a cognitive distinction separating domains of risk from uncertainty).

Agents with unbounded cognitive resources will have $\theta = \infty$ and $\delta = 1$, whereas for extremely bounded agents $\theta$ and $\delta$ will both approach zero. While $\theta$ and $\delta$ may exhibit some proportional correspondence over intermediate values, appreciable richness in describing judgments by cognitively-limited agents can stem from the independent variability of these parameter values.[134]

But note that the compound limit $l(\theta, \delta)$ is strictly a cognitive 'potential'. Pointedly, some agent might summon considerably less evidence to her mind than her particular limit $l(\theta, \delta)$ may allow (and this gap between summoned evidence and cognitive capacity might often be rational per Sims (2003): a bounded agent may come to improve her judgment by disregarding some irrelevant data).

The intermediate support ratio $\sigma(A, B)$ is not translated directly into subjective probability. Rather, QET presumes the ratio $\sigma(A, B)$ is transformed into a probability via quantization. Quantization processes partition an output space or scale into discrete subintervals. In QET, $\theta$ operates as a *forward quantizer*; it divides the probability scale (i.e., unit interval: $C = [0,1] \subset \mathbb{R}$) into discrete subintervals (*decision intervals*).[135] More specifically, $\theta$ partitions $C$ into disjoint decision intervals $c_v \subset C$, with each $c_v$ of length $\theta^{-1}$.[136] The final probability judgment is only allowed to

[134] Moreover, it has become a popular criticism of some models of bounded rationality which rely on 'cognitive parameters' that these parameters exhibit too much variability across subjects, time, and domains to be of practical use in applications. Even though many theoreticians are loathe to add parameters, it seems a pragmatic alternative here to include two cognitive parameters that separately might vary only slightly, but together may vary far more. To illustrate, consider how the single cognitive parameter $\lambda$ from McKelvey and Palfrey's (1995; 1998) quantal response equilibrium (QRE) model has been empirically demonstrated to often vary massively across individuals and settings. A joint cognitive parameter with only slight variability of the two component parameters may vary equally as substantially as a single parameter $\lambda$, and, from a psychological view, this joint variability appears as more plausible.

[135] Here decision intervals are taken for simplicity to be equally-sized, which makes the process described here (in the terminology of signal-processing theory) one of uniform quantization. Rook (2014a) demonstrates how non-uniform quantization (i.e., subdivision of the probability scale into non-equally-sized subintervals) might be modeled, but that approach is avoided here; it entails difficulties for later entropy methods which cloud exposition.

[136] Aside from decision intervals, the typical quantization process needs two other vital ingredients: a *quantization index*; and a *quantized value* (as output). Typical quantization can be described schematically as follows. An input

be extracted as specific quantized values within any $c_v$, which usually introduces errors. The endpoints of each decision interval are determined by a quantization index: $v \in \boldsymbol{V} = \{1, \ldots, \theta\} \subseteq \mathbb{Z}_+$. The lower endpoint of each $c_v$ is then the value $\left((v-1)\theta^{-1}\right)$; and the upper endpoint is $(v\theta^{-1})$. $\theta^{-1}$ serves as an incrementing scalar (as in ANSs).[137] An *index transmission function* completes the mapping $\sigma(A,B) \rightarrow C$ and produces a quantized probability $Q\big(\sigma(A,B)\big)$. One candidate specification for such an index transmission function is:

$$\textbf{(4)} \qquad Q\big(\sigma(A,B)\big) = \theta^{-1}\left(\sum\nolimits_{v=1}^{\theta} 1_{c_v}\big(\sigma(A,B)\big)S\right)$$

In Equation 4 above, $1_{c_v}(\cdot)$ serves as an indicator function, where $1_{c_v}\big(\sigma(A,B)\big) = 1$ if $\sigma(A,B) \in c_v$, but $1_{c_v}\big(\sigma(A,B)\big) = 0$ if $\sigma(A,B) \notin c_v$. This indicator function identifies which particular decision interval $c_v$ corresponds to the received input ratio $\sigma(A,B)$, and thus operates as a decision interval *lookup function*. Upon identifying the appropriate decision interval $c_v$, an *encoder mapping* step then picks out a value $v$ from the quantization index which corresponds to that $c_v$.[138]

value is employed to retrieve a target decision interval, with each such interval having associated to it an integer-valued quantization index number. The step is called *encoder mapping*. The returned value from the quantization index is then used to transform the input value into a quantized output value through an *index transmission* step. Rook (2014a) argues that quantization is analogous to the mental mapping (i.e., conversion) process $\sigma(A,B) \rightarrow C$ that gets performed by agents who face cognitive constraints imposed by $l(\theta,\delta)$ (i.e., a joint limit upon memory and attention).

[137] $\theta^{-1}$ can also be seen to indicate a particular agent's level of probabilistic sophistication. Additionally, for each decision interval $c_v$, only one endpoint is inclusive; the other is open. Nevertheless, whether this inclusive endpoint is the upper or lower endpoint only matters when $\delta = 1$.

[138] The behavioral reality reflected here is that people typically express probability judgments as point values, even though these point values may represent ranges of probabilities that an agent would find acceptable for the problem faced. QET thus accepts the notion of multiple priors. According to some multiple priors models, agents select, from among the multiple probabilities which they entertain at any given time, one single probability as being representative of their beliefs for the judgment problem faced (see, e.g.: Gilboa and Schmeidler (1989); Hansen and Sargent (2001; 2008)). QET embeds a similar logic; hence another step beyond encoder mapping is needed. That is, encoder mapping simply identifies a value $v$ from the quantization index which is the upper endpoint of the decision interval that gets selected as containing $\sigma(A,B)$. But $(v\theta^{-1})$ might not be the most 'representative' value of $c_v$ for an agent. Some adjustment may be needed to increase its representativeness. Thus, the quantization index must be adjusted to a quantized value through index transmission to obtain this representative value. The terms $\theta^{-1}$ and $S$ from the index transmission function in Equation (4) are able to accomplish such an adjustment.

Just as before, $\theta^{-1}$ is merely an incrementing scalar determined by the degree of cognitive discrimination that a subject can muster in thinking. The term $S$ is a 'switching' function equaling:

**(5)** $s_{up} = \delta(v-1) + v(1-\delta) = v - \delta, \qquad \text{if } \left(\left(\widehat{\boldsymbol{d}}_A + \widehat{\boldsymbol{d}}_B\right) / \, l(\theta,\delta)\right) \geq \tau$

$s_{down} = (v-1)(1-\delta) + v\delta = v + \delta - 1, \qquad \text{if } \left(\left(\widehat{\boldsymbol{d}}_A + \widehat{\boldsymbol{d}}_B\right) / \, l(\theta,\delta)\right) < \tau$

And, in turn:

**(6)** $c_v = [(v-1)\theta^{-1}, v\theta^{-1}), \qquad \text{for } s_{up}$

$c_v = \left((v-1)\theta^{-1}, v\theta^{-1}\right], \qquad \text{for } s_{down}$

The degree of *quantization distortion* resulting from quantization is: $\left(Q\left(\sigma(A,B)\right) - \sigma(A,B)\right)$. And quantization distortion is bounded by: $\left(Q\left(\sigma(A,B)\right) - \sigma(A,B)\right) < \theta^{-1}$. Thus, $\theta$ serves to both induce *and* bound the degree of error in subjective probability distortion an agent can make.[139]

The selection of either $s_{up}$ or $s_{down}$ then becomes a matter of which endpoint of $c_v$ (whether open or closed) garners more weight from the local-thinking parameter $\delta$ relative to the opposite endpoint of $c_v$. $\delta$ thereby allocates relative weights to both of the lower and upper endpoints of $c_v$.

Whether $S = s_{up}$ or $S = s_{down}$ holds is determined by the size of the quantity $\left(\left(\widehat{\boldsymbol{d}}_A + \widehat{\boldsymbol{d}}_B\right) / \, l(\theta,\delta)\right)$ relative to a threshold value $\tau \in (0,1] \subset \mathbb{R}_+$.[140] $\left(\left(\widehat{\boldsymbol{d}}_A + \widehat{\boldsymbol{d}}_B\right) / \, l(\theta,\delta)\right)$ describes how near to her cognitive limit an agent is operating. When $\left(\left(\widehat{\boldsymbol{d}}_A + \widehat{\boldsymbol{d}}_B\right) / \, l(\theta,\delta)\right)$ is high, the agent has

[139] Note that this degree of error precludes influences due to preferences, which may further affect how much the 'probabilities' which agents ultimately act upon diverge from the value $\sigma(A,B)$. It seems somewhat appropriate to classify quantized beliefs as decision weights (as appear in, e.g., prospect theory (Kahneman and Tversky 1979; Tversky and Kahneman 1992), or salience theory (Bordalo, Gennaioli, and Shleifer 2012)). This paper follows Wakker's (2004) suggestion that, under uncertainty, such weights may include influences caused by cognitive limits.

[140] Two instances exist wherein $s_{up} = s_{down}$. First, $s_{up} = s_{down}$ holds when $\delta = 0.5$. (For expositional clarity, this paper is only concerned with that case, and $\delta$ will be taken as fixed at 0.5 hereafter.) Second, $s_{up}$ and $s_{down}$ will converge as $\theta \to \infty$ (and $\theta^{-1} \to 0$). This second instance of $s_{up} = s_{down}$ with infinite $\theta$ describes a fictional agent with perfect cognitive abilities for discrimination. For all other combinations of values for $\theta$ and $\delta$, $s_{up} \neq s_{down}$ holds.

nearly (or even exactly) as much evidence $\left(\widehat{\boldsymbol{d}}_A + \widehat{\boldsymbol{d}}_B\right)$ accessible in mind as her cognitive capacity $l(\theta, \delta)$ can accommodate. But when $\left(\left(\widehat{\boldsymbol{d}}_A + \widehat{\boldsymbol{d}}_B\right) / l(\theta, \delta)\right)$ is low, an agent possesses spare mental capacity. $\tau$ demarcates where the ratio $\left(\left(\widehat{\boldsymbol{d}}_A + \widehat{\boldsymbol{d}}_B\right) / l(\theta, \delta)\right)$ is relatively high or low, and this highness or lowness can entail differing impacts for agents who possess distinct values for $\delta$ (see Rook (2014a) for discussion). QET proposes the threshold $\tau$ to be a function of the accessible evidence itself. A candidate for the functional form of $\tau$ is:

(7) $$\tau\left(\widehat{\boldsymbol{d}}_A, \widehat{\boldsymbol{d}}_B\right) = 1 - \frac{|\widehat{\boldsymbol{d}}_A - \widehat{\boldsymbol{d}}_B|}{\widehat{\boldsymbol{d}}_A + \widehat{\boldsymbol{d}}_B}$$

The above specification is essentially Bordalo, Gennaioli, and Shleifer's (2012; 2013) salience function. [141] This form embeds both size and distance effects that might arise from a difference in the discretized values of evidential strengths, $\widehat{\boldsymbol{d}}_A$ and $\widehat{\boldsymbol{d}}_B$.[142] The threshold function $\tau\left(\widehat{\boldsymbol{d}}_A, \widehat{\boldsymbol{d}}_B\right)$ is basically a distillation of the total cardinality-weighted symmetry of evidence accessible to an agent. One interpretation is: the more asymmetric is the cardinality-weighted balance of evidence, the more overconfident an agent will tend to be in her judgments. Conversely, with greater symmetry her confidence lessens. As the salience functional specification measures degrees of asymmetry it thus seems well-suited to govern the direction in which the probabilities get slanted.

The original presentation of QET (Rook 2014a), however, only treats situations wherein agents entertain a single focal hypothesis at any one time. Moreover, it only concerns judgments by

[141] Quantized beliefs are somewhat comparable to "robust beliefs" (Hansen and Sargent 2001; 2008). Agents with robust beliefs are permitted to doubt their own subjective probability estimates, and resultantly come to surround these doubted probabilities with an 'entropy cloud', and the size of the cloud varies with the degree of uncertainty (this cloud may be considered as a set of beliefs the agent would find plausible). Quantized beliefs, meanwhile, reverse that logic. Rather than formulating a doubtable probability and buffering it with a set of 'nearby' probabilities, quantized beliefs essentially take the cloud as primitive (that is, its size is dictated by $\theta^{-1}$), and the doubtable belief emerges from the cloud (as a function of $\delta$, $\tau$, and $\left(\widehat{\boldsymbol{d}}_A + \widehat{\boldsymbol{d}}_B\right)$ ). Thus, while similar, quantized beliefs differ from robust ones.

[142] Within psychophysics, size effects involve increasing error as estimated magnitudes increase; distance effects involve increasing inability to distinguish between two magnitudes 'near' to one another (see e.g. Purves et al. (2008)).

individual agents and does not address belief dynamics (learning). The next two subsections extend the original version of QET to incorporate multiple hypotheses and agents, plus (social) learning.

### **5.3.2**\\ Doxastic Spaces and Entropy

To begin, allow the number of focal hypotheses $A$ which may be considered by any agent for any one judgment problem to be $F \in \mathbb{Z}_+$, such that any specific focal hypothesis under consideration $A_f$ can be indexed by $f \leq F$, where $f \in \mathbb{Z}_+$. For an unbounded agent $F \to \infty$ is typical.[143] Meanwhile, for bounded agents, $F$ should be relatively small (e.g., a single-digit integer).[144] These $F$ hypotheses taken together can be understood as comprising a hyperspace with $F$ dimensions.[145]

Call such a hyperspace a *doxastic space*, and, moreover, recognize that an agent may (for purposes herein) become represented by her position inside such a doxastic space through a vector combination of her subjective beliefs regarding the truthfulness of the $F$ focal hypotheses. If such beliefs are posed as subjective probabilities, then a given doxastic space might be described by $C^F$ (where, again: $C = [0,1] \subset \mathbb{R}$). But if such probabilities are generated by quantizing beliefs, then we might instead describe such a space by $\theta^F$. If the same values of $\theta$ and $F$ are shared by all agents in a particular judgment setting, then we can place all of those agents into a single doxastic space $\theta^F$, wherein the distances between the positions of any two agents within the space reflects the 'distance' between their respective beliefs for the $F$ focal hypotheses that are being considered.

[143] Or, it is at least typical if one assumes that each hypothesis is specific enough to represent a possible state of the world, with there being an infinity of states; after a hypothesis has been refuted by the evidence, however, even an unbounded agent might then ignore it going forward and merely entertain unrefuted ones, as in Gilboa, Samuelson, and Schmeidler (2013), or Kurz (1994a; 1994b).

[144] While it is tempting to make $F$ some function of the cognitive limit $l(\theta, \delta)$ (or else one of its component parameters, $\theta$ or $\delta$) in order to reduce the number of model parameters, for extra flexibility $F$ is kept here as an independently-determined (and exogenous) parameter separate from other limits. Also realize that, if $F$ is infinite, then this setup becomes equivalent to that of an agent who can construct an infinite state space (see further remarks in Gilboa (2009)).

[145] Note that this hyperspace is contained inside the more economically-familiar space $\mathbb{R}^F$.

Let $N = \theta^F$ indicate the doxastic space's 'size'. Because both $\theta$ and $F$ can only assume the values of positive integers, the size of any given doxastic space in terms of the number of distinct (discrete) positions $p_n$ within it can be described by the value of $\theta$ raised to the power $F$, such that the position index $n$ within $\theta^F$ obeys $n \leq N$. This setup makes it straightforward to explore multidimensional distributions of beliefs within such a space, and specifically to measure entropy.

Entropy is a workable aid for analyzing distributions of agents within doxastic spaces because it addresses simultaneously two considerations about diversity: 1) the number of positions within the space occupied by any agents whatsoever (i.e., spaces with positive occupancy); and 2) the relative sizes of the occupancy proportions among those positions having positive values for occupancy (see Page (2011)). Using the preceding notation, one can measure Shannon's (1948) informational entropy (denoted herein by $H_N$) for any given doxastic space according to the following equation:

**(8)** $$H_N = -\sum_{n=1}^{N} p_n(\ln(p_n))$$

Shannon's entropy supplies an indication of heterogeneity in terms of not only the number of different types (i.e., distinct positions occupied by agents within a doxastic space), but also the proportional distribution of agents across types.[146] By using the Shannon specification, the approach here follows Harsanyi (1967) because it sees an agent's type as her 'portfolio' of beliefs.

Helpfully, Shannon's measure is also additively decomposable into a direct measure of inter-agent agglomeration in beliefs (i.e., the total extent to which all agents collocate with one another inside the doxastic space), plus an indicative measure – *juncture* – of inter-dimensional structure between the beliefs of single agents across the different hypotheses being considered which comprise the doxastic space. When total agglomeration occurs among the beliefs of a collection of agents within

[146] Following convention, where $p_{n \leq N} = 0$, $\ln(p_n) = 0$ is taken to hold; see, e.g., Hanson (1998) and Sims (2003)).

such a space, then all agents become collocated and no belief diversity exists for agents in that space. This *total doxastic agglomeration* minimizes $H_N$ for the space. Formally, total doxastic agglomeration occurs when: $p_j = 1$ for some position $j = n$ within $\theta^F$, *and* $p_k = 0$ holds for all $j \neq k \leq N$. One plausible interpretation of total doxastic agglomeration among a large collection of agents implies that all of those agents' beliefs are constrained in (effectively) the same way(s).[147]

The force of juncture may be understood as manifesting itself in attempts by agents to coherently reconcile those beliefs which they hold across various focal hypotheses (even if such hypotheses are themselves objectively independent). Nonexistence of juncture between beliefs on (e.g.) two focal hypotheses may indicate that an agent might change her beliefs on one hypothesis, but not the other, without experiencing cognitive dissonance. Conversely, high relative juncture should signify that changing one belief and not the other may cause incoherence for the agent (and thus dissonance), as some dependencies exist between the two hypotheses within the agent's mind.[148]

A decomposition of Shannon's entropy $H_N$ is able to capture measurements of both doxastic agglomeration and juncture inside $\theta^F$. Denote by $H_N^{max}$ the maximum hypothetical Shannon entropy with number of positions $N$ in the doxastic space specified, such that $H_N^{max} = \ln(N)$. Also let $\widetilde{H}_{N,X}$ denote the measured (or estimated) value for Shannon's entropy in $\theta^F$ with $X$ agents distributed across $N$ discrete positions in the doxastic space (with $x$ denoting some individual agent). $\widetilde{H}_{N,X}$ can be usefully interpreted as the extent to which detectable constraint is lacking, due

[147] Or else they may be treated as so in the standard "as if" sense of Friedman (1953).

[148] Hence, juncture can (like agglomeration) be seen as potentially contributing toward constraining those positions in a doxastic space which an agent might occupy. In short, it would, given a degree of belief in a particular hypothesis, dictate the degrees of freedom that an agent should enjoy in changing her beliefs about other hypotheses that she is also considering. For a bounded agent, changes to one belief may only be coherent if they are "compatible with the entire web of beliefs held by the individual" (Kahneman and Tversky 1974, p. 185). Thus, a bounded agent may (in the absence of new information on several beliefs) insufficiently (relative to Bayesian ideals) update a given belief to maintain coherence with other beliefs that she holds.

to absence of either (or both) of doxastic agglomeration or juncture, in a particular doxastic space (i.e., it increases with lessening constraint). Under this view, measured *syntropy* $\tilde{\Gamma}_{N,X}$ is the degree of observable structure from the (additive) combination of both measured agglomeration and juncture. It can be defined residually according to the following:

(**9**) $\quad H_N^{max} - \tilde{H}_{N,X} \equiv \tilde{\Gamma}_{N,X}$

Syntropy (which is essentially negative measured entropy, or degree of detectable structure in a doxastic space) is therefore bounded from below by zero and from above by $H_N^{max}$. It can be decomposed into the sub-measures of doxastic agglomeration $\tilde{\xi}_{N,X}$ and juncture $\tilde{\eta}_{N,X}$ according to:

(**10**) $\quad \tilde{\Gamma}_{N,X} = \tilde{\xi}_{N,X} + \tilde{\eta}_{N,X}$

Agglomeration and juncture can also be alternatively (and perhaps more intuitively) exposited by envisioning $\theta^F$ as a multidimensional contingency table (akin to those described by Kullback (1959)).[149] For simplicity, imagine that such a table has only two dimensions; ergo $F = 2$ (the further extension of the following to any finite number of dimensions is straightforward). Let $J$ and $K$ be the number of rows and columns (respectively) in the table, and let $\theta = J = K = 3$ (at least for the example here), such that $N = 9$ (i.e., there are nine 'cells' in the table).[150] (Realize that the table need not be symmetric, but that this election is made from convenience.) Take $p_{j,k}$ as the proportion of the $X$ agents whose beliefs cause them to occupy the cell in the $j^{\text{th}}$ row and the $k^{\text{th}}$ column of the table (with: $j \leq J$; $k \leq K$; and $j, k \in \mathbb{Z}_+$). Furthermore, let the marginal sums across such proportions within the table be reflected within the following equations:

[149] The following alternative definition was suggested within sociology by Martin (1999a; 2002). He uses the terms "consensus" and "tightness" instead of agglomeration and juncture, and his treatment is not cognitively motivated.

[150] In what follows (and where no confusion might arise), subscripts are generally suppressed for measured and hypothetical values of entropy, syntropy, doxastic agglomeration, and juncture in order to reduce the notational clutter.

(**11.1**) for rows: $p_j = \sum_{k=1}^{K} (p_{j,k}), \quad \forall j$

(**11.2**) for columns: $p_k = \sum_{j=1}^{J} (p_{j,k}), \quad \forall k$

Using both (11.1) and (11.2), agglomeration $\tilde{\xi}$ and juncture $\tilde{\eta}$ can then be calculated according to:

(**12**) $\tilde{\Gamma} = \ln(N) + \left( \sum_{j=1}^{J} \sum_{k=1}^{K} \left( p_{j,k} \left( \ln \left( p_{j,k} \right) \right) \right) \right)$

(**13**) $\tilde{\xi} = \ln(J) + \ln(K) + \left( \sum_{j=1}^{J} \left( p_j \left( \ln \left( p_j \right) \right) \right) \right) + \left( \sum_{k=1}^{K} \left( p_k (\ln(p_k)) \right) \right)$

(**14**) $\tilde{\eta} = \tilde{\Gamma} - \tilde{\xi}$

Realized measurements for agglomeration and juncture (juncture may be helpfully interpreted as "constraint beyond the marginals" (Martin 1999a)) can indicate strength and structure of constraints on agents' beliefs, when such beliefs are measured collectively within a doxastic space.[151] These constraints can specify belief dynamics in the space in ways dovetailing with QET.

## **5.4**\\ Learning Dynamics

To better illustrate the social-learning dynamics inside a given doxastic space for agents who engage in evidential quantization, first consider how a single agent might update her beliefs upon receiving some fresh evidence. Assume that the basis for the agent's updating is her intermediate

[151] One immediate objection that might be levelled against the foregoing decomposition technique is that it does not actually measure belief interdependencies and compatibilities at the intrapersonal (individual agent) scale. This criticism is surely valid and extends the conjecture made by Kahneman and Tversky (1974) that "no simple procedure" can exist to empirically measure or verify inter-belief compatibility at the individual level. I maintain that their conjecture still holds true, and that its continued validity highlights another primary merit of the entropy decomposition here: it circumvents the impossibility of such a measurement by instead focusing on statistically-implied interdependencies at the collective level, which is also seemingly more important within many social settings anyway.

support ratio $\sigma(A,B) = \frac{\widehat{\boldsymbol{d}}_A}{\widehat{\boldsymbol{d}}_A + \widehat{\boldsymbol{d}}_B}$ defined by Equation (1). The right-hand side of (1) is the same as that for the mean from the beta distribution (where $\widehat{\boldsymbol{d}}_A$ and $\widehat{\boldsymbol{d}}_B$ are respectively substituted for the two typical shape parameters, $\alpha$ and $\beta$, which sculpt the beta distribution).[152] Require that $\widehat{\boldsymbol{d}}_A, \widehat{\boldsymbol{d}}_B > 0$ holds true. This positivity requirement restricts $\widehat{\boldsymbol{d}}_A$ and $\widehat{\boldsymbol{d}}_B$ to the set of positive integers. The intuition here is that this learning model describes situations where agents do not think any single hypothesis can be either certain or impossible, based upon accessible evidence (or the lack of it).[153]

### **5.4.1**\\ Anticipatory Updating

Assume that fresh evidence arrives at time $t$ in the form of an exogenous signal $g_t \in \mathbb{Z}$.[154] The signal $g_t$ is perceived by all agents to contribute (integer-valued) evidence to either the focal hypothesis $A_f$ (if $g_t > 0$) or its corresponding alternative $B_f$ (if $g_t < 0$) for each hypothesis $f$ under consideration by the agents ($g_t$ is taken to not convey any evidence if it equals zero). Given the indicated correspondence between $\sigma(A,B)$ and the beta distribution, we might define the agent's updated intermediate support ratio (in response to $g_t$), $\sigma\big(A_f, B_f\big)|g_t$, by the following:

**(15)** $$\sigma\big(A_f, B_f\big)|g_t = \frac{\widehat{\boldsymbol{d}}_A + (z g_t)}{\widehat{\boldsymbol{d}}_A + \widehat{\boldsymbol{d}}_B + |g_t|}, \quad \text{where: } z = 0 \text{ if } g_t < 1, \; z = 1 \text{ if } g_t \geq 1$$

[152] The shape parameters $\alpha$ and $\beta$ of the beta distribution are not, however, typically restricted to integer values, even though both are strictly positive. Both $\alpha$ and $\beta$ are usually set equal to 1 when reflecting an ignorance prior.

[153] Such a positivity requirement avoids dogmatic priors, and is similar to the presumption in the paper by Radner (1972, p. 290), that each agent will not attach zero-probability to any observable event (to complete connection with Radner, take hypotheses here to describe only observable events, even when such observability might be only hypothetical). Moreover, we might become further convinced that the beta is a suitable distribution to employ here because it is the conjugate prior distribution to the binomial distribution. If an agent is trying to judge if evidence counts or not in contributing to whether a focal hypothesis $A$ is either true or false, then this binomial distribution assumption for the arrival of fresh evidence would seem to be fitting. Recall as well that the standard beta distribution is asymmetric and skewed for all instances wherein its two shape parameters are non-equal (i.e., where $\alpha \neq B$ holds).

[154] We could allow a continuous-valued signal, but as all bounded agents within the model will discretize any accessible evidence anyway, we can prevent hassles by modeling $g_t$ as the agent would perceive it in her evaluation of evidence: as discrete. Nonetheless, the evidential content of $g_t$ is not capped here, and it might be arbitrarily large.

In (15), as $|g_t|$ increases, the signal strengthens (either positively or negatively). For purposes of this model, the arrival of the signal $g_t$ is taken to be shared knowledge across all of the agents. That is, $g_t$ is an exogenous, public signal that is the same across all agents, and all agents have knowledge that all other agents have also received the identical public signal during time period $t$.

Assume that each agent forms quantized beliefs prior to the signal's arrival not only for herself; she also forms beliefs on what the average belief among all of her peers – *including herself* – is.[155] Denote these quantized priors for agent $x$'s own beliefs and her beliefs on the average belief of all agents as $Q_{x_0}\left(\sigma\left(A_f, B_f\right)\right)$ and $Q_{x_{\bar{X}}}\left(\sigma\left(A_f, B_f\right)\right)$, respectively (where subscripted $x_o$ and $x_{\bar{X}}$ respectively stand for agents $x$'s own beliefs and belief on the average belief across all agents). No *a priori* restrictions are placed on the relationship between $Q_{x_0}\left(\sigma\left(A_f, B_f\right)\right)$ and $Q_{x_{\bar{X}}}\left(\sigma\left(A_f, B_f\right)\right)$, so that they may be equal, but likewise may be distinct, as long as they obey the foregoing procedure of quantization, and respect cognitive constraints imposed by: $l(\theta, \delta)$ and $F$.

Similarly, let $Q_{x_0}\left(\sigma\left(A_f, B_f\right)|g_t\right)$ and $Q_{x_{\bar{X}}}\left(\sigma\left(A_f, B_f\right)|g_t\right)$ be the quantized posteriors of agent $x$ for her own beliefs and the beliefs she has about the average belief across all agents immediately after $g_t$'s arrival, but *before* she observes any reaction by others agents to the arrival of the signal $g_t$.[156]

[155] That is, each agent forms two sets of beliefs, which may be equal to one another, but need not. This dual-beliefs setup is similar to that of Chakrabarti and Roll (1999), although the combination procedure used here is very different.

[156] As noted by Hoch (1987), a person's clearest knowledge is usually about her own beliefs, and her estimation of others' beliefs often stems from some adjustment made to her own beliefs. An agent's judgment of others' evidence thereby need not reflect anything but adjustment to her evidence. Furthermore, in adapting QET to a dynamic setting (as is undertaken here), the threshold $\tau$ might be modified to reflect receipt of $g_t$, as this signal's arrival should be understood as increasing accessible evidence. We might therefore elect to describe the threshold which controls the quantizing of evidence according to the following relationship: $\tau\left(\widehat{\boldsymbol{d}}_A, \widehat{\boldsymbol{d}}_B, g_t\right) = 1 - \frac{|\widehat{\boldsymbol{d}}_A - \widehat{\boldsymbol{d}}_B + g_t|}{\widehat{\boldsymbol{d}}_A + \widehat{\boldsymbol{d}}_B + |g_t|}$.

If it is true that $Q_{x_0}\left(\sigma(A_f, B_f)\right) = Q_{x_{\bar{X}}}\left(\sigma(A_f, B_f)\right)$, then agent $x$ should encounter little trouble in updating her beliefs immediately following $g_t$'s arrival; it simply holds for her that $Q_{x_0}(\sigma(A_f, B_f)|g_t) = Q_{x_{\bar{X}}}(\sigma(A_f, B_f)|g_t)$. Yet what happens for those agents for whom it is true that $Q_{x_0}\left(\sigma(A_f, B_f)\right) \neq Q_{x_{\bar{X}}}\left(\sigma(A_f, B_f)\right)$? It would seem the case that, given homogenous $\theta$ and $\delta$ across all agents (as will, for the present time, be assumed here), $Q_{x_0}\left(\sigma(A_f, B_f)\right) \neq Q_{x_{\bar{X}}}\left(\sigma(A_f, B_f)\right)$ would only hold if agent $x$ felt that some private evidence had not become public, and either she possesses it and (some) others do not, or else (some) others possess it and she does not. Likewise, for any two agents $x$ and $y$: if $Q_{x_0}\left(\sigma(A_f, B_f)\right) \neq Q_{y_o}\left(\sigma(A_f, B_f)\right)$, then this non-equivalency is presumed to stem from differences in accessible evidence between $x$ and $y$. (Note: differing interpretations of evidence may operate as a form of differing access, even if they do not *per se* involve differences in private information; this view is contrary to De Long et al. (1990b)).

Moreover (and to return to the problem set out in Section 5.2.1), even if the true average of beliefs across all agents were visible to all agents, then they still might lack common knowledge of what evidence each other's beliefs were based upon (consult, for instance, Monderer and Samet (1989)).

Furthermore, with limited visibility, for any two agents $x$ and $y$, the following non-equivalency might hold (although need not): $Q_{x_{\bar{X}}}\left(\sigma(A_f, B_f)\right) \neq Q_{y_{\bar{X}}}\left(\sigma(A_f, B_f)\right)$. Yet, as was stated in Section 5.2.1, people may hold accurate views about extents of consensus (or else disagreement) in social contexts without necessarily knowing precisely what average opinion across everyone actually is.

This sense of the degree of consensus among agents can be captured via doxastic agglomeration $\tilde{\xi}$. *Total doxastic agglomeration* occurs when: $\tilde{\xi} = H_N^{max}$. And agglomeration is at a minimum when: $\widetilde{H}_N = H_N^{max}$. For simplicity, it is assumed here that $\tilde{\xi}$ is publicly-observable (or at least

accurately inferable) by all agents (as are $\widetilde{H}_N$ and $\tilde{\eta}$). Moreover these values are taken to be the same across all agents.[157] More specifically, assume that the values of $\tilde{\xi}$, $\widetilde{H}_N$, and $\tilde{\eta}$ which pertain to the distribution of all agents' beliefs within the doxastic space prior to the arrival of the signal $g_t$ (say, at time $t-1$; i.e., with a one-period lag) are 'accessible statistics' known to all agents (as in, e.g.: Jehiel and Koessler (2008); Van Huyck, Battalio, and Beil (1990; 1991)). They might, e.g., be extracted from "buzz" that circulates within financial settings (Clark and Monk 2013).[158]

Recall that $\tilde{\xi}$ measures agglomeration in a doxastic space, and $\widetilde{H}_N$ indicates dispersion of (or lack of constraint on) agents' beliefs in that space, but that these statistics do not always suffice for determining the location of the average belief. Given the foregoing notation and rationale, one learning rule for agents' beliefs in a doxastic space in response to arrival of some public signal is:

$$\textbf{(16)} \quad \pi_{x_o,t}\left(\sigma\left(A_f,B_f\right)\middle|g_t\right)=\left(\frac{1}{\widetilde{H}_{N,t-1}+\tilde{\xi}_{t-1}}\right)\left[\left(Q_{x_o,t}\left(\sigma\left(A_f,B_f\right)\middle|g_t\right)\right)\widetilde{H}_{N,t-1}+\left(Q_{x_{\bar{X}},t}\left(\sigma\left(A_f,B_f\right)\middle|g_t\right)\right)\tilde{\xi}_{t-1}\right]$$

In the above equation, the measured entropy and doxastic agglomeration parameters, $\widetilde{H}_{N,t-1}$ and $\tilde{\xi}_{t-1}$, are those from before the signal's arrival (i.e., from the initial distribution of agents' own quantized priors, $Q_{x_o,t-1}\left(\sigma\left(A_f,B_f\right)\right)$ for all $x$, in time period $t-1$). Thus, Equation (16) essentially describes a rule for changing beliefs whereby each agent updates both her own beliefs, and her beliefs about the average of all agents' beliefs, in response to the public signal $g_t$, but then takes a convex combination of these two updated beliefs to be her 'amalgamated' posterior belief $\pi_{x,t}\left(\sigma\left(A_f,B_f\right)\middle|g_t\right)$ in response to the signal. The weighting of the two component posteriors,

[157] Realize that it is not being suggested here that agents actually compute entropy, agglomeration, or juncture values, but rather that these values proxy their impressions of disagreement/consensus and coherence among beliefs across all agents in the system (and that such proxy impressions are 'public' due to their being equal across all of the agents).
[158] Financial media might in this way substantially influence perception of the level of consensus in markets. Mullainathan and Shleifer (2005), DellaVigna and Kaplan (2007), and Gentzkow and Shapiro (2010) find that media sources may benefit from presenting slanted views. The implications of the EBSL approach proposed here suggest that such biased views may potentially be less impactful than could be a misrepresentation of overall consensus levels.

($Q_{x_o,t}(\sigma(A_f, B_f)|g_t)$ and $Q_{x_{\bar{X}},t}(\sigma(A_f, B_f)|g_t)$), to achieve an amalgamated posterior is undertaken by multiplying her own (quantized) posterior by the lagged entropy value (relative to lagged agglomeration) and her belief about the average belief of all agents (including herself) by the lagged agglomeration value (relative to lagged entropy). Such "anticipatory" updating that includes expectations about others' updating thus emulates the anticipation dynamics in the model of social learning due to Gul and Lundholm (1995).[159] Given this process for anticipatory updating of agents' beliefs (i.e., before agents observe others' updates) a few specific cases merit comment.

First, when *total doxastic agglomeration* exists prior to the arrival of $g_t$ (i.e., when $\tilde{\xi}_{t-1} = H_N^{max}$) each agent will place positive weight only on $Q_{x_{\bar{X}},t}(\sigma(A_f, B_f)|g_t)$. As this situation is one in which all agents are collocated inside the doxastic space, then each agent knows (or should know) the true location of $Q_{x_{\bar{X}},t}(\sigma(A_f, B_f)|g_t)$ (but this knowledge need not be common), and total doxastic agglomeration is preserved from $t-1$ through to $t$: all agents remain collocated in the doxastic space before *and* after the public signal's arrival. Notice that this collocation will often be more likely for smaller doxastic spaces (that is, it will be more likely as values for parameters $\theta$ and $F$ are decreasing). Further note that (16) collapses to standard Bayesian updating by a representative agent (albeit with quantized beliefs) when total doxastic agglomeration occurs during time $t-1$.

Second, when maximum entropy exists (i.e., when $\widetilde{H}_{N,t-1} = H_N^{max}$) before $g_t$'s arrival, every agent will place positive weight exclusively on her own quantized posterior beliefs, $Q_{x_o,t}(\sigma(A_f, B_f)|g_t)$. Nevertheless, this situation only obtains under one condition: when agents

[159] It also follows from the classical distinction between behavior in initial and later periods noticed by Nagel (1995) in some of the original studies of "guessing" games (which EBSL seems to be capable of modeling). Likewise recognize that this conditioning of beliefs by an agent on both her own personal evidence and some distillation of others agents' evidence also features in Radner's (1979) approach to rational expectations equilibria (although he uses market prices as the mechanism through which others agents' evidence gets publicly conveyed).

are uniformly-distributed throughout the doxastic space. With non-uniformity, all agents will place some positive weight on the belief $Q_{x_{\bar{X}},t}\left(\sigma\left(A_f, B_f\right)\middle| g_t\right)$.[160] Notice that if there are too few distinct 'types' of agents within the doxastic space, then this uniformity condition will remain unfulfilled.

Such a failure of uniformity might occur, for instance, when any one discrete location in the doxastic space does not have positive occupancy. The general rarity of this condition that all locations within a doxastic space have equal *and* positive occupancies implies that agents will tend to place some weight on $Q_{x_{\bar{X}},t}\left(\sigma\left(A_f, B_f\right)\middle| g_t\right)$ when forming an amalgamated posterior $\pi_{x,t}\left(\sigma\left(A_f, B_f\right)\middle| g_t\right)$.[161] (I assume that $\pi_{x,t}\left(\sigma\left(A_f, B_f\right)\middle| g_t\right)$ becomes agent $x$'s 'own' belief for time $t$; this assumption becomes important for explaining belief dynamics after anticipatory updating.)

Realize that from $t-1$ until $t$ the values of $\widetilde{H}_N$ and $\tilde{\xi}$ may well change as agents update their quantized beliefs in response to $g_t$, and anticipate how other agents are likely to respond to $g_t$. Thus, the respective weights on each agent's own beliefs and her beliefs about the average of all agent's beliefs will change endogenously in response to updating of agents' beliefs inside the doxastic space (i.e., these parameters will 'self-tune' through time). Recall as well that these are not free parameters within the model, but are summary statistics of the initial belief distributions. The only truly exogenous parameters in the EBSL model here are the cognitive ones: $\theta$, $\delta$, and $F$.

Further recognize that, although agglomeration will tend to increase within the doxastic space as a response to an exogenous signal, the path of its increase can exhibit some interesting dynamics. These dynamics can become particularly complex if we permit agents to react to changes in one

[160] Failure of this non-uniformity condition may bear upon the results of experiments by Dominitz and Hung (2009).
[161] It is interesting to observe how this general tendency for agents to nearly always place some weight on the perceived consensus jibes with findings from social psychology (see, e.g.: Hoch (1987)) that people regularly place insufficient weight on their own personal judgments in social settings (i.e., those in which they are aware others will also operate).

another's beliefs after period $t$ (but without arrival of any other exogenous public signals). I turn to that possibility shortly, but first it stands important to consider the purpose that $\tilde{\eta}$ comes to serve.

### **5.4.2**\\ Juncture

Juncture $\tilde{\eta}$ has a crucial role to play as a 'friction' against agglomerative forces. Agglomeration and juncture might be likened to the centripetal and centrifugal forces, respectively, that are described as highlights from the new economic geography of Krugman (1998b). Recall that $\tilde{\eta}$ indicates (at the collective level) the degree of inter-dimensional dependencies among agents' beliefs within the doxastic space. With relatively high $\tilde{\eta}$, and, (e.g.) $F = 3$, then, *ceteris paribus*, an agent receiving a signal that pertains to only one hypothesis under consideration, but not the other two, may endure cognitive dissonance from updating only the one belief affected by the signal, but not the other two. Conservatism is henceforth presumed to operate here: agents do not update any dimension of their beliefs unless they receive a signal relating to that dimension, which may account for some market under-reactions (see Barberis, Shleifer, and Vishny 1998), or confirmatory bias (Rabin and Schrag 1999). If agents receive a signal on some, but not all, dimensions, then they reduce the degree of updating on belief dimensions which the signal affects.

The foregoing gives rise to a need for the concept of the public signal's dimensionality, which gets denoted here by $\Delta_{g_t}$, with $F \geq \Delta_{g_t} \in \mathbb{N}$. Assume that $\Delta_{g_t}$ is identical across all of the agents (for it is a property of $g_t$, which is assumed to be public). $\Delta_{g_t}$ governs how strongly $\tilde{\eta}$ impedes the belief-updating process. It might therefore be used to define a (linear) 'penalty' function, such as:

(**17**) $\quad \Psi = \frac{F-\Delta_{g_t}}{F-1}$

(Other such penalty functions may well be applicable, but the one above keeps things pleasingly simple.) This penalty function may impede belief updating by interacting with $\tilde{\eta}_{t-1}$ according to:

$$\textbf{(18)} \qquad \pi'_{x_o,t}\left(\left(A_f, B_f\right)\middle|g_t\right) = Q\left[Q_{x_o,t-1}\left(\sigma\left(A_f, B_f\right)\right) + \left[1 - \left[\Psi\left(\frac{\tilde{\eta}_{t-1}}{\ln(N)}\right)\right]\right]\left[\pi_{x_o,t}\left(\sigma\left(A_f, B_f\right)\middle|g_t\right) - Q_{x_o,t-1}\left(\sigma\left(A_f, B_f\right)\right)\right]\right]$$

The end value emanating from (18) is the quantized output of its right-hand-side components (denoted through bracketing of the right-hand-side expression with a quantization function: $Q[\cdot]$).

From both (17) and (18) above, recognize that, when $\Delta_{g_t} = F$, the impact of $\tilde{\eta}_{t-1}$ upon $\pi'_{x_o,t}\left(\left(A_f, B_f\right)\middle|g_t\right)$ is nil, and therefore $\pi'_{x_o,t}\left(\left(A_f, B_f\right)\middle|g_t\right) = \pi_{x_o,t}\left(\sigma\left(A_f, B_f\right)\middle|g_t\right)$ (i.e., Equation (18) collapses to Equation (16)).[162] Yet, for lower-dimensional signals (i.e., $\Delta_{g_t} < F$), $\tilde{\eta}_{t-1}$ induces what might be analogized by a modeling "trick" used in the new economic geography: "iceberg transportation costs" (see, e.g.: Fujita, Krugman, and Venables (1999)). With a high value of $\tilde{\eta}_{t-1}$ and low signal dimensionality $\Delta_{g_t}$ (less than $F$), some of the updating that would have occurred "melts away".[163] (Realize that, under total doxastic agglomeration, $\tilde{\Gamma} = \tilde{\xi}$, and thus $\tilde{\eta} = 0$, which implies that maintaining coherence is unimportant in such situations, and may describe a general facet of crowd psychology in social learning: it can be more unstable than is individual learning.)

[162] That Equation (18) can collapse to (16) is important to recognize in light of the fact that Equation (16) itself can collapse to standard Bayesian updating by a representative agent who uses quantized beliefs. Yet, those quantized beliefs may become the same as standard, continuous Bayesian beliefs if the discrimination parameter $\theta$ is made sufficiently large. Hence, the model here can bridge a wide spectrum of circumstances, making it satisfyingly portable.

[163] Note that a high value for $\tilde{\eta}$ which is shared by all agents might well arise from a common framework for understanding a problem, even if agents disagree on the problem's appropriate resolution. To invoke an (only somewhat stylized) example, imagine that all agents in a system understand a (American) political problem, involving three hypotheses, in terms of a Republican-Democrat dichotomy of perspective. It is plausible that a group of agents who share this 'red-blue' framing might also share a single value for $\tilde{\eta}$ because their beliefs should tend to settle along a linear continuum from Republican to Democratic. Thus, even with significant diversity of belief among agents (i.e., relatively low $\tilde{\xi}$), a predominant framework for understanding may impose exogenous constraints on beliefs (which need not even require beliefs on distinct hypotheses to be mutually reconciled in agents' minds, other than, e.g., by the justification: "that's just what Republicans believe…"; such justifications show the normative role that beliefs may play in helping to determine identities for purposes of economic decision-making: cf. Akerloff and Kranton (2000)).

### **5.4.3**\\ Echoing Mechanics

Treatment of learning dynamics has so far only considered the anticipatory updating by agents in response to the signal $g_t$ (à la Gul and Lundholm (1995)): it has only covered updating undertaken by agents before they observe any updating that their peers also undertake. This initial stage of learning is anticipatory in the respect that agents not only integrate evidence from the signal into their own troves of accessible evidence, but also take account of how they anticipate the average belief among all agents in the system will respond. Under full visibility where agents can perfectly see (or infer) one another's beliefs this anticipation may be accurate, so long as each agent can adequately approximate the weights of prior evidence used by other agents in forming their priors.

Yet, if either the strengths or weights of evidence are not perfectly visible for all agents, then some rich dynamics may arise if agents are able to (at least partially) observe others' anticipatory updating. Specifically, less-than-accurate anticipatory updating can precipitate subsequent *reactive* social learning wherein agents realize where (at least some of the) other agents' beliefs are actually located in the doxastic space. Secondary –and higher – reactions to reactive updating may then occur in such a way that one may say the arrival of $g_t$ can continue 'echoing' over several periods beyond just the period of anticipatory updating through an iterated tâtonnement process.[164]

As might be guessed, this echoing/ tâtonnement process can have a reflexive effect on entropy, agglomeration, and juncture: as agents reactively update their beliefs, they become redistributed around the doxastic space, which can change the entropy, agglomeration, and juncture parameters endogenously, and alter the learning rules agents face through shifting weights they place on their own beliefs versus that belief which they believe is the average belief being held across all agents.

[164] Note that Radner's (1979) notions of rational expectations equilibria also involve (series of) tâtonnement processes.

Thankfully, echoing dynamics can be captured through a fairly simple and iterative (re)application of the earlier learning rules (Equations (16), (17), and (18)), and a few other minor assumptions.[165]

First (as indicated earlier), assume that after the initial updating (i.e., in period $t$ after the arrival of $g_t$) each agent's own belief about each hypothesis $A_f$ under consideration becomes $\pi'_{x_o,t}\big((A_f, B_f)\big|g_t\big)$ (as a result of applying (18)). The values of entropy, agglomeration, and juncture then become updated as $\widetilde{H}_{N,t}$, $\tilde{\xi}_t$, and $\tilde{\eta}_t$ to reflect the updated distribution of all agents' own beliefs inside $\theta^F$. These new values for $\widetilde{H}_{N,t}$, $\tilde{\xi}_t$, and $\tilde{\eta}_t$ are assumed to become publicly known by each agent, such that all agents share an updated impression of these values at time $t$ (these parameters thus become accessible statistics to agents in the model, as with Jehiel (2005)).

Second, assume that only those agents who change their quantized beliefs as a response to receiving $g_t$ will become visible to all other agents. That is, only belief *changes* are publicly-visible to all.[166] Recognize, however, that, as agents here use quantized beliefs, changes pertain to movements from one decision interval to the next, and that the evolution of beliefs is not smooth for bounded agents.[167] And it occurs in discrete jumps when it does happen. Moreover, it is assumed that the ending beliefs (and not the starting beliefs) of such changes are what become visible to all agents (i.e., the beliefs of all agents $x$ for whom $\pi'_{x_o,t}\big(\sigma\big(A_f, B_f\big)\big|g_t\big) \neq Q_{x_o,t-1}\left(\sigma\big(A_f, B_f\big)\right)$ for period $t$). Thus, the average across all of these freshly-changed quantized

[165] This iterative application, including (what is equal to) public announcement of updated entropy, agglomeration, and juncture statistics, is similar to McKelvey and Page's (1986) setup, but with (homogeneous) cognitive constraints.
[166] This visibility-through-change assumption intends to reflect, for market settings, the notion that price is based on marginal trades. In non-price settings a departure from business-as-usual might be captured due to its visibility by competitors (as in, e.g., Caplin and Leahy (1994)).
[167] While there is not space here to connect this non-smoothness to stickiness in prices and information (see, e.g. Mankiw and Reis (2002)) that possibility still remains attractive to entertain and could be pursued in other future work.

beliefs becomes the new estimate of average belief (across all agents) for every agent during $t$. This new average belief $E_{\bar{X},t}\left(\sigma\left(A_f, B_f\right)\middle| g_t\right)$ across all agents within the system is describable as:

(**19**) $$E_{\bar{X},t}\left(\sigma\left(A_f, B_f\right)\middle| g_t\right) = \frac{\sum_{x=1}^{X} \left[\pi'_{x_0,t}\left(\sigma\left(A_f,B_f\right)|g_t\right)\right] \mathrm{I}\left(\left(\pi'_{x_0,t}\left(\sigma\left(A_f,B_f\right)|g_t\right) - Q_{x_0,t-1}\left(\sigma\left(A_f,B_f\right)\right)\right)\neq 0\right)}{\sum_{x=1}^{X} I_x}$$

where: $I_x = 1$ if $\left(\pi'_{x_0,t}\left(\sigma\left(A_f, B_f\right)\middle| g_t\right) - Q_{x_0,t-1}\left(\sigma\left(A_f, B_f\right)\right)\right) \neq 0$;

$I_x = 0$ if $\left(\pi'_{x_0,t}\left(\sigma\left(A_f, B_f\right)\middle| g_t\right) - Q_{x_0,t-1}\left(\sigma\left(A_f, B_f\right)\right)\right) = 0$

(Notice: $E_{\bar{X},t}\left(\sigma\left(A_f, B_f\right)\middle| g_t\right)$ is an un-quantized value as it is observed, and not mentally generated.) This new average belief is then reintroduced into Equation (16) to produce the value:

(**20**) $$\pi_{x_o,t+1}\left(\sigma\left(A_f, B_f\right)|g_t\right) = \left(\frac{1}{\widetilde{H}_{N,t}+\tilde{\xi}_t}\right)\left[\left(\pi'_{x_0,t}\left(\sigma\left(A_f, B_f\right)|g_t\right)\right)\widetilde{H}_{N,t} + \left(E_{\bar{X},t}\left(\sigma\left(A_f, B_f\right)|g_t\right)\right)\tilde{\xi}_t\right]$$

To account for effects of the new juncture value $\tilde{\eta}_t$, the output of Equation (20), $\pi_{x_o,t+1}\left(\sigma\left(A_f, B_f\right)\middle| g_t\right)$, then enters Equation (18) as input according to the following (note that the value $\Psi = \frac{F-\Delta_{g_t}}{F-1}$ will remain unchanged from Equation (17), as no new signals are taken to arrive):

(**21**) $$\pi'_{x_o,t+1}\left(\left(A_f, B_f\right)|g_t\right) = Q\left[\pi'_{x_o,t}\left(\sigma\left(A_f, B_f\right)\right) + \left[1 - \left[\Psi\left(\frac{\tilde{\eta}_t}{\ln(N)}\right)\right]\right]\left[\pi_{x_o,t+1}\left(\sigma\left(A_f, B_f\right)|g_t\right) - \pi'_{x_o,t}\left(\sigma\left(A_f, B_f\right)\right)\right]\right]$$

(Where the right-hand side is transformed by the quantization function $Q[\cdot]$ from Equation (4).) Application of (21) may result in new quantized values of beliefs for period $t+1$, which would cause the values of $\widetilde{H}_{N,t+1}$, $\tilde{\xi}_{t+1}$, and $\tilde{\eta}_{t+1}$ to be different from those in $t$. These changes would drive the beginning of the next iteration of the echoing/tâtonnement process with a new value for (19), and subsequent application of Equations (20) and (21).[168] The process eventually stops

[168] Notice that this reactive adjustment process emulates learning direction theory (Selten and Stoecker 1986).

whenever no further updating occurs. Note: this process's duration and the final distribution of the agents is determined entirely by: the initial belief distributions; the size of the public signal $g_t$; and the cognitive parameters $\theta$, $\delta$, and $F$. Recognize as well that this stripped-down, iterative process for belief updating is obviously stylized in the respect that it does not explicitly account for preferences, endowments, or other elements that characterize genuine learning processes. Yet the absence of those elements is intentional for clarifying the study here. One could layer-in such elements later (or, to start the other way around, introduce this social learning process into other models wherein those elements are already established; Rabin (2013a; 2013b) advocates the latter approach). Yet those complications are not undertaken here in favor of expositional simplicity.[169] I now turn to a few examples that help to illustrate some core properties from the full EBSL model.

### **5.4.4**\\ Basic Examples

The foregoing EBSL model of social learning remains tractable enough to be worked out by hand for situations involving comparatively few distinct types of agents and small doxastic spaces. This subsection demonstrates the EBSL model's behavior in several such simplified situations in order to build some deeper intuitions for its performance in richer settings which are explored thereafter.

To set the stage for subsequent sections, assume that the examples hereafter apply to investors who are trying to gauge the impact of an uncertain governmental intervention on the future state of the economy. And, as in the model by Pástor and Veronesi (2012), these investors must form beliefs on not only what the impact of such an intervention might be, but also what the likely intervention will itself be. Knowing that the economic impact of this uncertain governmental intervention may

[169] See Vives (1995) for a comparable simplifying approach.

harm their investments, all investors in this example contemplate the hypothesis that the chosen (but as yet unknown) intervention will have a negative impact upon future economic performance.

As the intervention in these examples pertains to a novel political situation, there is limited historical evidence on which investors can form beliefs on the hypothesis's truthfulness. They worry that they may either lack some key evidence that others possess, or that their own interpretations of what little evidence is accessible to them may be misguided in some way. But they also worry that other investors may likewise have erroneous beliefs. They express these uncertainties, to the best of their abilities, as beliefs both for themselves, as well as for what they believe to be the 'average' investor across all of the investors in the market (including themselves).

These beliefs follow the quantization process described earlier, because each investor only entertains five categories of belief on the hypothesis that the eventual intervention will have negative future economic consequences: 'strongly disagree'; 'disagree'; 'neutral'; 'agree'; and 'strongly agree'. These quantized beliefs therefore partition the probability scale into five equal subintervals, each of length $0.2$. Moreover, assume that these quantized beliefs correspond to the following subintervals for ranges of probability: $[0, 0.2)$; $[0.2, 0.4)$; $[0.4, 0.6)$; $[0.6, 0.8)$; $[0.8, 1)$.

Thus, for each investor, $\theta = 5$. Further, let $F = 1$ (they only consider the single hypothesis of a negative economic outcome from whatever intervention obtains; this assumption of a single hypothesis allows us to initially ignore juncture) and $\delta = 0.5$ (a 'neutral' value, as demonstrated earlier). Thus, $l(\theta, \delta) = 10$ and $\theta^F = 5$. And, to begin, let there be two types of investor, $x_1$ and $x_2$ (hence, $X = 2$), with investors being homogenous (for purposes here) within each type, and otherwise indistinguishable apart from type (i.e., each investor 'matters' as much as any other investor). (Throughout this paper, neither entry nor exit is presumed to occur within the market;

nor are agents presumed to engage in any form of direct communication with one another.) To start, let the number of investors be equal across both types. Furthermore, let both types of agents in this example form their quantized beliefs according to the table (Table 5.1) appearing below.[170]

*Table 5.1: Initial Evidence and Belief Distribution (Pre-Signal) for 2-Agent Example*

| | **Own Belief** | | | **Belief about Average Belief** | | |
|---|---|---|---|---|---|---|
| **Type** | $\hat{d}_A$ | $\hat{d}_B$ | Quantized | $\hat{d}_A$ | $\hat{d}_B$ | Quantized |
| $x_1$ | 1 | 9 | 0.1 | 1 | 9 | 0.1 |
| $x_2$ | 9 | 1 | 0.9 | 9 | 1 | 0.9 |

Given the five 'cells' in the single-dimensional doxastic space for this example, we find $x_1$ and $x_2$ have diametrically-opposed personal beliefs on the focal hypothesis: $x_1$ believes it true with a $90\%$ degree of faith, but $x_2$ places only $10\%$ faith in it (the neutral value of $\delta = 0.5$ causes both agents to place their beliefs in the center of the relevant decision interval). Each agent also suffers from poor visibility on the average belief in the market, and believes that the average matches her own.

If we were to depict this initial belief distribution via a tabular format, we might do so as follows:

| | **Quantized Probability** | | | | |
|---|---|---|---|---|---|
| **Belief** | **0.1** | **0.3** | **0.5** | **0.7** | **0.9** |
| Own | $x_1$ | | | | $x_2$ |
| Average | $x_1$ | | | | $x_2$ |

Given that each investor type accounts for half of all investors, calculating entropy and agglomeration values is straightforward: the two occupied cells each contribute values of

[170] Realize that either (or of both) $x_1$ or $x_2$ might be cast as 'noise traders'. While the commonplace distinction between noise traders and 'smart money' is unimportant for the example here, notice that one could easily layer-in the traditional noise-versus-smart money dichotomy (see, e.g., Shleifer and Summers (1990)) by specifying the location within the doxastic space in question that corresponds to the smart-money beliefs (conditional upon the available/accessible evidence). All agents not occupying that smart-money position inside the doxastic space would then (by default) be noise traders on the basis of their beliefs (albeit to varying degrees relative to their distance from the smart-money position). The EBSL model can thus embrace and enrich more conventional noise-trader approaches.

$-p_n(\ln(p_n)) = -0.5(\ln(0.5)) \approx 0.35$ toward entropy. The three unoccupied cells contribute nothing. Thus, $\widetilde{H}_{N,t-1} \approx 0.69$ (due to rounding). $H_N^{max} = \ln(5) \approx 1.61$, so syntropy equals $\tilde{\Gamma}_{t-1} \approx 0.92$. With no juncture (the doxastic space here is single-dimensional), agglomeration equals syntropy, and hence: $\tilde{\xi}_{t-1} \approx 0.92$. How might these values then respond to new evidence?

At time $t = 0$ both $x_1$ and $x_2$ receive a public signal $g_t$ (say, for example, that a particular government intervention is announced, and this action conveys exogenous information to the investors, although the economic impacts of such an intervention may remain unclear; the signal's absolute strength relative to initial evidence reflects the degree to which the economic impact of the chosen intervention remains unclear). And both know that the other receives the same $g_t$ at $t = 0$.[171] Say that $g_t = 2$, so that both investors initially update the two sets of beliefs which they hold (i.e., their own beliefs and their beliefs about average market belief), as per Table 5.2 below.

*Table 5.2: Updated Evidence and Belief Distribution (Post-Signal) for 2-Agent Example*

| | **Own Belief** | | | **Belief about Average Belief** | | |
|---|---|---|---|---|---|---|
| **Type** | $\hat{d}_A$ | $\hat{d}_B$ | Quantized | $\hat{d}_A$ | $\hat{d}_B$ | Quantized |
| $x_1$ | 3 | 9 | 0.3 | 7 | 5 | 0.5 |
| $x_2$ | 11 | 1 | 0.9 | 7 | 5 | 0.5 |

To complete the anticipatory updating in response to $g_t = 2$, both investors must now combine these beliefs. Each does so by applying the weight $\widetilde{H}_{N,t-1}/(\tilde{\xi}_{t-1} + \widetilde{H}_{N,t-1}) \approx 0.57$ to her own updated quantized belief, and the weight $\tilde{\xi}_{t-1}/(\tilde{\xi}_{t-1} + \widetilde{H}_{N,t-1}) \approx 0.43$ to her belief about the average belief in the market (i.e., the average of her peer's belief and her own), and then taking the summed product to be her amalgamated posterior (which she then quantizes). These

[171] Indeed, anticipation of the arrival of some public announcement may be one of the sources of heterogeneity, for an impending news event may motivate traders to try to generate new information by, e.g., undertaking analysis and research not previously done to "get an edge" (see, for example: Cao, Coval, and Hirshleifer (2002)).

amalgamated posteriors become (approximately) 30% and 90% for $x_1$ and $x_2$, respectively, (which also happen to match the quantized values). Updated beliefs for each agent appear below.

| | Quantized Probability | | | | |
|---|---|---|---|---|---|
| **Belief** | **0.1** | **0.3** | **0.5** | **0.7** | **0.9** |
| Own | | $x_1$ | | | $x_2$ |
| Average | | $x_1$ | | | $x_2$ |
| Amalgamated | | $\mathbf{x_1}$ | | | $\mathbf{x_2}$ |

As only $x_1$ changed her quantized beliefs in response to the (modest) signal $g_t = 2$, it is only her belief change that becomes visible to the market. The average belief change for $t = 0$ therefore becomes the (un-quantized) value of her amalgamated posterior: 30%. This updated value conveys information to both investors. Given this visible belief change, both investors now engage in reactive updating and apply Equation (20) to merge their own amalgamated posteriors with the updated (visible) value for the market average: 30%. Given that $x_1$ and $x_2$ remain in separate cells within the doxastic space, the entropy and agglomeration values remain unchanged from $t = -1$, even though both $x_1$ and $x_2$ calculate the average market belief as 30%. Applying Equation (20) we find that the amalgamated beliefs for each in $t = 1$ become 30% and 56% for $x_1$ and $x_2$, respectively. After quantizing, these beliefs then become (respectively) 30% and 50%, which indicates that (although $x_1$ does not change her beliefs between the periods $t = 0$ and $t = 1$), $x_2$ changes her amalgamated posterior, and the new market average therefore becomes that value.[172]

[172] The behavior of $x_2$ in response to $x_1$ seems to emulate the behavior observed in the famous Asch (1951; 1952; 1956) experiments from social psychology, wherein some 'visible' opinion overpowers other individuals' prior beliefs. On classic experiments similar to Asch's but where participants cannot see one another, see Crutchfield (1955).

| | Quantized Probability | | | | |
|---|---|---|---|---|---|
| **Belief** | **0.1** | **0.3** | **0.5** | **0.7** | **0.9** |
| Own | | $x_1$ | | | $x_2$ |
| Average | | $x_1, x_2$ | | | |
| Amalgamated | | $\mathbf{x_1}$ | $\mathbf{x_2}$ | | $\mathbf{x_2}$ |

Reapplying Equation (20) (again, with unchanged entropy and agglomeration values, because the investors remain in separate cells in the doxastic space), we find that the investors' new amalgamated posteriors in $t = 2$ become $41\%$ and $53\%$, for $x_1$ and $x_2$, respectively; these values now both quantize to $50\%$, and cause $x_1$ and $x_2$ to collocate in the doxastic space: total doxastic agglomeration takes hold thereafter. As between $t = 1$ and $t = 2$ only investor $x_1$ moved, her amalgamated posterior now becomes the average market belief that both investors presume as the 'true' average. Yet agglomeration has now achieved a maximum, and entropy a minimum, so no further echoing occurs (absent further signals). Interestingly, two echoing periods were observed.[173] Would such echoing have happened if instead visibility among beliefs were clearer?

As it turns out, if we leave the initial distribution of investors' own beliefs untouched from before (i.e., $x_1$ and $x_2$ hold beliefs amounting to $10\%$ and $90\%$, respectively), but now allow perfect visibility on the average belief ($50\%$), then reworking the first example finds no echoing (i.e., no updating beyond the initial anticipatory updating), and the distribution in the doxastic space settles with $x_1$ and $x_2$ having amalgamated posteriors (after quantization) of $50\%$ and $70\%$, respectively.

What might a stronger signal do? Keeping the same distribution of beliefs as in the original example, but increasing the signal's strength by an order of magnitude (to $g_t = 20$) fails to eliminate echoing. At $t = 0$, $x_1$ and $x_2$ have respective (quantized) amalgamated posteriors of

[173] Realize that the model here therefore characteristically exhibits the three hallmarks of "disagreement models" identified by Hong and Stein (2007): slow information revelation; limited attention; and diversity among prior beliefs.

70% and 90%. But the move by $x_1$ suffices to cause $x_2$ to move afterward, such that both have quantized amalgamated posterior values of 70% at $t = 1$ (and with no further echoing thereafter).

As it turns out, for the particular initial distribution in the original example, it takes a signal value of at least $g_t = 35$ (for positive signals) to eradicate the echoing effects from updating.[174] This result therefore replicates the finding by Smith and Sørenson (2000) that arbitrarily strong signals preclude false herding (which echoing might be understood as being) even in finite action spaces.

More interesting (and complex) dynamics can occur with still greater diversity among agents within the doxastic space, which stems from allowing a wider variety across the agent types (i.e., beliefs held). Realize that, for the single-dimensional, five-celled doxastic space featuring in the previous examples, unless there are at any point in time five distinct types of agents (i.e., all of the five cells within the doxastic space are initially occupied), and that the proportions of each of these five types are equivalent, then all agents will put at least some positive weight on the beliefs which they believe to be the market average. Starting with a uniform distribution of investors across the five cells in the doxastic space minimizes initial agglomeration (at zero) and, interestingly, requires a signal strength of at least $g_t = 25$ for positive signal values (and at most $g_t = -26$ for negative signal values) to instigate echoing with the starting beliefs distributed according to the table below (this signal threshold holds regardless of the specific level of initial visibility among investors' beliefs; notice, however, that the table below presumes relatively poor visibility, with each investor believing that the average belief among all of the investors matches that of her own belief exactly).

[174] But note that, for the particular initial distribution here, a value of $g_t = 1$ causes no echoing (because agents do not change their initial beliefs at all for that small of signal). Likewise signal values of $-1$ and 0 also do not echo. For negative signal values of greater (absolute) strength than $-1$, it requires $g_t = -36$ (or less) to eliminate any echoing.

*Table 5.3: Initial Evidence and Belief Distribution (Pre-Signal) for 5-Agent Example*

| | Own Belief | | | Belief about Average Belief | | |
|---|---|---|---|---|---|---|
| **Type** | $\hat{d}_A$ | $\hat{d}_B$ | Quantized | $\hat{d}_A$ | $\hat{d}_B$ | Quantized |
| $x_1$ | 1 | 9 | 0.1 | 1 | 9 | 0.1 |
| $x_2$ | 3 | 7 | 0.3 | 3 | 7 | 0.3 |
| $x_3$ | 5 | 5 | 0.5 | 5 | 5 | 0.5 |
| $x_4$ | 7 | 3 | 0.7 | 7 | 3 | 0.7 |
| $x_5$ | 9 | 1 | 0.9 | 9 | 1 | 0.9 |

Yet that same uniform distribution which maximizes initial entropy also will not echo for positive signal values of $g_t = 35$ or greater (and likewise will not echo for $g_t = -36$ or less for negative signal values). It seems, therefore that, under poor visibility, more-diverse initial distributions of beliefs may echo over a narrower range of signal values than do less-diverse (in terms of agents' beliefs) initial distributions for the 'toy' example of a single-dimensional doxastic space (save for a completely-agglomerated initial distribution where all investors use homogenous prior beliefs).

Indeed, analysis of a selection of other single-dimensional doxastic spaces (in Appendix 5.A) supports that inference (this same result will be replicated in the richer empirical settings of the subsequent sections), along with some other intriguing results. For example, whereas the maximally-diverse belief distribution (i.e., uniform over the space concerned) for a five-position doxastic space in one dimension exhibits echoing over only a relatively narrow band of signal strengths, a single-dimensional space of only three positions allows no echoing for a uniform distribution over the entire range (positive and negative) of signal sizes (likewise with $\delta = 0.5$; also, the above results, as well as the remainder of the results described in this paper, presume that investors are operating with maximal evidence: for example, $l(\theta, \delta) = 6$ in the three-location example just mentioned). Conversely, a seven-position doxastic space with one dimension and a maximally-diverse prior distribution of agents' beliefs permits echoing over a broader range of

signal values than does the five-position space. Yet, such an enlarged space then exhibits gaps between the ranges that will, and the ranges that will not, induce echoes. For example, over positive signal values, a uniform prior distribution of agent's beliefs within a seven-position, single-dimensional doxastic space exhibits echoing over only the (inclusive) ranges: $5-6$; $9-11$; $17-20$; $25-31$; and $63-76$. Positive signal values not within those ranges induce no echoing at all for the initial configuration of agents in the space as described (again, with investors presumed to be operating at their evidential limits of $l(\theta,\delta)=14$ for the seven-position doxastic space here).

What is more, there is an inverse correspondence between the numbers of periods of echoing over those specified ranges and the signal strengths: smaller (absolute) signals echo for more periods.[175] Yet, moving to an even larger single-dimensional space of $\theta=10$ (with $\delta=0.5$, and presuming investors operate at their evidential limits of $l(\theta,\delta)=20$), we find that this general inverse correspondence is only loosely monotonic: increasing signal strength only generally reduces the durations for belief echoing. Further, this relationship is not strictly monotonic (see Appendix 5.A).

Additionally, the example with $\theta=10$ and maximal initial diversity among agents' beliefs suggests that belief echoing may be more common in larger doxastic spaces than in smaller ones. In the single-dimensional example with $\theta=10$, more than $60\%$ of the signal strengths that could possibly have instigated echoing did so (in the sense that beyond a specific signal size – which is $g_t=170$ for $\theta=10$ – all agents are perfectly agglomerated after anticipatory updating takes place, and thus no further updating could occur thereafter; these results appear in Appendix 5.A). This finding suggests that, within large doxastic spaces, echoing might be more common than not.

[175] Here it is assumed that the signals are received without noise or ambiguity. Signal sizes, might, however, also represent noisy signals, wherein the 'size' of the signal corresponds with the noise-free component (ergo small signals may be those whose noisy components dominate their noiseless components). Compare these ideas with Black (1986).

Moreover, in such larger doxastic spaces, very weak signals might continue to echo over many periods, with only few investors changing their beliefs during the early periods, but a significant proportion of investors suddenly shifting their beliefs in later periods of echoing as agglomeration takes root. (Goeree et al. (2007) also find that relatively weak signals promote unstable dynamics.) These self-amplifying echoes from weak initial signals might carry important implications for explaining some curious financial market phenomena, as will be discussed within the next sections.

These results also matter in the respect that belief echoing appears to exhibit richer, more complicated dynamics as agents become *more* cognitively-uninhibited in terms of their discrimination capabilities on evidence $\theta$. Recall that the sizes of the dimensions (in terms of number of potentially-occupiable positions in each) in the doxastic spaces proposed here are governed by the cognitive parameter $\theta$. The basic examples here demonstrate that richer echoing phenomena may be more prevalent among not the most extremely-bounded agents, but among those agents with higher, albeit less-than-perfect, cognitive capacities (at least in terms of the parameter $\theta$; the penultimate section shows that similar findings apply for $F$). Thus, more interesting phenomena can often be described by the EBSL model not by 'dumbing down' agents, but instead by smartening them up (a common criticism of bounded-rationality models in financial economics is that they make bounded agents too 'un-clever'; EBSL therefore can squarely remedy that concern). As well, given the spatial expression of the cognitive limitations suggested here, it seems many of Krugman's (1994) observations on the intrigue of complex 'landscapes' for economists may apply (which is gratifying, given the model's inspiration by economic geography).

And while the foregoing basic examples with low dimensionality, and relatively modest diversity, presented in this section deserve a more complete characterization in the future, these simple situations do not appear to be the most interesting ones that the model can tackle. Instead, the

EBSL model seems more useful in its ability to cope with higher dimensionality and greater diversity (which include many realistic situations important to economists). I therefore elect to devote the remainder of this paper to exploring a more complicated example using empirical data.

## **5.5**\\ Fiscal Cliff Data

The so-called 'Fiscal Cliff' episode which unfolded in the United States over the final weeks of 2012 seems to offer an ideal test-bed scenario for the ideas on agglomeration, entropy, and quantization presented in preceding sections. The Cliff situation appeals because it presented investors with a double-layer of uncertainty (of the type described by Pástor and Veronesi (2012)): political uncertainty *and* impact uncertainty. The former concerns lack of certain knowledge about what political outcome will be realized. But resolution of political uncertainty does not necessarily entail immediate resolution of impact uncertainty, which itself concerns the macroeconomic aftermath of the specific political outcome that obtains. That is, consequences of a policy might remain uncertain despite the removal of uncertainty on which specific policy path will be chosen.

Given the relative novelty of the Cliff situation, and the presence of that double-layer of uncertainty, it seems plausible that substantial diversity among investors' beliefs about the resolution of uncertainty may have existed.[176] To test this supposition, beliefs about the macroeconomic impacts of prospective action by the US Congress on the Fiscal Cliff situation were elicited from individuals employed by institutional asset owners, investment managers, and financial consultancies between 17 and 24 December 2012. These respondents divulged their

[176] Empirical evidence (both formal and anecdotal) on substantial heterogeneity in forecasts of market and macroeconomic variables for less-novel circumstances appears (to mention but a very small subset of the relevant literature): Muth (1961, p. 316); Townsend (1983); Kurz, Jin, and Motolese (2005); Piazessi and Schneider (2010); Nagel (2012); Malmendier and Nagel (2011); Wieland and Wolters (2011). Mankiew, Ries, and Wolfers (2003) even go so far as to suggest that such "disagreement may be key to macroeconomic dynamics" (p. 2). Kandel and Pearson note that one reason for dropping assumptions of homogenous beliefs is "that it is obviously not literally true" (1995, p. 834); they suggest that it may perhaps be better to embrace belief diversity more fully.

expectations about the impact of possible Congressional action (what action specifically was purposely left unspecified to capture political uncertainty) over both the short and long terms for five aspects of the economy: consumer confidence; economic growth; economic stability; employment; and international trade (for brevity, these aspects are referred to (and abbreviated) hereafter as, respectively: *confidence* (C); *growth* (G); *stability* (S); *employment* (E); and *trade* (T)). Respondents also reported expectations about what "most investors (i.e., the market)" believed the macroeconomic impacts of prospective Congressional action would be over the same periods of time (i.e., short and long terms). Four belief sets were thus obtained for each respondent.

Respondents were all digital subscribers to *Pensions & Investments*, a major professional periodical. In total, 813 respondents completed the survey instrument, with 680 respondents' answers suitable for analytical purposes (see Appendix 5.B on detailed data-preparation protocols).

Given the historically skewed distribution of market reactions to policy changes (see Pástor and Veronesi (2012)), and the general intuition that reactions to upside shocks are less destabilizing than their downside counterparts, respondents' beliefs on potential negative economic impacts were gathered. This negative framing was accomplished through asking respondents to report their degree of agreement with a negatively-framed proposition phrased in the following which follows:

*"*[I / most investors] *believe that the outcome of Congressional action on the U.S. fiscal cliff will have a negative impact on the following aspects of the U.S. economy in the* [short / long] *term".*

The "aspects" referred to above are those five economic aspects mentioned earlier (confidence, growth, stability, employment, and international trade). And there were four versions of the proposition (see tables below for specific wordings, as well as response distributions). The propositional nature of the response prompts purposely aligns with ideas that people attach

probabilities to descriptions of events rather than events themselves. The hypothesis-based substrate for beliefs also seems adequate for the Fiscal Cliff situation, given the scenario-based presentations of the matter that were commonly put into the public domain by the media (note, for example, the popular two-scenario projections publicized by the US Congressional Budget Office).

Furthermore, respondents were asked to express their support for the truth of the propositions along a five-point Likert scale, consisting of the following response options: "strongly agree", "agree", "neutral", "disagree", and "strongly disagree". A further (sixth) option of "don't know" was also included to allow respondents to express the distinction between ambiguity of belief and neutrality of belief, although "don't know" responses were excluded from the present analysis.[177] The five response options were assumed to frame an equal partition of the unit interval, with each option covering a 'length' of 0.2 (i.e., respondents were presumed to be operating *as if* having $\theta = 5$).[178]

Despite these several shortcomings of the data, there are three distinguishing facets of it which make this empirical set of beliefs unique and suggest that it should not be dismissed out-of-hand.[179]

First, to have nearly 700 direct beliefs from professionals who are active participants inside the institutional investment community constitutes a rarity among many of the prominent studies within financial economics that use survey approaches for empirical data acquisition. While the scale of response, and identities, of the respondents (i.e., actual players in the institutional investment community, and not merely members of the general public) help make this particular dataset precious, the timing of its collection also causes it to be a rarity in meaningful ways. The

[177] A related problem was much earlier raised by Ellsberg (1961).
[178] No attempt was made to verify the accuracy of this assumption, and therefore the results which rely on this data should be taken as indicative rather than definitive. Nor were respondents provided monetary incentives (although see Tversky and Kahneman (1992) regarding non-necessity of incentives provision).
[179] For a recent study about potential research value from using survey-derived beliefs see Greenwood and Shleifer (2014). On survey-derived beliefs see also: Nyarko and Schotter (2002); Manski (2004); Barberis *et al.* (2014).

survey inquired *ex ante* about a specific event which entailed two layers of uncertainty (political and impact) and with potential material implications for financial market performance. Earlier landmark surveys of market participants on specific events generally inquired about such events only *ex post facto* (see, e.g.: Shiller (1990); Shiller, Kon-Ya, and Tsutsui (1991); Campello, Graham, and Harvey (2010)). Further, surveys about specific events with economic consequences do not typically yield information on the beliefs that respondents hold about other market participants in any form which easily lends itself to extensive formal analysis. Finally, the general novelty of the Fiscal Cliff episode means that beliefs which respondents conveyed as responses would not likely have been subjected to substantial previous 'testing' (i.e., they were relatively un-trialed expectations about a comparatively unique event in economic history). The evidence which respondents could have used to construct these beliefs (whether quantitative or qualitative) was therefore unlikely to have been substantial (this assumption will be further explored shortly). In total, then, the significant novelty of this data suggests some value in retaining it for the ensuing analysis, under the caveats that the attending results should only be understood as being indicative.

The tables below should supply the reader with at least two impressions. First, meaningful heterogeneity exists in the reported beliefs. Whether these beliefs were captured *en route* to convergence to equilibrium is not discernable from this latitudinal data. But the observable diversity suggests that considerable uncertainty existed among respondents about resolution of the Fiscal Cliff issue. Second, for many respondents, non-negligible divergences were present between their own professed beliefs and beliefs they ascribed to others in the market, (see Appendix 5.C), which bolsters earlier assertions that agents might hold distinct beliefs for themselves and others.

*Table 5.4: Reported Beliefs of Outcomes on Congressional Action on U.S. Fiscal Cliff Issue*

| Self / Short Term (SST) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **I believe that the outcome of congressional action on the U.S. fiscal cliff will have a negative impact on the following aspects of the U.S. economy in the short term:** | | | | | | | | |
| | Strongly Agree | Agree | Neutral | Disagree | Strongly Disagree | Don't Know | No Response | TOTAL |
| Consumer Confidence | 156 | 276 | 94 | 96 | 34 | 24 | 0 | 680 |
| | *23%* | *41%* | *14%* | *14%* | *5%* | *4%* | *0%* | *100%* |
| Economic Growth | 181 | 258 | 87 | 106 | 21 | 23 | 4 | 680 |
| | *27%* | *38%* | *13%* | *16%* | *3%* | *3%* | *1%* | *100%* |
| Economic Stability | 174 | 227 | 100 | 120 | 33 | 24 | 2 | 680 |
| | *26%* | *33%* | *15%* | *18%* | *5%* | *4%* | *0%* | *100%* |
| Employment | 160 | 254 | 121 | 101 | 20 | 22 | 2 | 680 |
| | *24%* | *37%* | *18%* | *15%* | *3%* | *3%* | *0%* | *100%* |
| International Trade | 74 | 159 | 243 | 129 | 37 | 37 | 1 | 680 |
| | *11%* | *23%* | *36%* | *19%* | *5%* | *5%* | *0%* | *100%* |

| Self / Long Term (SLT) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **I believe that the outcome of congressional action on the U.S. fiscal cliff will have a negative impact on the following aspects of the U.S. economy in the long term:** | | | | | | | | |
| | Strongly Agree | Agree | Neutral | Disagree | Strongly Disagree | Don't Know | No Response | TOTAL |
| Consumer Confidence | 107 | 201 | 139 | 153 | 55 | 23 | 2 | 680 |
| | *16%* | *30%* | *20%* | *23%* | *8%* | *3%* | *0%* | *100%* |
| Economic Growth | 154 | 210 | 92 | 142 | 59 | 22 | 1 | 680 |
| | *23%* | *31%* | *14%* | *21%* | *9%* | *3%* | *0%* | *100%* |
| Economic Stability | 141 | 208 | 89 | 149 | 67 | 22 | 4 | 680 |
| | *21%* | *31%* | *13%* | *22%* | *10%* | *3%* | *1%* | *100%* |
| Employment | 144 | 214 | 110 | 137 | 51 | 21 | 3 | 680 |
| | *21%* | *31%* | *16%* | *20%* | *8%* | *3%* | *0%* | *100%* |
| International Trade | 79 | 138 | 215 | 142 | 58 | 38 | 10 | 680 |
| | *12%* | *20%* | *32%* | *21%* | *9%* | *6%* | *1%* | *100%* |

| Market / Short Term (MST) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Most investors (i.e., the market) believe that the outcome of congressional action on the U.S. fiscal cliff will have a negative impact on the following aspects of the U.S. economy in the short term:** | | | | | | | | |
| | Strongly Agree | Agree | Neutral | Disagree | Strongly Disagree | Don't Know | No Response | TOTAL |
| Consumer Confidence | 141 | 293 | 91 | 108 | 16 | 30 | 1 | 680 |
| | *21%* | *43%* | *13%* | *16%* | *2%* | *4%* | *0%* | *100%* |
| Economic Growth | 158 | 299 | 79 | 101 | 12 | 29 | 2 | 680 |
| | *23%* | *44%* | *12%* | *15%* | *2%* | *4%* | *0%* | *100%* |
| Economic Stability | 155 | 261 | 106 | 111 | 14 | 29 | 4 | 680 |
| | *23%* | *38%* | *16%* | *16%* | *2%* | *4%* | *1%* | *100%* |
| Employment | 130 | 307 | 101 | 99 | 9 | 29 | 5 | 680 |
| | *19%* | *45%* | *15%* | *15%* | *1%* | *4%* | *1%* | *100%* |
| International Trade | 80 | 204 | 220 | 103 | 18 | 48 | 7 | 680 |
| | *12%* | *30%* | *32%* | *15%* | *3%* | *7%* | *1%* | *100%* |

| Market / Long Term (MLT) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Most investors (i.e., the market) believe that the outcome of congressional action on the U.S. fiscal cliff will have a negative impact on the following aspects of the U.S. economy in the long term:** | | | | | | | | |
| | Strongly Agree | Agree | Neutral | Disagree | Strongly Disagree | Don't Know | No Response | TOTAL |
| Consumer Confidence | 98 | 224 | 153 | 134 | 35 | 35 | 1 | 680 |
| | *14%* | *33%* | *23%* | *20%* | *5%* | *5%* | *0%* | *100%* |
| Economic Growth | 119 | 237 | 122 | 131 | 36 | 34 | 1 | 680 |
| | *18%* | *35%* | *18%* | *19%* | *5%* | *5%* | *0%* | *100%* |
| Economic Stability | 114 | 223 | 135 | 131 | 39 | 34 | 4 | 680 |
| | *17%* | *33%* | *20%* | *19%* | *6%* | *5%* | *1%* | *100%* |
| Employment | 98 | 263 | 124 | 123 | 32 | 35 | 5 | 680 |
| | *14%* | *39%* | *18%* | *18%* | *5%* | *5%* | *1%* | *100%* |
| International Trade | 64 | 178 | 218 | 125 | 38 | 49 | 8 | 680 |
| | *9%* | *26%* | *32%* | *18%* | *6%* | *7%* | *1%* | *100%* |

From these initial distributions of beliefs, starting values for entropy, syntropy, agglomeration, and juncture can all be calculated. Because the belief sets used are samples, however, one must be concerned about whether or not the observed values are statistically different from random values. To be able to conduct such tests of statistical difference one must generally obey a 'coverage rule' of five respondents per cell within a given doxastic space. With five response levels per dimension (i.e., $\theta = 5$), the response volumes permit comfortable results for doxastic spaces of up to three dimensions (i.e., $F = 3$), such that $\theta^F = 125$. As there are five possible hypotheses that might be combined together, the values for entropy, syntropy, agglomeration, and juncture were calculated for all possible triples for both self and market responses over both the long and short terms. The results of such calculations appear in the tables below which are split between short and long terms.

*Table 5.5: Starting Entropy Statistics Calculated from Initial Belief Distributions*

| | *Maximum* | **Self - Short Term** | | | | **Market - Short Term** | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Triple* | *Possible Entropy* | *Entropy* | *Syntropy* | *Agglomeration* | *Juncture* | *Entropy* | *Syntropy* | *Agglomeration* | *Juncture* |
| C/G/S | 4.83 | 3.39 | 1.44 | 0.55 | 0.89 | 2.58 | 2.25 | 0.76 | 1.49 |
| C/G/E | 4.83 | 3.34 | 1.48 | 0.60 | 0.88 | 2.63 | 2.20 | 0.84 | 1.36 |
| C/G/T | 4.83 | 3.54 | 1.29 | 0.57 | 0.72 | 2.84 | 1.99 | 0.76 | 1.23 |
| C/S/E | 4.83 | 3.51 | 1.32 | 0.52 | 0.80 | 2.73 | 2.10 | 0.77 | 1.33 |
| C/S/T | 4.83 | 3.58 | 1.25 | 0.50 | 0.75 | 2.88 | 1.95 | 0.69 | 1.26 |
| C/E/T | 4.83 | 3.52 | 1.31 | 0.55 | 0.76 | 2.88 | 1.95 | 0.77 | 1.18 |
| G/S/E | 4.83 | 3.24 | 1.59 | 0.54 | 1.05 | 2.50 | 2.33 | 0.81 | 1.52 |
| G/S/T | 4.83 | 3.39 | 1.44 | 0.52 | 0.92 | 2.70 | 2.13 | 0.73 | 1.40 |
| G/E/T | 4.83 | 3.35 | 1.48 | 0.58 | 0.90 | 2.71 | 2.12 | 0.80 | 1.31 |
| S/E/T | 4.83 | 3.41 | 1.42 | 0.50 | 0.92 | 2.79 | 2.04 | 0.74 | 1.30 |

| | *Maximum* | **Self - Long Term** | | | | **Market - Long Term** | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Triple* | *Possible Entropy* | *Entropy* | *Syntropy* | *Agglomeration* | *Juncture* | *Entropy* | *Syntropy* | *Agglomeration* | *Juncture* |
| C/G/S | 4.83 | 3.08 | 1.74 | 0.22 | 1.52 | 2.64 | 2.18 | 0.38 | 1.81 |
| C/G/E | 4.83 | 3.14 | 1.69 | 0.24 | 1.45 | 2.66 | 2.17 | 0.43 | 1.74 |
| C/G/T | 4.83 | 3.30 | 1.53 | 0.25 | 1.28 | 2.91 | 1.92 | 0.42 | 1.50 |
| C/S/E | 4.83 | 3.15 | 1.68 | 0.22 | 1.45 | 2.76 | 2.07 | 0.42 | 1.65 |
| C/S/T | 4.83 | 3.32 | 1.51 | 0.24 | 1.27 | 2.95 | 1.88 | 0.42 | 1.46 |
| C/E/T | 4.83 | 3.29 | 1.54 | 0.25 | 1.29 | 2.85 | 1.98 | 0.47 | 1.51 |
| G/S/E | 4.83 | 2.89 | 1.94 | 0.24 | 1.70 | 2.49 | 2.34 | 0.42 | 1.92 |
| G/S/T | 4.83 | 3.11 | 1.72 | 0.25 | 1.47 | 2.74 | 2.09 | 0.42 | 1.68 |
| G/E/T | 4.83 | 3.12 | 1.70 | 0.26 | 1.44 | 2.73 | 2.10 | 0.47 | 1.64 |
| S/E/T | 4.83 | 3.11 | 1.72 | 0.26 | 1.47 | 2.79 | 2.03 | 0.46 | 1.57 |

Note: C = “Consumer Confidence”; G = “Economic Growth”; S = “Economic Stability”; E = “Employment”; T = “International Trade”.

As is visible from the above tables, measured syntropy is much lower than measured entropy for all possible triples, which indicates substantial belief diversity and can be understood as considerable interpersonal uncertainty at the collective level. Additionally, syntropy is uniformly higher for ‘market’ beliefs than for what agents profess as their own beliefs: respondents agree more upon what they think about a representative market agent believes than what they do themselves believe. It is difficult to say whether or not this sample is reflective of the population of beliefs at the time of collection (i.e., if it is indeed representative of the true overall market belief distribution). Yet if such were the case, then it might be said that meaningful visibility issues existed, and investors were mistaken in their estimates about the location of average market belief.

The relationship between syntropy over the long and short terms is less clear. What is interesting to recognize, however, is that in all instances, juncture accounts for a higher proportion of the organization among respondents’ beliefs than agglomeration does. A first suspicion, then, might be that, even though little initial consensus exists among these beliefs, there are, *in the minds of the respondents*, some dependencies between the hypotheses upon which these beliefs are founded.

With these calculated values for inputs, we can now undertake various investigatory simulations.

## **5.6**\\ Analysis and Discussion

### **5.6.1**\\ Setup

The simulations are meant to demonstrate the model's performance in a richer setting than the basic setups presented within Section 5.4.4. These simulations take aboard several simplifying assumptions (which are detailed below) that are useful for keeping the analysis tidy and straightforward, and could be jettisoned for future studies, but are retained here to maintain focus.

First, as mentioned, the usable volume of belief responses (680) comfortably permits at most combinations of three hypotheses simultaneously (for statistical reasons; a psychological motivation might be given to this limitation at three dimensions, however, as humans typically have a hard time mentally-visualizing more than three dimensions at any one time). With five response options per hypothesis to indicate degree of belief, this restriction allows for doxastic spaces with sizes of $N = 5^3 = 125$ discrete positions. It is further assumed here that $\delta = 0.5$, so that this local-thinking parameter takes it neutral value (as described in 4.3.1), and causes each quantized belief to take the midpoint values for the decision interval within which it is located. Together, these assumptions imply a joint cognitive limit of $l(\theta, \delta) = \theta(1 - \delta)^{-1} = 10$. No efforts were made to test whether this limit indeed held for respondents, yet such calibration is not necessary for the indicative analysis undertaken here (it must be then borne in mind, however, that the results in this section should therefore be taken as more illustrative than necessarily veridical).

Imposing a discrimination limit of $\theta = 5$ causes the EBSL model to treat agents as if they segment the probability scale into five uniform decision intervals. A table relating the quantized probabilities pertaining to each of the resulting decision intervals, as well as matching the values of evidence assumed to give rise to such probabilities with their Likert responses, appears below.

*Table 5.6: Evidence and Quantized Probability Conversion Scale*

| Likert response | Quantized probability | Assumed $\widehat{d}_A$ | Assumed $\widehat{d}_B$ |
|---|---|---|---|
| Strongly agree | 0.9 | 9 | 1 |
| Agree | 0.7 | 7 | 3 |
| Neutral | 0.5 | 5 | 5 |
| Disagree | 0.3 | 3 | 7 |
| Strongly disagree | 0.1 | 1 | 9 |

The above values and quantized probabilities are taken to apply for both short- and long-term beliefs expressed by respondents as their own beliefs, as well as those about others within the market. That $\theta = 5$ applies to agents' beliefs on others' beliefs signifies that such agents are only able to engage in limited classifications of other agents' beliefs, such that resulting classifications are "coarse", as it is also with Mullainathan, Schwartzstein, and Shleifer (2008) and Jehiel (2005).

The simulations were set up to reflect as closely as possible the sequence of the basic model of learning dynamics contained in Section 5.4: agents observe the initial values of entropy, agglomeration, and juncture, and then update their beliefs for themselves and (the average of) all others in response to an exogenous signal; subsequent reactive updating (echoing) may occur thereafter as all agents in the simulations observe when any agent's beliefs change from occupying one decision interval (in any particular dimension of the doxastic space) to another. Thus, the true (updated) market belief distribution becomes (partially) revealed only after a signal gets received.

In the initial period of each simulation (period $t = 0$) a signal $g_0$ arrives. The signal may take integer values in the range $[-10,10] \subset \mathbb{Z}$ for each of the three hypotheses. (Again, one could introduce continuous-valued signals, but because agents discretize evidence anyway under QET, this assumption of discrete signal values avoids unnecessary hassle, given that the aim here is to just demonstrate the model's performance in an environment of high belief diversity.) Zero-valued signals for a hypothesis are treated as no signal having been received for the hypothesis (with the

relevant impact upon updating through the juncture term; see Subsection 5.4.3). All ten possible triples of hypotheses were explored for both the short and long terms (making for twenty sets of simulations in total), and all possible combinations of signal values across each triple were simulated (9261 combinations for each triple). In total, 185,200 distinct starting conditions were simulated. Simulations were run via a program written by the author in the R computing language.

### **5.6.2**\\ Results

Table 5.7 below presents the distribution of simulation durations. These durations reflect the total number of periods (from the initial period in which the exogenous signal arrives and anticipatory updating takes place onwards) in which belief changes took place among the 680 simulated agents (i.e., duration reflects the last period in which beliefs changed in the simulated doxastic spaces). It should be noted that "zero" periods of updating corresponds to simulations for which no changes in agents' beliefs took place after the signal's arrival (these zero-updating simulations all were connected to extremely weak signals with one or more dimensions of the signal having strengths equal to zero, and thus agents treating such dimensions as non-arrivals of a signal). It should also be noticed that simulations lasting for only a single period involved only anticipatory updating. That is, simulations with durations of only one period involved no reactive updating by agents in response to observed changes in others' beliefs, and exhibited no belief echoing. As can be seen, the vast majority of simulations (usually ninety percent or above) involved belief echoing, with average simulations witnessing around three periods of updating activity (i.e., two periods of echoing). A significant number of simulations saw as many as four periods of activity (three echoing periods), with a very small minority of simulations having up to five periods of updating.

*Table 5.7: Distribution of Simulation Durations and Other Selected Statistics*

| Short Term | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Total Duration** | **Hypothesis Combination (Triple)** | | | | | | | | | |
| **(Periods)** | **CGS** | **CGE** | **CGT** | **CSE** | **CST** | **CET** | **GSE** | **GST** | **GET** | **SET** |
| 0 | 7 | 27 | 27 | 27 | 27 | 27 | 15 | 27 | 27 | 27 |
| 1 | 36 | 420 | 507 | 579 | 645 | 687 | 499 | 624 | 642 | 866 |
| 2 | 1338 | 1159 | 1263 | 1577 | 1612 | 1137 | 888 | 1411 | 936 | 1018 |
| 3 | 5832 | 5496 | 5348 | 5014 | 4807 | 4941 | 5502 | 5290 | 5237 | 5009 |
| 4 | 2048 | 2159 | 2116 | 2064 | 2170 | 2469 | 2327 | 1909 | 2419 | 2341 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 30 | 0 | 0 | 0 |
| Total | 9261 | 9261 | 9261 | 9261 | 9261 | 9261 | 9261 | 9261 | 9261 | 9261 |
| Average Duration | 3.07 | 3.01 | 2.97 | 2.92 | 2.91 | 2.99 | 3.05 | 2.91 | 3.01 | 2.95 |
| % with Echoing | *99.5%* | *95.2%* | *94.2%* | *93.5%* | *92.7%* | *92.3%* | *94.4%* | *93.0%* | *92.8%* | *90.4%* |

| Long Term | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Total Duration** | **Hypothesis Combination (Triple)** | | | | | | | | | |
| **(Periods)** | **123** | **124** | **125** | **134** | **135** | **145** | **234** | **235** | **245** | **345** |
| 0 | 27 | 27 | 27 | 27 | 27 | 27 | 27 | 27 | 27 | 27 |
| 1 | 795 | 1005 | 1005 | 1120 | 1099 | 1281 | 1066 | 1058 | 1281 | 1296 |
| 2 | 2132 | 1582 | 1778 | 1985 | 1898 | 2054 | 1322 | 1314 | 1635 | 1540 |
| 3 | 3723 | 3717 | 3487 | 3222 | 3219 | 3793 | 3808 | 3613 | 4508 | 4726 |
| 4 | 2584 | 2930 | 2964 | 2907 | 3018 | 2106 | 3038 | 3249 | 1810 | 1672 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Total | 9261 | 9261 | 9261 | 9261 | 9261 | 9261 | 9261 | 9261 | 9261 | 9261 |
| Average Duration | 2.87 | 2.92 | 2.90 | 2.85 | 2.87 | 2.72 | 2.95 | 2.97 | 2.73 | 2.73 |
| % with Echoing | *91.1%* | *88.9%* | *88.9%* | *87.6%* | *87.8%* | *85.9%* | *88.2%* | *88.3%* | *85.9%* | *85.7%* |

*Table 5.8: Average Number of Occupied Positions in Doxastic Space after each Period*

| Short Term | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Hypothesis Combination (Triple)** | | | | | | | | | |
| **Time Period** | **CGS** | **CGE** | **CGT** | **CSE** | **CST** | **CET** | **GSE** | **GST** | **GET** | **SET** |
| t = -1 | 72.00 | 68.00 | 72.00 | 78.00 | 74.00 | 69.00 | 68.00 | 67.00 | 64.00 | 67.00 |
| 0 | 32.79 | 31.51 | 33.08 | 34.15 | 33.78 | 31.31 | 30.73 | 30.48 | 30.16 | 31.19 |
| 1 | 8.38 | 12.23 | 12.81 | 13.32 | 13.55 | 14.82 | 12.22 | 12.80 | 14.72 | 16.03 |
| 2 | 5.07 | 8.55 | 9.14 | 10.00 | 10.20 | 9.67 | 8.18 | 8.97 | 9.47 | 10.53 |
| 3 | 4.58 | 7.95 | 8.61 | 9.42 | 9.57 | 9.09 | 7.61 | 8.57 | 8.88 | 9.94 |
| 4 | 4.58 | 7.95 | 8.61 | 9.42 | 9.57 | 9.09 | 7.60 | 8.57 | 8.88 | 9.94 |

| Long Term | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Hypothesis Combination (Triple)** | | | | | | | | | |
| **Time Period** | **CGS** | **CGE** | **CGT** | **CSE** | **CST** | **CET** | **GSE** | **GST** | **GET** | **SET** |
| t = -1 | 62.00 | 63.00 | 65.00 | 61.00 | 65.00 | 66.00 | 57.00 | 63.00 | 62.00 | 57.00 |
| 0 | 29.09 | 30.24 | 31.54 | 28.86 | 31.10 | 39.57 | 27.91 | 30.53 | 40.31 | 37.00 |
| 1 | 13.01 | 15.85 | 15.74 | 15.43 | 15.60 | 22.62 | 15.35 | 15.99 | 23.84 | 22.16 |
| 2 | 10.37 | 11.60 | 11.81 | 11.79 | 11.99 | 17.18 | 10.97 | 11.45 | 17.60 | 15.97 |
| 3 | 9.77 | 10.82 | 11.01 | 11.03 | 11.19 | 16.49 | 10.33 | 10.71 | 16.96 | 15.46 |
| 4 | 9.77 | 10.82 | 11.01 | 11.03 | 11.19 | 16.49 | 10.33 | 10.71 | 16.96 | 15.46 |

Table 5.8 demonstrates that this significant degree of observed belief echoing from the simulations may be driven substantially by the increasing levels of agglomeration as the simulations progress

after the exogenous signal's arrival. Because agglomeration values may not (yet) be a familiar statistic to many economists, values in Table 5.8 are instead presented in terms of average number of occupied positions in the simulated doxastic spaces at the end of each period (each of size $N = 125$ positions, spread across $F = 3$ dimensions; note that these occupancy figures pertain to agents' own beliefs, i.e., their personal priors and amalgamated posteriors at the end of each period). Table 5.8 indicates sizable levels of initial diversity present at the start of each simulation, due to more than half of the possible positions in the doxastic space being occupied from the onset (which indicates considerable diversity in agents' priors before the signal's arrival). But it also shows that, despite the non-decreasing nature of doxastic agglomeration, some residual diversity remains at the end of most simulations. Specifically, while many simulations conclude with agents' beliefs highly-agglomerated within the doxastic space once all updating finishes (both anticipatory updating and echoing), most simulations do not conclude with total doxastic agglomeration and perfect heterogeneity among agents' beliefs. As will be discussed shortly, this high-but-imperfect doxastic agglomeration could result in substantial subsequent instability among the agents' beliefs.

Results from the simulations also reflect an interesting (and anticipated) relationship between signal size and duration (in periods) of updating among agents' beliefs. Table 5.9 shows the distribution of all simulation durations for given signal strengths (where total signal strength is the sum of the absolute values of each signal for the three hypothesis-dimensions involved in each simulation run). As can be seen, average duration is not strictly increasing with total signal strength: average duration increases at first with total signal strength (the intuition here is that initial updating activity and echoing need 'activation energy', as in chemical reactions), but beyond a total signal strength of around 15, average duration begins to slowly decline. While, above a total signal strength of 21, all simulations exhibit some echoing (at least for the signal strengths

simulated here of up to |10| for any one hypothesis-dimension), the proportion of simulations with longer durations (e.g., four periods) declines. A straightforward interpretation of this finding is that (in alignment with Smith and Sørenson (2000)) sufficiently-large signal strengths tend to diminish echoing (and large enough signals might even eliminate it) by removing a substantial amount of uncertainty during the anticipatory updating period. Weaker signals, meanwhile, do not remove enough compositional uncertainty, and may even strengthen it as agents suddenly find themselves discovering unanticipated belief-updating by their peers as reaction to the public signal.

Yet, Table 5.9 shows distributions for signals that are both mixed (i.e., signals whose values for some hypotheses are positive while others are negative) and pure (that is, signals that are either exclusively positive or exclusively negative for all of those hypotheses to which such signals pertain). Table 5.10 therefore shows duration distributions for only purely positive and negative signals (separately). While for weaker strengths purely positive and negative signals are roughly symmetric (i.e., they have similar distributions across durations and comparable average durations), stronger signals seem to reflect an asymmetry related to the asymmetry of the initial belief distribution among agents. Recall that the *ex ante* distributions of both agents' own beliefs and their beliefs about average beliefs in the market were pessimistic: they were slightly skewed toward agreement with the proposition that Congressional action would have a negative effect on the macroeconomic variables in question. It would seem, in light of results in Table 5.10, that purely positive signals, which would provide confirmatory evidence on this slightly negative skew in beliefs, therefore reduce echoing. Yet slightly disconfirming evidence from negative signals seems to increase echoing. For the Fiscal Cliff example here, this echoing from disconfirming evidence (i.e., news that Congressional action might have a non-detrimental effect on the macroeconomic variables), may have had a sustained and self-reinforcing impact on reversing

pessimism (although the model below is certainly too stylized to generate any firm conclusions; while the simulations here may supply indicative suggestions and intuition, further details would be required to be able to confidently derive assertions about any roles which echoing or entropy-based social learning may have played in reactions by investors to a 'resolution of the Fiscal Cliff).

*Table 5.9: Distribution of All Simulation Durations for Each Signal Strength*

| Signal Strength | Average Duration | % of Simulations Per Duration (in Periods) 0 | 1 | 2 | 3 | 4 | 5 | % with Echoing | Number of Simulations |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.000 | *100.0%* | -- | -- | -- | -- | -- | *0.0%* | *20* |
| 1 | 0.000 | *100.0%* | -- | -- | -- | -- | -- | *0.0%* | *120* |
| 2 | 0.606 | *62.2%* | *23.3%* | *6.1%* | *8.3%* | -- | -- | *14.4%* | *360* |
| 3 | 1.255 | *18.9%* | *51.1%* | *15.5%* | *14.5%* | -- | -- | *30.0%* | *760* |
| 4 | 1.755 | -- | *54.4%* | *20.7%* | *19.9%* | *5.0%* | -- | *45.6%* | *1,320* |
| 5 | 2.092 | -- | *42.0%* | *19.2%* | *26.3%* | *12.5%* | -- | *58.0%* | *2,040* |
| 6 | 2.264 | -- | *34.6%* | *21.1%* | *27.6%* | *16.7%* | -- | *65.4%* | *2,920* |
| 7 | 2.413 | -- | *29.6%* | *21.6%* | *26.9%* | *22.0%* | -- | *70.4%* | *3,960* |
| 8 | 2.559 | -- | *24.1%* | *21.9%* | *27.9%* | *26.1%* | -- | *75.9%* | *5,160* |
| 9 | 2.671 | -- | *20.0%* | *21.3%* | *30.2%* | *28.4%* | *0.0%* | *80.0%* | *6,520* |
| 10 | 2.756 | -- | *16.9%* | *20.4%* | *33.2%* | *29.4%* | *0.1%* | *83.1%* | *8,040* |
| 11 | 2.834 | -- | *14.1%* | *19.2%* | *36.0%* | *30.6%* | *0.1%* | *85.9%* | *9,600* |
| 12 | 2.941 | -- | *10.5%* | *17.7%* | *39.1%* | *32.6%* | *0.1%* | *89.5%* | *10,960* |
| 13 | 3.028 | -- | *7.8%* | *16.2%* | *41.3%* | *34.7%* | -- | *92.2%* | *12,000* |
| 14 | 3.054 | -- | *7.6%* | *14.9%* | *41.9%* | *35.5%* | -- | *92.4%* | *12,720* |
| 15 | 3.055 | -- | *7.2%* | *14.9%* | *43.3%* | *34.7%* | -- | *92.8%* | *13,120* |
| 16 | 3.042 | -- | *6.7%* | *15.4%* | *44.9%* | *33.0%* | -- | *93.3%* | *13,200* |
| 17 | 3.036 | -- | *5.8%* | *15.6%* | *47.8%* | *30.8%* | -- | *94.2%* | *12,960* |
| 18 | 3.037 | -- | *4.8%* | *15.0%* | *52.0%* | *28.2%* | -- | *95.2%* | *12,400* |
| 19 | 3.034 | -- | *3.7%* | *14.4%* | *56.6%* | *25.3%* | -- | *96.3%* | *11,520* |
| 20 | 3.034 | -- | *2.6%* | *13.9%* | *61.0%* | *22.5%* | -- | *97.4%* | *10,320* |
| 21 | 3.046 | -- | *1.2%* | *12.9%* | *66.0%* | *19.9%* | -- | *98.8%* | *8,800* |
| 22 | 3.047 | -- | -- | *11.8%* | *71.6%* | *16.6%* | -- | *100.0%* | *7,200* |
| 23 | 2.988 | -- | -- | *12.3%* | *76.5%* | *11.1%* | -- | *100.0%* | *5,760* |
| 24 | 2.948 | -- | -- | *12.7%* | *79.8%* | *7.5%* | -- | *100.0%* | *4,480* |
| 25 | 2.911 | -- | -- | *13.3%* | *82.3%* | *4.4%* | -- | *100.0%* | *3,360* |
| 26 | 2.889 | -- | -- | *14.3%* | *82.6%* | *3.1%* | -- | *100.0%* | *2,400* |
| 27 | 2.882 | -- | -- | *14.9%* | *81.9%* | *3.1%* | -- | *100.0%* | *1,600* |
| 28 | 2.875 | -- | -- | *15.6%* | *81.3%* | *3.1%* | -- | *100.0%* | *960* |
| 29 | 2.875 | -- | -- | *15.6%* | *81.3%* | *3.1%* | -- | *100.0%* | *480* |
| 30 | 2.875 | -- | -- | *15.6%* | *81.3%* | *3.1%* | -- | *100.0%* | *160* |

Note: Signal strengths are sums of absolute values of signal sizes for all three hypotheses in each simulation run.

*Table 5.10: Distribution of Simulation Durations for Purely Positive and Negative Signals*

| Signal Size | | Average Duration | | % of Simulations per Duration (in Periods) | | | | | | | | | | | | Number of Simulations | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | 0 | | 1 | | 2 | | 3 | | 4 | | 5 | | | |
| + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - |
| 3 | -3 | 0.150 | 0.150 | 90.0% | 90.0% | 5.0% | 5.0% | 5.0% | 5.0% | -- | -- | -- | -- | -- | -- | 20 | 20 |
| 4 | -4 | 1.567 | 1.467 | -- | -- | 58.3% | 61.7% | 26.7% | 30.0% | 15.0% | 8.3% | -- | -- | -- | -- | 60 | 60 |
| 5 | -5 | 2.367 | 2.058 | -- | -- | 34.2% | 44.2% | 15.0% | 25.0% | 30.8% | 11.7% | 20.0% | 19.2% | -- | -- | 120 | 120 |
| 6 | -6 | 2.505 | 2.375 | -- | -- | 27.5% | 28.0% | 17.5% | 30.5% | 32.0% | 17.5% | 23.0% | 24.0% | -- | -- | 200 | 200 |
| 7 | -7 | 2.673 | 2.507 | -- | -- | 24.0% | 25.0% | 16.0% | 30.0% | 28.7% | 14.3% | 31.3% | 30.7% | -- | -- | 300 | 300 |
| 8 | -8 | 2.812 | 2.605 | -- | -- | 18.3% | 22.4% | 15.7% | 28.6% | 32.4% | 15.2% | 33.6% | 33.8% | -- | -- | 420 | 420 |
| 9 | -9 | 2.896 | 2.746 | -- | -- | 12.5% | 17.9% | 15.2% | 26.1% | 42.5% | 19.6% | 29.8% | 36.4% | -- | -- | 560 | 560 |
| 10 | -10 | 2.972 | 2.839 | -- | -- | 8.8% | 14.3% | 13.9% | 25.1% | 48.8% | 23.2% | 28.6% | 37.1% | -- | 0.3% | 720 | 720 |
| 11 | -11 | 2.977 | 2.942 | -- | -- | 8.2% | 11.8% | 14.1% | 22.9% | 49.4% | 24.9% | 28.2% | 40.2% | -- | 0.2% | 900 | 900 |
| 12 | -12 | 3.002 | 3.083 | -- | -- | 7.9% | 8.5% | 13.0% | 18.6% | 50.1% | 29.0% | 29.0% | 43.6% | -- | 0.2% | 1100 | 1100 |
| 13 | -13 | 3.070 | 3.229 | -- | -- | 5.0% | 4.8% | 11.9% | 15.1% | 54.2% | 32.5% | 28.9% | 47.6% | -- | -- | 1260 | 1260 |
| 14 | -14 | 3.062 | 3.280 | -- | -- | 5.4% | 4.3% | 10.0% | 12.4% | 57.5% | 34.4% | 27.0% | 48.9% | -- | -- | 1380 | 1380 |
| 15 | -15 | 3.051 | 3.257 | -- | -- | 5.5% | 3.7% | 9.0% | 14.4% | 60.3% | 34.5% | 25.1% | 47.5% | -- | -- | 1460 | 1460 |
| 16 | -16 | 3.034 | 3.216 | -- | -- | 5.8% | 3.3% | 8.5% | 16.7% | 62.3% | 35.1% | 23.5% | 44.9% | -- | -- | 1500 | 1500 |
| 17 | -17 | 3.050 | 3.157 | -- | -- | 5.1% | 3.4% | 8.1% | 18.4% | 63.5% | 37.3% | 23.3% | 40.9% | -- | -- | 1500 | 1500 |
| 18 | -18 | 3.063 | 3.144 | -- | -- | 4.6% | 2.7% | 8.3% | 16.8% | 63.4% | 44.0% | 23.8% | 36.5% | -- | -- | 1460 | 1460 |
| 19 | -19 | 3.063 | 3.126 | -- | -- | 4.1% | 2.0% | 8.9% | 15.0% | 63.5% | 51.5% | 23.5% | 31.5% | -- | -- | 1380 | 1380 |
| 20 | -20 | 3.056 | 3.109 | -- | -- | 3.0% | 1.4% | 10.6% | 13.7% | 64.0% | 57.4% | 22.3% | 27.5% | -- | -- | 1260 | 1260 |
| 21 | -21 | 3.043 | 3.112 | -- | -- | 1.7% | 0.8% | 11.5% | 11.4% | 67.5% | 63.6% | 19.3% | 24.2% | -- | -- | 1100 | 1100 |
| 22 | -22 | 3.037 | 3.136 | -- | -- | -- | -- | 11.8% | 7.9% | 72.8% | 70.7% | 15.4% | 21.4% | -- | -- | 900 | 900 |
| 23 | -23 | 2.971 | 3.104 | -- | -- | -- | -- | 11.7% | 6.3% | 79.6% | 77.1% | 8.8% | 16.7% | -- | -- | 720 | 720 |
| 24 | -24 | 2.920 | 3.127 | -- | -- | -- | -- | 11.8% | 3.4% | 84.5% | 80.5% | 3.8% | 16.1% | -- | -- | 560 | 560 |
| 25 | -25 | 2.879 | 3.145 | -- | -- | -- | -- | 12.1% | 1.0% | 87.9% | 83.6% | -- | 15.5% | -- | -- | 420 | 420 |
| 26 | -26 | 2.870 | 3.150 | -- | -- | -- | -- | 13.0% | -- | 87.0% | 85.0% | -- | 15.0% | -- | -- | 300 | 300 |
| 27 | -27 | 2.860 | 3.150 | -- | -- | -- | -- | 14.0% | -- | 86.0% | 85.0% | -- | 15.0% | -- | -- | 200 | 200 |
| 28 | -28 | 2.850 | 3.150 | -- | -- | -- | -- | 15.0% | -- | 85.0% | 85.0% | -- | 15.0% | -- | -- | 120 | 120 |
| 29 | -29 | 2.850 | 3.150 | -- | -- | -- | -- | 15.0% | -- | 85.0% | 85.0% | -- | 15.0% | -- | -- | 60 | 60 |
| 30 | -30 | 2.850 | 3.150 | -- | -- | -- | -- | 15.0% | -- | 85.0% | 85.0% | -- | 15.0% | -- | -- | 20 | 20 |

How much impact might belief echoing actually have for those situations where it occurs? While the EBSL model simulated here cannot answer this question in terms of price impacts, it is nevertheless capable of providing some potential insights into the relevant volumes of belief changes among agents. Table 5.11 shows the proportions of agents in the simulations that changed their positions in the doxastic space over each period. While it is observable that the average percentage of agents changing their positions due to changes in their (quantized) beliefs generally decreases with increasing duration, we can see that the associated maxima drop off much less rapidly from one period to the next. That is, although reactive updating (i.e., belief echoing in periods beyond $t = 0$) tends to find a smaller proportion of agents changing their beliefs in each period $t \geq 1$ than for the anticipatory updating period $t = 0$, there exist periods of echoing wherein the vast majority of agents change their beliefs and move within the doxastic space for periods $t > 0$. As will be seen shortly, these situations wherein large proportions of agents change locations in

the doxastic space several periods after the signal's arrival may exhibit little activity during their early periods, and therefore the shifts which results may come as unheralded or otherwise 'abrupt'. Lee (1998) suggests that these sudden and mass shifts in beliefs may eventuate in market crashes.

The results from Table 5.11 can also be viewed in another way, as demonstrated by Table 5.12. Table 5.12 shows the average total proportion of movements for each hypothesis triple (across all periods of each simulation) that are attributable to exogenous and endogenous 'evidence', respectively. That is, exogenous attributions are those movements by agents within the doxastic space that take place during period $t = 0$ and correspond to anticipatory updating solely in response to the arrival of the exogenous signal. Meanwhile, endogenous attributions are those movements by agents due to reactive updating during periods $t \geq 1$, and are thus due only to agents responding (endogenously) to visible changes in one another's beliefs, as well as the attending changes in agglomeration, entropy, and juncture values within the overall collection of agents. As can be seen, exogenous and endogenous movements measured in such ways are similar.

Nonetheless, we might be suspicious that, in more realistic settings, changes in beliefs when beliefs are highly-diverse may have a different impact on the system than when change in beliefs are highly-homogenous. Indeed, the EBSL model here predicts as much: changes to other agents' beliefs in high-entropy situations should generally be less impactful to any one agent's beliefs than changes to other agents' beliefs in high-agglomeration (i.e., low-entropy) settings. Table 5.13 accounts for this fact by providing a 'density adjustment' to values from Table 5.11 (i.e., it rescales the percentage of agents moving in any period by the number of positions in the doxastic space with positive occupancy for that period).[180] This adjustment produces the surprising finding (albeit

[180] This density-adjustment is similar in principle to accounting for the "intensity of choice" phenomenon among investors within financial markets described by Brock and Hommes (1998).

predicted by the EBSL model) that the density-adjusted movements by agents in periods $t > 0$ tend to have more impact than those from period $t = 0$. That is, given the initial belief distributions used here, belief-echoing tends to be more impactful than does anticipatory updating. Moreover, such echoes can involve a very significant (density-adjusted) fraction of the agents, with up to a quarter of the density-adjusted market moving at once for periods as far out as $t = 3$. While the setup here is indubitably stylized, the possibility of one quarter of all market participants transacting at once (as might be a consequence of such belief-updating) would fairly likely have a large impact on financial markets, especially if all such transacting agents held very similar (but not fully-homogenous) beliefs. Importantly, for those situations wherein such a significant fraction of the density-weighted movement occurs in outer periods (e.g., nearly $25\%$ in $t = 3$ for the triple GSE regarding short-term beliefs; see Table 5.13), instigating signals are (in the EBSL model here) generally weak and mixed. Moreover, such sizable shifts due to belief-echoing in outer periods may be relatively unheralded by agents' belief-updating in earlier periods, as Table 5.14 suggests.

*Table 5.11: % of Agents Changing Position in Doxastic Space per Period*

| | | Short Term | | | | | Long Term | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Hypothesis Triple** | **Period, t =** | **0** | **1** | **2** | **3** | **4** | **0** | **1** | **2** | **3** | **4** |
| CGS | Maximum | *94.1%* | *94.0%* | *90.3%* | *32.9%* | -- | *93.4%* | *93.1%* | *91.8%* | *51.9%* | -- |
| | Average | *87.5%* | *74.9%* | *52.0%* | *2.9%* | -- | *86.9%* | *67.0%* | *32.4%* | *3.4%* | -- |
| CGE | Maximum | *94.1%* | *94.0%* | *90.3%* | *78.5%* | -- | *93.4%* | *92.8%* | *92.5%* | *81.5%* | -- |
| | Average | *87.7%* | *66.7%* | *46.0%* | *3.8%* | -- | *87.0%* | *65.7%* | *32.7%* | *4.4%* | -- |
| CGT | Maximum | *94.1%* | *92.9%* | *90.3%* | *75.0%* | -- | *93.4%* | *91.3%* | *91.8%* | *81.5%* | -- |
| | Average | *88.1%* | *70.7%* | *46.7%* | *3.4%* | -- | *87.2%* | *68.6%* | *32.7%* | *4.4%* | -- |
| CSE | Maximum | *94.1%* | *93.8%* | *90.7%* | *57.2%* | -- | *93.2%* | *92.6%* | *91.6%* | *81.2%* | -- |
| | Average | *87.7%* | *66.1%* | *42.2%* | *2.8%* | -- | *86.8%* | *65.5%* | *30.4%* | *4.5%* | -- |
| CST | Maximum | *94.1%* | *92.4%* | *90.7%* | *57.2%* | -- | *93.2%* | *91.8%* | *90.9%* | *81.2%* | -- |
| | Average | *88.0%* | *70.6%* | *41.8%* | *3.0%* | -- | *87.0%* | *68.6%* | *31.5%* | *5.1%* | -- |
| CET | Maximum | *94.1%* | *92.5%* | *92.1%* | *75.0%* | -- | *93.2%* | *90.1%* | *92.1%* | *86.3%* | -- |
| | Average | *88.3%* | *70.0%* | *44.7%* | *4.5%* | -- | *87.0%* | *66.7%* | *34.3%* | *7.9%* | -- |
| GSE | Maximum | *94.0%* | *93.8%* | *93.2%* | *74.6%* | *4.1%* | *93.4%* | *92.8%* | *92.5%* | *51.9%* | -- |
| | Average | *87.3%* | *68.0%* | *46.3%* | *3.9%* | *0.0%* | *86.4%* | *66.0%* | *35.6%* | *4.8%* | -- |
| GST | Maximum | *93.8%* | *92.6%* | *92.4%* | *75.0%* | -- | *93.2%* | *91.3%* | *92.1%* | *87.1%* | -- |
| | Average | *87.7%* | *70.7%* | *43.2%* | *2.5%* | -- | *86.6%* | *69.0%* | *35.5%* | *5.6%* | -- |
| GET | Maximum | *94.0%* | *92.2%* | *93.7%* | *62.8%* | -- | *93.4%* | *90.0%* | *92.6%* | *87.4%* | -- |
| | Average | *88.0%* | *70.2%* | *47.5%* | *4.3%* | -- | *86.8%* | *66.6%* | *35.4%* | *7.2%* | -- |
| SET | Maximum | *93.8%* | *92.4%* | *93.7%* | *83.4%* | -- | *93.4%* | *90.4%* | *92.1%* | *87.2%* | -- |
| | Average | *87.8%* | *69.9%* | *43.1%* | *3.9%* | -- | *86.6%* | *66.7%* | *33.1%* | *6.9%* | -- |

*Table 5.12: Attribution of Movements within Doxastic Space*

| Short Term | CGS | CGE | CGT | CSE | CST | CET | GSE | GST | GET | SET |
|---|---|---|---|---|---|---|---|---|---|---|
| Exogenous | *41.3%* | *47.1%* | *46.5%* | *48.6%* | *47.8%* | *47.5%* | *46.8%* | *47.5%* | *46.8%* | *47.9%* |
| Endogenous | *58.7%* | *52.9%* | *53.5%* | *51.4%* | *52.2%* | *52.5%* | *53.2%* | *52.5%* | *53.2%* | *52.1%* |
| **Long Term** | **CGS** | **CGE** | **CGT** | **CSE** | **CST** | **CET** | **GSE** | **GST** | **GET** | **SET** |
| Exogenous | *50.4%* | *51.1%* | *50.2%* | *51.8%* | *50.5%* | *51.1%* | *50.3%* | *49.3%* | *51.1%* | *51.5%* |
| Endogenous | *49.6%* | *48.9%* | *49.8%* | *48.2%* | *49.5%* | *48.9%* | *49.7%* | *50.7%* | *48.9%* | *48.5%* |

*Table 5.13: Density-Adjusted % of Agents Changing Position in Doxastic Space per Period*

| | | Short Term | | | | | Long Term | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Hypothesis Triple** | **Period, t =** | **0** | **1** | **2** | **3** | **4** | **0** | **1** | **2** | **3** | **4** |
| CGS | Maximum | *1.3%* | *5.2%* | *19.6%* | *12.9%* | -- | *1.5%* | *5.8%* | *19.4%* | *11.0%* | -- |
| | Average | *1.2%* | *2.4%* | *11.1%* | *4.9%* | -- | *1.4%* | *2.8%* | *7.4%* | *3.6%* | -- |
| CGE | Maximum | *1.4%* | *5.2%* | *19.6%* | *12.9%* | -- | *1.5%* | *5.8%* | *14.8%* | *20.4%* | -- |
| | Average | *1.3%* | *2.4%* | *9.9%* | *5.0%* | -- | *1.4%* | *2.7%* | *5.2%* | *3.4%* | -- |
| CGT | Maximum | *1.3%* | *5.3%* | *19.6%* | *12.9%* | -- | *1.4%* | *4.9%* | *14.8%* | *20.4%* | -- |
| | Average | *1.2%* | *2.5%* | *10.2%* | *5.0%* | -- | *1.3%* | *2.7%* | *5.5%* | *3.2%* | -- |
| CSE | Maximum | *1.2%* | *4.5%* | *19.4%* | *12.6%* | -- | *1.5%* | *5.4%* | *14.5%* | *20.3%* | -- |
| | Average | *1.1%* | *2.3%* | *9.5%* | *3.5%* | -- | *1.4%* | *2.9%* | *5.4%* | *3.4%* | -- |
| CST | Maximum | *1.3%* | *4.8%* | *19.4%* | *12.6%* | -- | *1.4%* | *5.0%* | *14.5%* | *20.3%* | -- |
| | Average | *1.2%* | *2.5%* | *9.4%* | *3.4%* | -- | *1.3%* | *2.8%* | *5.7%* | *3.7%* | -- |
| CET | Maximum | *1.4%* | *5.0%* | *18.8%* | *11.3%* | -- | *1.4%* | *3.0%* | *11.4%* | *21.6%* | -- |
| | Average | *1.3%* | *2.7%* | *7.9%* | *5.1%* | -- | *1.3%* | *2.1%* | *3.5%* | *7.2%* | -- |
| GSE | Maximum | *1.4%* | *4.9%* | *20.1%* | *24.9%* | *0.5%* | *1.6%* | *6.1%* | *17.4%* | *11.6%* | -- |
| | Average | *1.3%* | *2.5%* | *9.1%* | *4.5%* | *0.4%* | *1.5%* | *3.0%* | *5.7%* | *3.8%* | -- |
| GST | Maximum | *1.4%* | *6.1%* | *20.1%* | *12.9%* | -- | *1.5%* | *5.3%* | *17.4%* | *21.8%* | -- |
| | Average | *1.3%* | *2.8%* | *9.3%* | *3.4%* | -- | *1.4%* | *2.9%* | *5.7%* | *3.8%* | -- |
| GET | Maximum | *1.5%* | *5.7%* | *19.5%* | *12.9%* | -- | *1.5%* | *3.0%* | *11.5%* | *21.8%* | -- |
| | Average | *1.4%* | *2.8%* | *7.7%* | *5.1%* | -- | *1.4%* | *2.0%* | *2.8%* | *7.7%* | -- |
| SET | Maximum | *1.4%* | *5.4%* | *19.2%* | *20.8%* | -- | *1.6%* | *3.3%* | *13.1%* | *21.8%* | -- |
| | Average | *1.3%* | *2.8%* | *7.1%* | *3.9%* | -- | *1.5%* | *2.2%* | *2.6%* | *8.2%* | -- |

*Table 5.14: Examples of Large Outer-Periods Density-Weighted Shifts from Weak Signals*

| **Signal composition** | | | t = 0 | t = 1 | t = 2 | t = 3 |
|---|---|---|---|---|---|---|
| -2 | -1 | -2 | *0.3%* | *2.1%* | *4.1%* | *24.9%* |
| -2 | -1 | 3 | *1.3%* | *2.2%* | *4.1%* | *24.9%* |
| -2 | -1 | 4 | *1.3%* | *2.2%* | *4.1%* | *24.9%* |
| -2 | 1 | 2 | *0.3%* | *2.1%* | *4.1%* | *24.9%* |
| -2 | 1 | 3 | *1.3%* | *2.2%* | *4.1%* | *24.9%* |
| -2 | 1 | 4 | *1.3%* | *2.2%* | *4.1%* | *24.9%* |

Note: Initial conditions involved the hypothesis triple GSE over the short term.

How might such density-adjusting affect the attribution of movements to exogenous and endogenous influences? Table 5.15 demonstrates that such density-rescaling makes endogenous

movements far more influential than those which are attributable to the influence of the exogenous signal. Despite the stripped-down and simplistic nature of the model, it is interesting that these figures roughly agree with proportions calculated by Cutler, Poterba, and Summers (1993), who find that exogenous news only explains about one-third of activity in financial markets. Further, the fact that some of the more dramatic outer-period (echoing) activity can be generated with relatively muted signals suggests that EBSL may have an explanatory role to play in some of the more sizable and sudden moves by financial markets that were not preceded by noticeable arrivals of news. Specifically, given the several intervening periods that can lapse between the arrival of a relatively quiet and/or mixed signal and a large, endogenous shift among agents' beliefs due to the self-reinforcing nature of echoing, EBSL might be an influential force behind such peculiar events.

*Table 5.15: Density-Adjusted Attribution of Movements within Doxastic Space*

| Short Term | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **CGS** | **CGE** | **CGT** | **CSE** | **CST** | **CET** | **GSE** | **GST** | **GET** | **SET** |
| Exogenous | *14.1%* | *22.3%* | *22.5%* | *23.6%* | *24.3%* | *25.0%* | *21.9%* | *24.2%* | *25.0%* | *26.3%* |
| Endogenous | *85.9%* | *77.7%* | *77.5%* | *76.4%* | *75.7%* | *75.0%* | *78.1%* | *75.8%* | *75.0%* | *73.7%* |

| Long Term | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **CGS** | **CGE** | **CGT** | **CSE** | **CST** | **CET** | **GSE** | **GST** | **GET** | **SET** |
| Exogenous | *28.1%* | *30.2%* | *29.6%* | *30.9%* | *29.6%* | *36.6%* | *29.7%* | *28.4%* | *38.2%* | *39.1%* |
| Endogenous | *71.9%* | *69.8%* | *70.4%* | *69.1%* | *70.4%* | *63.4%* | *70.3%* | *71.6%* | *61.8%* | *60.9%* |

Finally, we might also consider how the simulations capture evolution of the average belief across all agents throughout the duration of the simulation. One way to undertake such an investigation is to measure total distance travelled in the doxastic space by the average belief across all agents. That is, we might track the movement of this average belief (or 'barycenter') within the doxastic space from one period to the next during each simulation. Recalling that the doxastic spaces here are three-dimensional (with effectively the unit interval comprising each dimension), we can

recognize that, in the space of the three-dimensional real numbers, the size of the simulated doxastic spaces is the same as for a unit cube. Using measures of Hamming (i.e., city-block) distance, then the furthest possible distance that the average an travel in any one period is 3 'units'. As the initial belief distributions are, however, rather diverse, maximum possible distance of travel will be substantially less, as the average in any one dimension will be nearer the midpoint of that dimension than an endpoint. Table 5.16 uses this approach to depict how far the average belief travels for the various hypothesis triples. As can once again be discerned, anticipatory updating accounts for only roughly half of the movement by the average relative to the overall Hamming distance which it moves over the simulation's duration. And, in many instances, this anticipatory updating accounts for only a small minority of such distance, with echoing (endogenous updating) accounting for as much as nearly 98% of the Hamming distance travelled overall for any duration.

*Table 5.16: Hamming Distances of Travel for True Average Belief*

| | | | CGS | CGE | CGT | CSE | CST | CET | GSE | GST | GET | SET |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Short Term** | | | | | | | | | | | | |
| **Travel During Period:** | t = 0 | Average | 0.47 | 0.47 | 0.47 | 0.47 | 0.47 | 0.47 | 0.47 | 0.47 | 0.47 | 0.47 |
| | | Maximum | 0.96 | 0.97 | 0.91 | 0.95 | 0.89 | 0.90 | 0.96 | 0.90 | 0.91 | 0.89 |
| | | Avg. % Total Travel | *48.2%* | *55.2%* | *53.3%* | *56.6%* | *54.4%* | *54.7%* | *54.9%* | *54.1%* | *54.3%* | *55.2%* |
| | | Min. % Total Travel | *2.3%* | *14.7%* | *14.1%* | *15.5%* | *14.5%* | *14.0%* | *10.5%* | *14.6%* | *14.1%* | *14.5%* |
| | t >= 1 | Average | 0.37 | 0.45 | 0.50 | 0.44 | 0.48 | 0.48 | 0.46 | 0.49 | 0.48 | 0.47 |
| | | Maximum | 0.96 | 0.96 | 1.02 | 0.95 | 0.98 | 0.98 | 0.98 | 0.99 | 0.99 | 0.98 |
| **Net Travel** | All Periods | Average | 0.47 | 0.52 | 0.50 | 0.55 | 0.53 | 0.53 | 0.52 | 0.52 | 0.51 | 0.55 |
| | | Minimum | 1.24 | 1.26 | 1.19 | 1.29 | 1.21 | 1.20 | 1.25 | 1.20 | 1.18 | 1.21 |
| **Long Term** | | | | | | | | | | | | |
| **Travel During Period:** | t = 0 | Average | 0.47 | 0.46 | 0.47 | 0.46 | 0.47 | 0.40 | 0.46 | 0.46 | 0.38 | 0.38 |
| | | Maximum | 0.85 | 0.86 | 0.81 | 0.85 | 0.81 | 0.60 | 0.87 | 0.82 | 0.55 | 0.54 |
| | | Avg. % Total Travel | *57.2%* | *58.6%* | *56.8%* | *59.2%* | *57.0%* | *56.3%* | *57.9%* | *56.1%* | *56.0%* | *56.5%* |
| | | Min. % Total Travel | *20.2%* | *20.9%* | *13.6%* | *20.9%* | *13.6%* | *13.7%* | *21.0%* | *13.6%* | *13.7%* | *13.7%* |
| | t >= 1 | Average | 0.43 | 0.40 | 0.44 | 0.40 | 0.44 | 0.40 | 0.42 | 0.46 | 0.39 | 0.38 |
| | | Maximum | 0.96 | 0.97 | 0.97 | 0.96 | 0.98 | 0.96 | 0.97 | 0.98 | 0.93 | 0.90 |
| **Net Travel** | All Periods | Average | 0.68 | 0.70 | 0.68 | 0.72 | 0.70 | 0.74 | 0.67 | 0.66 | 0.73 | 0.76 |
| | | Minimum | 1.34 | 1.35 | 1.35 | 1.37 | 1.36 | 1.35 | 1.33 | 1.33 | 1.31 | 1.33 |

### **5.6.3**\\ Discussion

Taken collectively, these investigatory simulations using the EBSL model suggest that endogenous belief changes in financial markets may be powerful forces behind some observable phenomena. Of course, the model here is rudimentary, and is simply a stripped-down tâtonnement process that abstracts away from risk preferences, endowments, prices, trading costs, and other realistic market features. Yet, simplicity of the EBSL model underscores some forces that other, more-detailed models can have difficulty pinpointing and studying clearly: mechanics of anticipation and reaction by agents of one another's beliefs when there exists substantial diversity among those beliefs (cf. Townsend (1983)). While further enrichment of the basic EBSL setup studied here will obviously be prudent for firmer conclusions to be comfortably asserted, it seems feasible that phenomena such as excess volatility and sudden, unheralded crashes in financial markets might be (at least partly) investigable in terms of the endogenous belief dynamics due to belief-echoing from investors' trust in their own beliefs shifting (sometimes abruptly) relative to that in others' beliefs.

And while the lack of some details in the EBSL model proposed here may dissuade definitive statements about whether or not it explains market reactions to the announcement by the US Government on a deal that 'resolved' the Fiscal Cliff matter, the results generated via simulation hint in just that direction: the strong and positive signal emitted by that announcement made for a pronounced upward surge in asset prices that persisted for a brief while. Yet the EBSL model is perhaps more insightful for exploring counterfactual histories, such as: what if instead a weak and mixed signal had been issued in the form of a failure to resolve the matter? The results here imply that such news could have been destabilizing for markets, either immediately, or after some delay. This statement requires some justification in terms of how stability manifests itself in EBSL terms.

As both the basic examples in Subsection 5.4.4 and the results from simulations in the current section suggest, systems wherein agents engage in entropy-based social learning are most stable at their extremes. When total doxastic agglomeration takes place agents' beliefs are homogenous, and they behave identically and follow Bayesian updating (albeit in adherence to the QET model of quantized beliefs) in response to exogenous signals. Likewise, when agglomeration is minimized (and entropy maximized) by having maximum heterogeneity in the distribution of initial beliefs, then echoing is more scarce, and of shorter duration when it does occur (see the results in Appendix 5.D, which repeat the simulated results in this section for a maximally-diverse initial belief distribution). What seems to be the most unstable configuration for systems subject to EBSL, however, are those of high (but not total) doxastic agglomeration with limited visibility.

As results from this paper demonstrate, that situation can give rise to substantial endogenous updating by agents of their beliefs that might persist for some time, cause dramatic shifts in the collective opinion, and evade *ex ante* detection for many periods. Such endogenous updating may appear relatively unpredictable if initial belief distributions are unknown, and go unheralded due to high sensitivity to weak signals. While the model here only dealt with one exogenous signal for each simulation, multiplicity of signals (with some signals arriving during or after anticipatory updating, for example; as well, early signals may leave belief distributions vulnerably unstable towards later signals) might cause further instability. But such a possibility requires further testing.

This section concludes with discussion of three points germane to foregoing observations on EBSL. The first involves lengths of the periods presumed in this setup. While it is most convenient, and perhaps obvious, to think of period lengths in the model as uniform, in natural economic settings the period lengths might be heterogeneous and potentially variable (see Mandelbrot and Hudson (2004), e.g., on "trading time"). Moreover, depending on the length of the periods chosen

for modelling, the interpretation of agents moving simultaneously within the period may be more or less realistic. It should be noted that the realism of results from the EBSL model (particularly in financial settings) may hinge crucially on such period lengths. It is also worth remarking that the EBSL model here could be converted into a sequential one, wherein each agent moves one-at-a-time, and does so with knowledge of the relevant agglomeration parameters and predecessors' moves. That extension may be suitable in various settings, but is, however, saved for later research.

The second point bears upon the relevance of geography. The EBSL model is inspired by geographic principles of agglomeration, distribution, and diversity, so it stands perhaps little wonder that it might generate the implication that location matters not in some isolated sense, but in a relational way.[181] That is, within the EBSL model, the position an agent occupies within doxastic space is not always of foremost importance. What often matters at least as much is her position in relation to the overall distribution of agents within that space. To rejigger a common contrivance from financial economics: whether some individual agent is a noise trader or smart-money trader is of relatively secondary importance to the agent, and what may matter first (and often utmost) is the specific proportion of noise traders to smart money (and diversity of the former); see, e.g., Shleifer and Summers (1990) on the sometimes murky border between the two.

Still, another implication of the EBSL approach is that location may matter outside the doxastic space: in the real world. An assumption of EBSL thus far has been that all agents access the same values for agglomeration, entropy, and juncture values. This assumption is fairly unrealistic, apart from highly-restricted settings wherein agents can directly observe one another's behaviors (which presents an interesting wrinkle in the home-bias puzzle surrounding preferences for local equity

[181] See, e.g., Bathelt and Glückler (2003; 2011).

investments: see Coval and Moskowitz (1999; 2001); Huberman (2001); Wójcik (2011)). But this restriction highlights exactly the significance of location: social learning may become even more erratic when observability is absent, especially if asymmetries are present in the extents to which some agents can observe others' behavior (e.g. via face-to-face contact; see, e.g.: Storper and Venables (2004)). Exactly how the EBSL model responds to heterogeneous assessments of agglomeration, juncture, and entropy by agents in it appears a subject ripe for future investigation.

Finally, the paper here, although not explicitly including the impact of prices, makes a case not for outright neglecting the role of price mechanisms within economic settings, but for instead rethinking the often near-complete hegemony granted to them in many financial and economic models and theories. The primacy of price as a mechanism for revealing the most fundamental mechanics of markets has recently been questioned by some, and the findings herein from exploring the EBSL model support the stance by a growing cadre of economic researchers calling for deeper investigation into the influences of other forces within markets (such as, e.g., volume; see specifically Hong and Stein (2007)). While the price mechanism surely cannot be discarded, (and the EBSL model may remain handicapped until it explicitly incorporates a role for price) it may be discerned from the current paper that other forces can be at least as interesting for research.

## **5.7**\\ Conclusion and Extensions

This paper has offered a psychological response to an economic question about the dynamics behind how much trust agents allocate to their own personal beliefs relative to the beliefs that they believe others hold. The approach taken here embedded one property of agents' beliefs, doxastic agglomeration, within the agents' environment itself. Combined with another property, the visibility to all agents of the average belief held among themselves (i.e., visibility of the average market belief), agglomeration of beliefs was used to construct an entropy-based social learning

(EBSL) model. The resulting EBSL model exhibits several pleasing properties for economic research, including: few free parameters (and endogenous self-tuning among some of these non-free parameters as the model evolves through time); high tolerance for belief diversity; and ability to collapse into the standard representative agent setup with Bayesian learning at an extreme. The model may resolve several longstanding issues with other approaches describing social learning, and gives rise to a new concept of endogenous learning: belief echoing. While effects from belief echoing are similar in many ways to those of informational cascades, various differences make distinctions between the two phenomena both helpful and worthwhile. In particular, belief echoing was shown to be useful in exploring and helping describe processes of endogenous belief changes.

The EBSL model erected in this paper was explored in two separate ways and settings. Firstly, some basic, easily-tractable examples were undertaken (in Subsection 5.4.4) to provide motivation and intuition. Then the model was deployed on a richer problem involving the beliefs of institutional investors (derived from a novel empirical dataset) about the impact of Congressional action on the US Fiscal Cliff episode from late 2012. The results of that numerical analysis of the model suggested that EBSL and its attending phenomena of belief echoing may be important forces behind some of the more anomalous, but impactful, occurrences within modern financial markets. Although the model presented in this paper is only a much-simplified recreation of actual market mechanics, it demonstrated that substantial richness can be extracted from the simple generative force of entropy-based social learning. Much work remains, however, to be done in formalizing, testing, and vetting both EBSL and belief echoing as more generally-applicable economic constructs. The remainder of this paper offers some ideas that advance that mission. (Other suggestions about future extensions and research likewise appear in the body of the paper.)

First, the paper has only dealt with a uniform partition of the probability scale by agents (due to a constant value of $\theta^{-1}$ across the entire range or probabilities). This approach was undertaken for simplicity, as a uniform value of $\theta^{-1}$ makes working with entropy concepts far easier. Importantly, however, agents have been empirically shown to often treat probabilities as nonlinear (as in prospect theory (Kahneman and Tversky 1979; Tversky and Kahneman 1992). Rook (2014a) offers preliminary suggestions for allowing nonlinear probabilities into QET (the psychological theory on which the EBSL model of this paper is largely based), and incorporation of those ideas and others into EBSL stands vital for future extensions of many ideas treated in the present paper.

Secondly, in the current EBSL model agents are all treated as having identical values for the cognitive parameters $\theta$, $\delta$, and $F$. Exploring plausible heterogeneity among these parameters across the agents poses a potential route for enriching EBSL, and could possibly be accomplished by fusing EBSL more neatly with ideas from studies of cognitive hierarchies and level-k thinking (see, e.g., Camerer, Ho, and Hong (2004), as well as Crawford, Costa-Gomes, and Iriberri (2013)).

Finally (as with its conceptual forebears of support theory (Tversky and Koehler 1994; Rottenstreich and Tversky 1997) and QET), the EBSL approach here does not explicitly specify the origins of either the hypotheses or evidence from which agents construct their beliefs. The sources and selection among various potential hypotheses and evidence is obviously vital for EBSL, and synthesis of ideas here with the sparsity-based techniques pioneered by Gabaix (2014) for modeling scarce allocation of attention may present one possible resolution to be pursued later.

In a more general vein, wider research by economists on the endogenous diversity of agents' beliefs, and how such endogenous heterogeneity may (reflexively) impact learning and wider economic systems seems an important future ambition to which EBSL may contribute in the future.

## APPENDIX 5.A\\ Extensions of Basic Examples (from Subsection 5.4.4)

### 5.A.1\\ Supplementary Example 1: 1-Dimensional, 3-Position Doxastic Space

Assumed values: $\theta = 3$; $\delta = 0.5$; $l(\theta, \delta) = 6$; $F = 1$; $\theta^F = 3$

*Table 5.A.1: Initial Evidence and Belief Distribution (Pre-Signal) for 3-Agent Example*

| | Own Belief | | | Belief about Average Belief | | |
|---|---|---|---|---|---|---|
| **Type** | $\hat{d}_A$ | $\hat{d}_B$ | Quantized | $\hat{d}_A$ | $\hat{d}_B$ | Quantized |
| $x_1$ | 1 | 5 | 0.167 | 1 | 5 | 0.167 |
| $x_2$ | 3 | 3 | 0.500 | 3 | 3 | 0.500 |
| $x_3$ | 5 | 1 | 0.833 | 5 | 1 | 0.833 |

- Echo durations (in parentheses) corresponding to signal strengths (in italics): *negative infinity through positive infinity* (0).

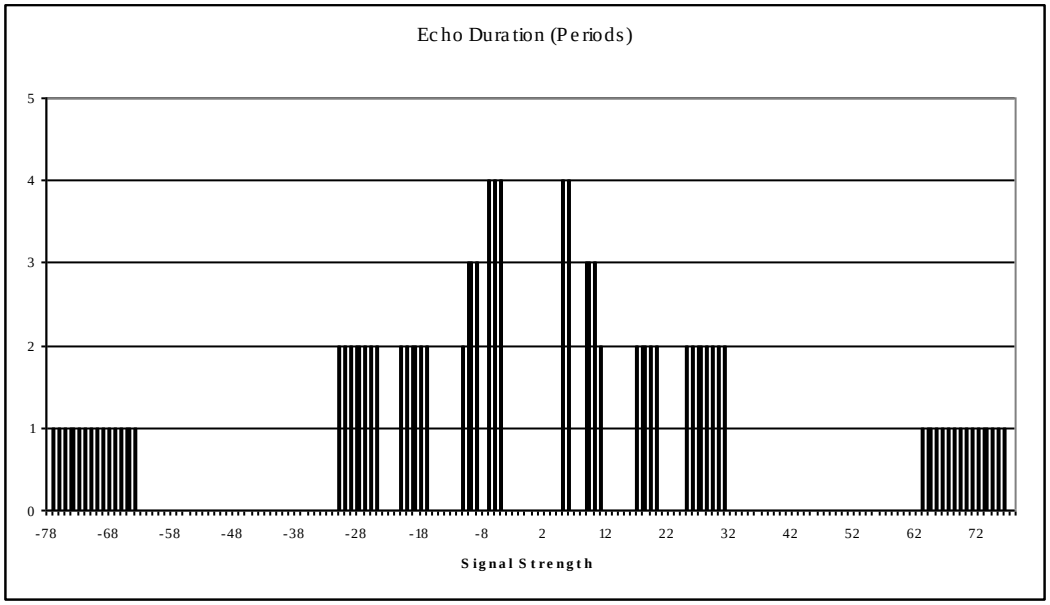
### 5.A.2\\ Supplementary Example 2: 1-Dimensional, 7-Position Doxastic Space

Assumed values: $\theta = 7$; $\delta = 0.5$; $l(\theta, \delta) = 14$; $F = 1$; $\theta^F = 7$

*Table 5.A.2: Initial Evidence and Belief Distribution (Pre-Signal) for 7-Agent Example*

| | Own Belief | | | Belief about Average Belief | | |
|---|---|---|---|---|---|---|
| **Type** | $\hat{d}_A$ | $\hat{d}_B$ | Quantized | $\hat{d}_A$ | $\hat{d}_B$ | Quantized |
| $x_1$ | 1 | 13 | 0.071 | 1 | 13 | 0.071 |
| $x_2$ | 3 | 11 | 0.214 | 3 | 11 | 0.214 |
| $x_3$ | 5 | 9 | 0.357 | 5 | 9 | 0.357 |
| $x_4$ | 7 | 7 | 0.500 | 7 | 7 | 0.500 |
| $x_5$ | 9 | 5 | 0.643 | 9 | 5 | 0.643 |
| $x_6$ | 11 | 3 | 0.786 | 11 | 3 | 0.786 |
| $x_7$ | 13 | 1 | 0.929 | 13 | 1 | 0.929 |

*Chart 5.A.1: Echo Durations for Signal-Strengths in 7-Agent Example*




- Echo durations (in parentheses) corresponding to signal strengths (in italics): *negative infinity through -78* (0); *-77 through -64* (1); *-63 through -32* (0); *-31 through -25* (2); *-24 through -21* (0); *-21 through -17* (2); *-16 through -12* (0); *-11* (2); *-10 through -9* (3); *-8* (0); *-7 through -5* (4); *- 4 through 0* (0); *1 through 4* (0); *5 through 6* (4); *7 through 8* (0); *9 through 10* (3); *11* (2); *12 through 16* (0); *17 through 20* (2); *21 through 24* (0); *25 through 31 (2)*; *32 through 62* (0); *63 through 76* (1); *77 through positive infinity* (0).

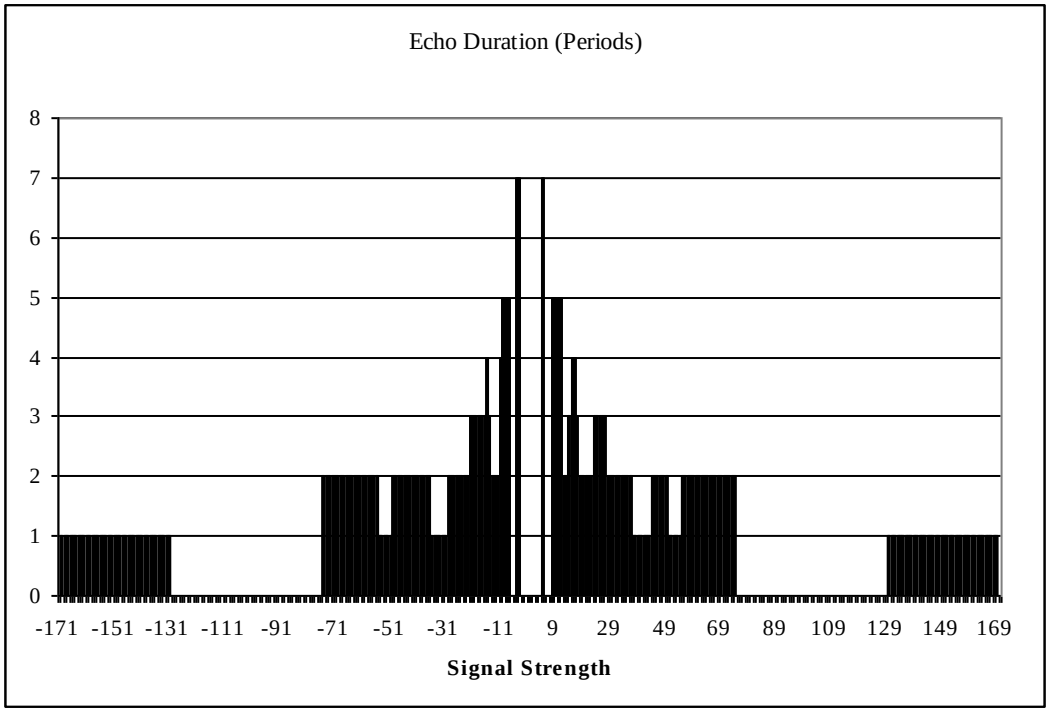
**5.A.3**\\ Supplementary Example 2: 1-Dimensional, 10-Position Doxastic Space

Assumed values: $\theta = 10$; $\delta = 0.5$; $l(\theta, \delta) = 20$; $F = 1$; $\theta^F = 10$

*Table 5.A.3: Initial Evidence and Belief Distribution (Pre-Signal) for 10-Agent Example*

| | **Own Belief** | | | **Belief about Average Belief** | | |
|---|---|---|---|---|---|---|
| **Type** | $\hat{d}_A$ | $\hat{d}_B$ | Quantized | $\hat{d}_A$ | $\hat{d}_B$ | Quantized |
| $x_1$ | 1 | 19 | 0.05 | 1 | 19 | 0.05 |
| $x_2$ | 3 | 17 | 0.15 | 3 | 17 | 0.15 |
| $x_3$ | 5 | 15 | 0.25 | 5 | 15 | 0.25 |
| $x_4$ | 7 | 13 | 0.35 | 7 | 13 | 0.35 |
| $x_5$ | 9 | 11 | 0.45 | 9 | 11 | 0.45 |
| $x_6$ | 11 | 9 | 0.55 | 11 | 9 | 0.55 |
| $x_7$ | 13 | 7 | 0.65 | 13 | 7 | 0.65 |
| $x_8$ | 15 | 5 | 0.75 | 15 | 5 | 0.75 |
| $x_9$ | 17 | 3 | 0.85 | 17 | 3 | 0.85 |
| $x_{10}$ | 19 | 1 | 0.95 | 19 | 1 | 0.95 |

*Chart 5.A.2: Echo Durations for Signal-Strengths in 10-Agent Example*




- Echo durations (in parentheses) corresponding to signal strengths (in italics): *negative infinity through -171* (0); *-170 through -131* (1); *-130 through -76* (0); *-75 through -56* (2); *-55 through -51* (1); *-50 through -37* (2); *-36 through -31* (1); *-30 through -23* (2); *-22 through -17* (3); *-16* (4); *-15* (3); *-14 through -12* (2); *-11* (4); *-10 through -8* (5); *-7 through -6* (0); *-5 through -4* (7); *-3 through 0* (0); *1 through 3* (0); *4* (7); *5 through 7* (0); *8 through 11* (5); *12 through 13* (2); *14* (3); *15 through 16* (4); *17* (3); *18 through 22* (2); *23 through 27* (3); *28 through 36* (2); *37 through 43* (1); *44 through 49* (2); *50 through 54* (1); *55 through 74* (2); *75 through 129* (0); *130 through 169* (1); *170 through positive infinity* (0).

## **APPENDIX 5.B**\\ Preparation Protocol for Fiscal Cliff Data

The original response pool from the survey of digital subscribers to the financial periodical, *Pensions & Investments*, included the stated beliefs of 813 individual (anonymous) respondents. Some of these respondents, however, either entered too many responses for one or more of the propositions posed to them, or else left blank a significant number of responses. The responses of two respondents were excluded because they entered multiple responses for one or more propositions. The responses of seven respondents were excluded because these individuals responded to fewer than four of the propositions about their own beliefs for either the short or long-term. The responses of 124 respondents were eliminated because they responded to fewer than four propositions about their beliefs on the average market belief for either the short or long terms. These exclusion rules were devised before the analysis was conducted, and resulted in a pool of usable responses consisting of beliefs for 680 respondents.

## **APPENDIX 5.C**\\ Supplementary Statistics on Fiscal Cliff Data

The distributions of beliefs reported by respondents on propositions related to the economic impact of Congressional action on the Fiscal Cliff episode (see Section 5.5 in the main text of the paper) exhibited significant differences between the reported beliefs for 'self' and 'the market'. To see this discrepancy in perhaps the most-simple format possible, see Table 5.C.1 below. Table 5.C.1 shows the percentages of respondents for each of the five macroeconomic variables that reported different beliefs for themselves and what they believed the market believes. As can be seen, between one-third and one-half of respondents reported differences between their own beliefs and the average belief in the market across all five variables and for both short- and long-term horizons.

*Table 5.C.1: Percentage of Respondents Reporting Different Beliefs for Self and Market*

| | Macroeconomic Variable | | | | |
|---|---|---|---|---|---|
| **Time Horizon** | **Confidence** | **Growth** | **Stability** | **Employment** | **Trade** |
| Short Term | *37.3%* | *42.2%* | *45.2%* | *43.4%* | *42.8%* |
| Lont Term | *44.5%* | *42.2%* | *43.0%* | *40.7%* | *37.9%* |

While such high proportions of divergence seem to present convincing evidence that the distributions of respondents' own beliefs and their beliefs on market belief are statistically distinct, some readers may crave a more statistically-rigorous approach. Such an approach, founded in differences in the entropy statistics of the two belief distributions, can be found in Miller and Maddow (1963). They supply a statistical test for identifying meaningful differences between the entropy values of two distributions, which relies on difference tests between the variances. Following Miller and Madow (see also Martin (1999a)), estimated entropies have limiting distributions that are normal, with zero-centered means, and variances which are equivalent to:

$$var(W) = \frac{1}{N}\sum_{n=1}^{N} p_n(\ln(p_n) + \rho_N)^2$$

Here the size of the doxastic space is equal to $N$, and $p_n$ is the proportion of agents $X$ in any one position $n$ within the space. Substitution of a particular entropy value (measured, estimated, or hypothetical) for $\rho_N$ provides an estimated variance that can be used for statistical tests of significant difference between two entropies (again, whether measured, estimated, or hypothetical). Specifically, the null hypothesis $\rho_1 = \rho_2$ can be tested using the combined variances $\sigma_1^2 + \sigma_2^2$ associated with each entropy $\rho_1$ and $\rho_2$. Note that variance increases with the value of $\rho$, and thus a conservative test might utilize the larger of $\rho_1$ and $\rho_2$ for $\rho$ in the above when computing the variances for $\sigma_1^2$ and $\sigma_2^2$.

We can conveniently use this statistical test to examine statistical differences between the two belief distributions in terms of their associated entropy values. Measuring the distributions in three dimensions at a time (i.e., following the "hypothesis triples" approach in Sections 5.5 and 5.6), with five positions per dimension (with $\theta^F = 5^3$, and therefore $N = 125$), we find that the initial distributions for respondents' own beliefs and their beliefs about the average market belief differ strongly: the null hypothesis is rejected comfortably for every such hypothesis triple across both the short and long terms for the following confidence levels: 90%; 95%; 99%; and 99.9%.

Likewise, the above tests of statistical differences in entropy values can be applied using the hypothetical maximum and minimum entropy values. Again applying such tests finds all four initial distributions of respondents' beliefs (i.e., their own beliefs on the short term, their own beliefs on the long term, their beliefs on the average market belief over the short term, and their beliefs on the average market belief over the long term) statistically differ from both maximum and minimum levels of heterogeneity for the confidence levels: 90%; 95%; 99%; and 99.9%.

## **APPENDIX 5.D**\\ Numerical Results – Maximum-Diversity Distribution

This appendix repeats the basic numerical analysis from Section 5.6 with a maximally-heterogeneous distribution for agents' own beliefs about the five macroeconomic variables. The setup and procedures follow nearly-exactly those from Section 5.6 (including ten possible hypothesis triples, a doxastic space of size $\theta^F = 5^3$, with $\delta = 0.5$, and $l(\theta, \delta)$ homogenous across all agents). The only distinction in this section is that the distribution of agents' own beliefs is presumed to meet the uniformity condition described in Subsection 4.4: the distribution of agents is uniform across the doxastic space, and all positions in the doxastic space have positive occupancy. For purposes here, there were presumed to be $X = 750$ such agents, with six agents in each position within the space initially (i.e., each position in the space having an 'occupancy rate' of 0.8% of the overall group ('the market') of agents. As was demonstrated for larger spaces in Subsection 5.4, it was found that the visibility condition was irrelevant at maximal heterogeneity: whether or not agents initially knew the correct average belief in the market had no impact whatsoever on the ensuing patterns of echoing or belief dynamics. Therefore the only results reported in this section are (for simplicity) assume each agent believes the average market belief corresponds exactly to her own personal beliefs.

The initial evidential distribution and quantized probabilities likewise correspond exactly to those from Table 5.6 within Section 5.5. As can be seen from Table 5.D.1, the numerical simulations conducted with maximally-diverse initial distributions tended to reduce echoing on average. Specifically, the average duration of echoes was significantly shorter (only about one period of echoing, on average across all simulations), whereas the likelihood of echoing (in terms of the percentage of simulations exhibiting some echoing) was equal to the lowest such percentage in the base simulations of Section 5.6. Also, it can be seen that the longest duration observed was four

periods (three echoing periods). Moreover, from Table 5.D.2, one can see that this increased initial diversity tended to persist: the average number of occupied positions in the doxastic space at the conclusion of every period was substantially greater than for the base simulations.

*Table 5.D.1: Distribution of Simulation Durations and Other Selected Statistics*

| **Total Duration (Periods)** | **0** | **1** | **2** | **3** | **4** | **Total** |
|---|---|---|---|---|---|---|
| Number of Simulations | 27 | 1296 | 5672 | 2222 | 44 | 9261 |
| % of Simulations | *0.3%* | *14.0%* | *61.2%* | *24.0%* | *0.5%* | *100.0%* |

| | |
|---|---|
| Average Duration | 2.10 |
| % with Echoing | *85.7%* |

*Table 5.D.2: Average Number of Occupied Positions in Doxastic Space after each Period*

| Time Period | t = -1 | t = 0 | t = 1 | t = 2 | t = 3 |
|---|---|---|---|---|---|
| Average Positions Occupied per Period | 125.00 | 47.40 | 19.80 | 19.04 | 19.02 |

Furthermore, it is visible from the distribution of simulations according to absolute total signal strength that substantial echoing (in terms of multiple periods of echoing) does not occur with the maximally-heterogeneous initial distribution without substantially-larger signal strengths (relative to the less-heterogeneous base-case from Section 5.6; notice that for that less-diverse case runs with multiple periods of echoing begin with absolute total signal strengths of $|2|$ and larger, whereas these longer echoes do not appear in the maximally-diverse case here until absolute total signal strengths of $|9|$ or greater are achieved. This finding suggests that highly-diverse initial belief distributions might be less likely to give rise to echoes which build in strength and result in sudden shifts in overall beliefs in a market, yet can go undetected for many periods before having any noticeable effects; instead, it seems that maximally-diverse initial distributions may be remarkably stable for EBSL. This stability is also evidenced by Tables 5.4 and 5.5, which show comparatively few agents (proportionally) changing their beliefs in outer periods relative to the

less-diverse situations explored in Section 5.6 of the main text. Significantly, these results suggest that large and sudden shifts multiple periods after arrival of an exogenous signal may be uncommon (for signal strengths explored here) with high diversity.

*Table 5.D.3: Distribution of All Simulation Durations for Each Signal Strength*

| Signal Strength | Average Duration | % of Simulations Per Duration (in Periods) 0 | 1 | 2 | 3 | 4 | 5 | % with Echoing | Number of Simulations |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.000 | 100.0% | -- | -- | -- | -- | -- | 0.0% | 1 |
| 1 | 0.000 | 100.0% | -- | -- | -- | -- | -- | 0.0% | 6 |
| 2 | 0.444 | 66.7% | 22.2% | 11.1% | -- | -- | -- | 11.1% | 18 |
| 3 | 1.053 | 21.1% | 52.6% | 26.3% | -- | -- | -- | 26.3% | 38 |
| 4 | 1.394 | -- | 60.6% | 39.4% | -- | -- | -- | 39.4% | 66 |
| 5 | 1.490 | -- | 51.0% | 49.0% | -- | -- | -- | 49.0% | 102 |
| 6 | 1.562 | -- | 43.8% | 56.2% | -- | -- | -- | 56.2% | 146 |
| 7 | 1.616 | -- | 38.4% | 61.6% | -- | -- | -- | 61.6% | 198 |
| 8 | 1.659 | -- | 34.1% | 65.9% | -- | -- | -- | 65.9% | 258 |
| 9 | 1.742 | -- | 30.7% | 65.6% | 2.5% | 1.2% | -- | 69.3% | 326 |
| 10 | 1.841 | -- | 27.9% | 62.2% | 8.0% | 2.0% | -- | 72.1% | 402 |
| 11 | 1.917 | -- | 25.0% | 60.0% | 13.3% | 1.7% | -- | 75.0% | 480 |
| 12 | 1.985 | -- | 21.2% | 60.6% | 16.8% | 1.5% | -- | 78.8% | 548 |
| 13 | 2.047 | -- | 17.3% | 62.0% | 19.3% | 1.3% | -- | 82.7% | 600 |
| 14 | 2.113 | -- | 14.5% | 61.0% | 23.3% | 1.3% | -- | 85.5% | 636 |
| 15 | 2.149 | -- | 12.2% | 60.7% | 27.1% | -- | -- | 87.8% | 656 |
| 16 | 2.200 | -- | 10.3% | 59.4% | 30.3% | -- | -- | 89.7% | 660 |
| 17 | 2.259 | -- | 8.6% | 56.8% | 34.6% | -- | -- | 91.4% | 648 |
| 18 | 2.329 | -- | 7.1% | 52.9% | 40.0% | -- | -- | 92.9% | 620 |
| 19 | 2.389 | -- | 5.6% | 50.0% | 44.4% | -- | -- | 94.4% | 576 |
| 20 | 2.411 | -- | 3.9% | 51.2% | 45.0% | -- | -- | 96.1% | 516 |
| 21 | 2.400 | -- | 1.8% | 56.4% | 41.8% | -- | -- | 98.2% | 440 |
| 22 | 2.356 | -- | -- | 64.4% | 35.6% | -- | -- | 100.0% | 360 |
| 23 | 2.250 | -- | -- | 75.0% | 25.0% | -- | -- | 100.0% | 288 |
| 24 | 2.143 | -- | -- | 85.7% | 14.3% | -- | -- | 100.0% | 224 |
| 25 | 2.048 | -- | -- | 95.2% | 4.8% | -- | -- | 100.0% | 168 |
| 26 | 2.000 | -- | -- | 100.0% | -- | -- | -- | 100.0% | 120 |
| 27 | 2.000 | -- | -- | 100.0% | -- | -- | -- | 100.0% | 80 |
| 28 | 2.000 | -- | -- | 100.0% | -- | -- | -- | 100.0% | 48 |
| 29 | 2.000 | -- | -- | 100.0% | -- | -- | -- | 100.0% | 24 |
| 30 | 2.000 | -- | -- | 100.0% | -- | -- | -- | 100.0% | 8 |

*Table 5.D.4: % of Agents Changing Position in Doxastic Space per Period*

| Time Period | t = 0 | t = 1 | t = 2 | t = 3 |
|---|---|---|---|---|
| Average % Agents Changing Location | *93.6%* | *73.2%* | *12.2%* | *0.2%* |
| Maximum % Agents Changing Location | *99.2%* | *100.0%* | *84.0%* | *36.0%* |

*Table 5.D.5: Density-Adjusted % of Agents Changing Position in Doxastic Space per Period*

| Time Period | t = 0 | t = 1 | t = 2 | t = 3 |
|---|---|---|---|---|
| Average % Agents Changing Location | *0.8%* | *2.1%* | *5.4%* | *9.0%* |
| Maximum % Agents Changing Location | *0.8%* | *3.7%* | *9.3%* | *9.0%* |

CHAPTER 6

# DISAGREEING ON AGREEING

## **6.1**\\ Introduction

Like proverbial beauty, consensus may (sometimes) lie in the eye of the beholder. Many focal economic problems are influenced by whether or not circumstances permit agents to agree to disagree. If agreeing to disagree is impossible, then, in general, consensus should eventually transpire via belief equilibration. Yet other situations of economic interest can exist wherein agents disagree about whether or not they even agree: asymmetric settings may arise wherein some agents falsely perceive a consensus, even though other agents recognize that no such agreement holds; and misperceived consensuses may frustrate processes of stable equilibration. This paper explores conditions in which disagreements on agreement stem from cognitive heterogeneity among agents.

Specifically, it studies social learning scenarios wherein agents face diverse constraints on their individual cognitive resources, and therefore may disagree about whether consensus indeed exists because they differ along one or more dimensions in their respective cognitive abilities. Identifying some specific (combinations of) dimensions of cognitive ability that can influence accuracy of agents' perceptions of consensus, and the extents to which they may do so, is this paper's objective.

To aid in the investigation, a 'working surrogate' is proposed for the concepts of common knowledge and common belief: *doxastic agglomeration*. Doxastic agglomeration was first introduced in a model of entropy-based social learning (EBSL) (Rook 2014b); it is a measure based on a decomposition of Shannon's (1948) informational entropy, and indicates the degree of dispersion among agents' beliefs in any suitably-defined belief (doxastic) space. Importantly, various components of the structure of such doxastic spaces are determined by cognitive parameters which, under EBSL, control how agents form and update beliefs (i.e., learn). The original exposition of doxastic agglomeration and EBSL treated these cognitive parameters as homogenous across agents; but this paper explores some crucial results from making them diverse.

A useful feature in representing common knowledge through doxastic agglomeration comes from how it allows fracturing of an otherwise infinite circularity in common knowledge. Simply put, conventional treatments of common knowledge require an infinite hierarchy of knowledge about each agent's own knowledge, and her knowledge of others' knowledge. Typically such a hierarchy is built from concatenated statements about her own knowledge of others' knowledge, and others' knowledge of her own knowledge, which volley back and forth *ad infinitum*. Likewise, the weaker concept of common belief also involves this infinite circularity, except with beliefs replacing knowledge (see: Monderer and Samet (1989); Brandenburger and Dekel (1987)). Doxastic agglomeration in the EBSL framework, however, maintains a clean division between each agent's own beliefs, and her beliefs about other agents' beliefs, by forging a connection between these two 'stacks' of beliefs only indirectly: via a mutual belief by all agents that an agglomeration (entropy) statistic summarizing the beliefs of all agents is accepted by all agents (i.e., they all believe it true).

This setup reinforces conclusions by Aumann (1976) and others that, with common knowledge about specific facets of their environments and identical priors, agents should not agree to disagree. Yet it adds the caveat that such results need not hold for agents who are on unequal cognitive footings: under cognitive diversity, agents with differing cognitive abilities may never reach agreement because they continually disagree about whether or not they are in a state of consensus.

Aumann (1976) famously proved that, given specific conditions (which include common knowledge of one another's posterior beliefs, and identical prior beliefs), consensus must arise. The ensuing economic literature on the impossibility of agreeing to disagree has in the interim become vast, and includes numerous elaborations on dynamics for convergence to consensuses.[182]

[182] Due to the magnitude of that literature only a selective listing is possible here. Some highlight works from that vein include, e.g.: Harrison and Kreps (1978); Milgrom (1981); Geanakoplos and Polemarchakis (1982); Milgrom and

These works on the dynamics of consensus regularly dovetail with, and reinforce, prominent theories concerning rational expectations equilibria (REE), especially in connection with revelation of private information in financial markets. Nonetheless, it is intriguing to observe (and is only infrequently remarked upon in the subsequent literature on common knowledge and consensus) that, in his seminal paper, Aumann (1976) passingly acknowledges that people's potentials for achieving consensus may occasionally be corrupted by systematic psychological biases (he specifically references the work of Kahneman and Tversky (1974), and also alludes to the content of some personal dialogue he had with Amos Tversky). Aumann's concerns should, however, largely apply not to systematic biases that affect all people equally, but rather that affect individuals differently. That is, Aumann expresses concerns that cognitive factors may preclude consensus; however, people subject to identical biases in their thinking may still be able to achieve consensus, as stands the case under *pluralistic ignorance*, whereby people privately reject some proposition for believing it untrue, but publicly accept it for falsely believing that all (or some majority of) others accept it, although in truth all (or a majority of) others also privately reject it.[183]

### **6.1.1**\\ Cognitive Heterogeneity and Consensus Misperception

Thus, rather than cognitive homogeneity, in which biases in thinking influence all agents equally, a more reasonable preoccupation for studies on emergence and stability of consensuses among bounded agents would seem to be with cognitive diversity: with agents heterogeneously-impacted

Stokey (1982); Tirole (1982); Bray (1982); Cave (1983); Mertens and Zamir (1985); Bacharach (1985); McKelvey and Page (1986); Brandenburger and Dekel (1987; 1993); Monderer and Samet (1989); Rubinstein (1989); Samet (1990; 2010); Rubinstein and Wolinsky (1990); Parikh and Krasucki (1990); Nielsen et al. (1990); Geanakoplos (1992); Vassilakis and Zamir (1993); Morris (1994; 1999); Sonsino (1995); Nielsen (1995); Neeman (1996a; 1996b); Heifetz (1996); Krasucki (1996); Hanson (1998; 2002); Ménager (2006; 2008); Cripps et al. (2008).

[183] An excerpt from the 1981 Sir Robert Menzies Lecture given at Monash University by former British Prime Minister Margaret Thatcher (in which she remarks on the topic of achieving political consensus) seems to aptly describe a situation of "consensus" reached through pluralistic ignorance: "To me consensus seems to be the process of abandoning all beliefs…in search of something in which no one believes, but to which no one objects".

by biases. Such diversity in how agents think then may make genuine consensuses elusive, and may enable agents to often disagree on whether or not some consensus truly exists (notice that in pluralistic ignorance all agents *publicly agree* that a consensus exists, while in reality it does not).

Still, we must recognize that cognitive diversity is not an idea unfamiliar to economics. Notably, in formulating cumulative prospect theory, Tversky and Kahneman (1992) observed considerable variability across experimental subjects in the values which were exhibited for the "diminishing sensitivity" parameter that prospect theory associates with an agent's nonlinear distortions of the probability scale (see also: Wu and Gonzalez (1996); Bruhin, Fehr-Duda, and Epper (2010)). Likewise, studies of quantal response equilibria (QRE) (McKelvey and Palfrey 1995; McKelvey and Palfrey 1998) find substantial diversity among subjects in the QRE error parameter (which is interpreted as a generic measure of subjects' cognitive abilities in playing games). More recently, empirically-motivated theories of cognitive hierarchy and level-k thinking (see the extensive review by Crawford, Costa-Gomes, and Iriberri (2013)) situate cognitive diversity front-and-center in explaining departures of players' collective behavior from what would be anticipated under cognitive homogeneity.[184] Still, despite the increasing acknowledgement of, and respect for, cognitive diversity in economic research on bounded rationality (in both theoretical and empirical work) the topic of diverse *consensus misperceptions* among agents remains largely underexplored.

Yet much psychological research on 'false-consensus effects' does focus on the interplay between cognitive diversity and misperceptions of the degree of agreement among people within groups. False-consensus effects (FCEs) have been a subject of both intense interest and puzzlement within

[184] Rogers, Palfrey, and Camerer (2009) discuss synthesizing QRE and cognitive hierarchy approaches. See, however, Haile, Hortaçsu, and Kosenok (2008) for some empirically-motivated criticisms of QRE concepts. Furthermore, realize that quantal response approaches depend crucially upon several common knowledge assumptions. Meanwhile, work on cognitive hierarchies and cognitive diversity in economics seems to have been renewed in earnest in the wake of papers by Stahl and Wilson (1994; 1995) and Nagel (1995) (although such ideas claim their roots in Keynes (1936)).

psychology for decades. Ostensibly-conflicting findings suggest that people sometimes overestimate, but at other times underestimate, the degree of consensus within a designated (sub)population (especially the extent to which others' beliefs will be the same as their own). This "relative" nature of FCEs has been noted by psychologists ever since FCEs were first described by Lee, Greene, and House (1977), who suggested that cognitive and perceptual mechanisms are the likely culprits behind people's distorted judgments about consensuses. And, in a famous meta-study of 115 experiments on FCEs, Mullen et al. (1985) found significant diversity in the extents to which people are susceptible to false perceptions of consensus, and lamented: "…the parameters of this phenomena are not yet well understood". Soon thereafter Marks and Miller (1987) offered one explanation by suggesting that "two or more specific mechanisms may operate simultaneously or in concert" to produce observed heterogeneity in FCEs, which suggests that the cognitive diversity driving FCEs may be multidimensional in nature (i.e., there may be more than one cognitive parameter or function that differs at any particular time, and thereby diverse mechanisms may exist through which cognitive diversity could drive agents' misperceptions of consensuses).[185]

In an important contribution, however, Dawes and Mulford (1996) furnished experimental evidence that people's perceptions of consensus need not always be false. Their studies showed that, when people perceive very high levels of consensus among a target group of other people, then such perceptions are often robustly accurate predictors of people's judgments about other people; but the studies also showed that such accuracy weakens when low levels of consensus (i.e., high levels of disagreement) are subjectively perceived. These findings prompted Dawes and Mulford to assert that the popular designation of 'false-consensus effect' can oftentimes be a

[185] Potential multidimensional heterogeneity in cognitive abilities should be unsurprising, as multidimensional diversity has been shown in other agent characteristics (e.g. preferences; see Blanco, Engelmann, and Norman (2008)).

misnomer. Hence, Dawes and Mulford suggested an alternative approach to studying the issue: comparing weights people give to their own beliefs against the weights they give to others' beliefs.

This redefinition, rooted in comparative belief-weighting, should appear reasonable to many behavioral economists. Specifically, in various non-cooperative settings, experimental research finds that agents sometimes behave as if overweighting their own beliefs and evidence relative to those of others; but at other times agents act as if they underweight their own evidence and beliefs (see Dominitz and Hung (2009) and references therein). Many of these economic experiments that suggest variability in the weighting of agents' own beliefs relative to others' beliefs do not, however, directly measure judgments of consensus, even though findings from psychology suggest that such data can help explain observed cross-contextual variability in own-vs-others weightings.

EBSL, however, directly embeds a primary role for consensus: it involves a convex combination of an agent's own (*primary*) beliefs with her beliefs about others' beliefs (*secondary* beliefs), whereby relative weights allocated to each are functions of her perception of the degree of consensus on the beliefs in question, which EBSL measures using the distributional entropy of all agents' beliefs. This relative allocation of weight to primary and secondary beliefs based upon the degree of perceived consensus allows EBSL to capture many rich psychological phenomena, such as the "comparative ignorance hypothesis" (Fox and Tversky 1995), which claims that agents trust their primary beliefs more when they see themselves as more knowledgeable than their peers, but trust their primary beliefs less when they feel ignorant relative to those peers. Through using the degree of perceived consensus to proxy agents' perceptions of relative ignorance (i.e., people often perceive agreement among others as more indicative of knowledgeability, and disagreement as more indicative of its absence) EBSL can capture this intriguing relativity in human judgment, which proves otherwise difficult to formulate within a model (Fox and Tversky 1995, p. 601-602).

Nevertheless, the original presentation of EBSL (Rook 2014b) took (as a first pass) perceptions of consensus by agents as uniformly accurate across agents. The present paper relaxes that simplifying assumption and allows agents to heterogeneously differ (but not necessarily all err) when perceiving consensuses. As will be shown, this alteration can give rise to situations wherein agents disagree (sometimes persistently) about whether or not some consensus exists among themselves; in this paper such situations are called *asymmetric consensus misperceptions* (ACMs).

### **6.1.2**\\ Dislodging Asymmetric Consensus Misperceptions

Situations of asymmetric consensus misperception (and thus disagreements about agreement) should be of especial interest to those economists who study forms of social interaction which may hinder all agents being fully-informed. In particular, financial markets might be one arena of application for concepts of consensus misperception dealt with by the present paper. Notions of consensus are inherent in ideas about market efficiency and rational expectations equilibria (as are common knowledge and/or common beliefs). Yet, if circumstances occur wherein one or more market participants is aware that no consensus exists, while the remaining participants erroneously believe all market participants' beliefs to be in agreement with each other, then this disagreement may go uncorrected for some time if the participants who are aware of it lack sufficient opportunity to correct it (e.g., face limits to arbitrage).[186] But, eventually, arrival of some new information (or other change in circumstance) may cause the disagreement to suddenly become revealed, which might suddenly increase uncertainty among those participants who were previously unaware of it. We could describe such a situation as the 'dislodging' of an asymmetric consensus misperception.

[186] This lack of common knowledge may give rise to (e.g.) price bubbles in financial markets, which might result in sudden and dramatic crashes (consider, e.g.: Allen, Morris, and Postlewaite (1993); Abreu and Brunnermeier (2003)).

Hence, if agents have preferences concerning uncertainty, and are sufficiently uncertainty-averse, then sudden injections of uncertainty about other participants' beliefs (due to revealed misperceptions of consensus; Crawford (1995) calls this uncertainty about other agents' beliefs "strategic uncertainty", whereas Avery and Zemsky (1998) label it as "compositional uncertainty"), might prompt reactions by those agents who previously – but falsely – perceived a consensus. And such reactions might easily become misinterpreted by a researcher as a change in uncertainty aversion (even though no genuine 'change' in uncertainty preference would have actually transpired) or risk aversion, unless it were otherwise accounted for. This paper does not attempt to distinguish between changes in uncertainty preferences among agents and changes in the degree of uncertainty perceived by agents. Conditions for dislodging asymmetric consensus misperceptions are, however, directly dealt with. It is observed that dislodging asymmetric misperceptions is an inherently dynamic process that can involve social learning, which makes EBSL an appropriate paradigm for studying such processes. Importantly, it is demonstrated that situations can occur which require a 'Goldilocks Condition' to obtain in order to dislodge an ACM.

It may be obvious that dislodging an asymmetric consensus misperception may sometimes require an informational signal of sufficient strength. Yet results in this paper show that such signals can sometime be too strong, in that they serve not to dislodge an ACM, but instead to simply push the misperceived consensus from one location (along the probability scale) to another. Succinctly, circumstances subject to the Goldilocks Condition require a (relative) signal strength that is 'just right' in order to dislodge an ACM. Arguments in later sections of this paper demonstrate that if such a range of just-right signal sizes is too narrow, then situations of asymmetrically misperceived consensuses may persist for quite some time, even in environments with high informational flows.

### **6.1.3**\\ Structure

This paper's remainder is as follows. Section 6.2 elaborates the problem of consensus misperception due to cognitive diversity by sketching a more formal depiction of the problem of agents' disagreements about consensuses among their subjective beliefs. Section 6.3 revisits the entropy-based social learning (EBSL) framework (Rook 2014b), and indicates various ways in which cognitive diversity is already permissible via the model's existing structure. Subsection 6.3.3 also introduces a new interpretation for total doxastic agglomeration and minimum entropy as the most noise-robust signals about the diversity among their peers' beliefs that agents can receive, and thus adds to the list of earlier justifications for why agents might resist departing from a strong consensus when they are aware of it. Section 6.4 explains how elements of the EBSL setup can substitute for conventional economic modeling tools of common knowledge and belief. Particularly, it discusses how breaking the otherwise inherent circularities in other approaches to common knowledge or belief may confer an operational advantage on the doxastic-agglomeration replacement proposed here. Section 6.5 looks at two further possibilities for injecting cognitive diversity into the EBSL setup: systematic misperceptions of consensus; and non-linearity in subjective probabilities. Section 6.6 explores conditions required for dislodging asymmetric consensus misperceptions under various configurations of cognitive heterogeneity. It identifies a 'Goldilocks Condition' that can be required to remove ACMs. Implications for rational expectations equilibria are discussed. Section 6.7 concludes and proposes links with other research.

## **6.2**\\ Disagreement from Diversity

This section provides a formal depiction of the problem of asymmetric consensus misperception, as well as unveils some of the ideas which this paper proposes for confronting the problem analytically. While a full model appears in Section 6.3, the section here may help with intuitions.

To fix ideas, let $i \in I$ index some finite collection of agents who are trying to discern the truthfulness of a single proposition (*focal hypothesis*) $A$ about the world (such a hypothesis will generally describe an event, but also may, with sufficient detail, describe a state). Each agent $i$ forms a belief regarding $A$'s truth based upon her accessible evidence (broadly understood as any information of which she is aware, possesses, and deems relevant in evaluating $A$'s truth). Such accessible evidence about $A$ involves: a set of evidence supporting the truth of $A$, $\boldsymbol{d}_{i,t}^{A}$ (where subscripts denote this evidence as accessible to agent $i$ at time $t \in \mathbb{Z}$); as well as a set of evidence, $\boldsymbol{d}_{i,t}^{\neg A}$, which supports an alternative hypothesis: $\neg A$ (the negation of $A$, where $A$ and $\neg A$ are mutually exclusive, and therefore collectively exhaustive, statements about the world: one must be true).[187] Assume both $\boldsymbol{d}_{i,t}^{A} \cap \boldsymbol{d}_{i,t}^{\neg A} = \emptyset$ and $\boldsymbol{d}_{i,t}^{A} \cup \boldsymbol{d}_{i,t}^{\neg A} \neq \emptyset$ are always true for every agent $i \in I$.

Also require that each agent express her subjective evaluation of accessible evidence as a belief $\pi_{i,t}\left(\boldsymbol{d}_{i,t}^{A}, \boldsymbol{d}_{i,t}^{\neg A}\right)$, which is a subjective probability locatable on the unit interval $C = [0,1] \subset \mathbb{R}$. That is, assume that each agent $i$ formulates a subjective probability about $A$'s truthfulness (relative to that for $\neg A$) by subjective assessment of $\boldsymbol{d}_{i,t}^{A}$ in relation to $\boldsymbol{d}_{i,t}^{A} \cup \boldsymbol{d}_{i,t}^{\neg A}$ (i.e., the stock of accessible evidence that supports $A$ relative to her entire stock of accessible evidence). Let $\theta_i \in \mathbb{Z}_+$ denote the 'discrimination' abilities of $i$: this cognitive parameter indicates how finely $i$ can evaluate the respective elements of $\boldsymbol{d}_{i,t}^{A}$ and $\boldsymbol{d}_{i,t}^{\neg A}$ when she forms $\pi_{i,t}\left(\boldsymbol{d}_{i,t}^{A}, \boldsymbol{d}_{i,t}^{\neg A}\right)$. Essentially, if an agent forms her subjective probability solely according to evidence, but has $\theta_i = \infty$, then she should enjoy (at least the potential for) infinite precision in formulating her belief regarding the truthfulness of $A$.

[187] Accessible evidence must be distinguished from available evidence, of which accessible evidence is a subset. Evidence available to an agent is all evidence that could *potentially* be used by that agent, but some evidence available to an agent may not be accessible to that (or any) agent, depending on various factors. The primary concern here is with cognitive factors that can make evidence differentially accessible to agents under a diversity of cognitive abilities.

Yet, for $\theta_i < \infty$, agent $i$'s precision in forming her belief may be limited if she faces restrictions on how much total evidence $\boldsymbol{d}_{i,t}^{A} \cup \boldsymbol{d}_{i,t}^{\neg A}$ she can summon to, and actively maintain in, her mind at once.[188] Particularly, under such dual limits, $i$ may be only able to formulate $\pi_{i,t}\left(\boldsymbol{d}_{i,t}^{A}, \boldsymbol{d}_{i,t}^{\neg A}\right)$ as a finite number of values on the unit interval. Effectively, $\theta_i$ partitions $C$ into $\theta_i$ (half-open) disjoint subintervals $c_v$, where $v \in \boldsymbol{V} = \{1, \ldots, \theta_i\} \subseteq \mathbb{Z}_+$ indexes these subintervals.[189] For simplicity, say (for the time being) that each such subinterval has uniform length $\theta_i^{-1}$, and also that $\theta_i$ is identical for all agents $i \in I$, with this shared discrimination parameter denoted by $\theta_I$. Under shared $\theta_I$, each agent's belief on $A$ (i.e., $\pi_{i,t}\left(\boldsymbol{d}_{i,t}^{A}, \boldsymbol{d}_{i,t}^{\neg A}\right)$) must fall into one of $\theta_I$ discrete, and equally-sized, subintervals $c_v$ of $C$. Let this 'discretization' of $i$'s beliefs be indicated by $\pi_{i,t}\left(\boldsymbol{d}_{i,t}^{A}, \boldsymbol{d}_{i,t}^{\neg A}\right)|\theta_I$. If $\pi_{j,t}\left(\boldsymbol{d}_{j,t}^{A}, \boldsymbol{d}_{j,t}^{\neg A}\right)|\theta_I = \pi_{k,t}\left(\boldsymbol{d}_{k,t}^{A}, \boldsymbol{d}_{k,t}^{\neg A}\right)|\theta_I$ for all agents $j \neq k \in I$, then it can be asserted that, at time $t$, all agents' beliefs are collocated on the unit interval *at the level of* $\theta_I$ (i.e., they all occupy the same subinterval $c_v$ based upon the partition of $C$ induced by an identical value of $\theta_i$ for all $i \in I$).

How might collocation (i.e., identical beliefs at the level of $\theta_I$, and thus a consensus among all of the agents' beliefs about truthfulness of $A$) become common knowledge among the entire assembly of these cognitively-bounded agents? Note that the partition induced by $\theta_I$ makes the distribution of agents among subintervals at $t$ amenable to description via Shannon's (1948) entropy formula:

$$\textbf{(1)} \qquad H_{\theta_I,t} = -\sum_{v=1}^{\theta_I} p_{v,t}\left(\log_2\left(p_{v,t}\right)\right)$$

[188] This latter limit we might say is a limit on memory (specifically: working memory), whereas the constraint due to $\theta_i$ restricts only the agent's attention, which, as shall be explicated later, is only one component of working memory.
[189] Experimental findings by De Giorgi and Reimann (2008) on beauty-contest games suggest that people may indeed use interval-based thinking in situations that demand reasoning about others under incomplete information. Indeed, much of the early empirical analysis of results from beauty-contest experiments uses methods based upon interval-based clustering of subjects' guesses; see, e.g.: Nagel (1995); Duffy and Nagel (1997); Ho, Camerer, and Weigelt (1998); Nagel et al. (2002). This interval-centricity of the outcomes (as well as the processes) in situations involving other-regarding strategic cognition seems to support using finite values for $\theta_i$ here. Likewise, one might also interpret the use of finite $\theta_i$ as a form of analogical reasoning with limited categories, such as in Jehiel (2005).

Above, $p_{v,t}$ is the proportion of agents occupying subinterval $c_v$ (indexed by $v$) at $t$.[190] (Per convention, let $p_{v,t}\left(\log_2\left(p_{v,t}\right)\right) = 0$ if $p_{v,t} = 0$; see Hanson (1998) and Sims (2003).) Note that entropy $H_{\theta_I,t}$ is minimized (at zero) when all agents' beliefs collocate in the same subinterval $c_v$. Call this minimum-entropy configuration of the $I$ agents' beliefs *total agglomeration*. There are $\theta_I$ total-agglomeration configurations when $\theta_I$ is identically-shared across all of the agents (i.e., total agglomeration could potentially occur within any one of the $\theta_I$ cells along the unit interval).

Assume (reasonably) that each agent has direct visibility on only her own beliefs. That all agents formulate beliefs – and must occupy *some* cell $c_v$ on $C$ at $t$ – remains required from above. Also let the value of $H_{\theta_I,t}$ be publicly announced.[191] With total agglomeration ($H_{\theta_I,t} = 0$) each agent then knows the locations of all agents, because she knows that all agents' beliefs are collocated at the level of $\theta_I$. Since each agent knows her own beliefs, and (presumably) knows she is a member of the set of all agents $I$, she therefore knows the beliefs of all $I$ agents (including here own), and also knows that all $I$ agents' beliefs agree: she knows there is unanimous, level-$\theta_I$ belief consensus. But can this knowledge become common? Given two knowledge conditions, that possibility exists.

**Condition K1**: All agents know that all agents know their own beliefs.

**Condition K2**: All agents know that $H_{\theta_I,t}$ is known to all agents.

Note that both conditions require mutual knowledge (adopting Brandenburger's (1992) definition of mutual knowledge as a fact known by all agents), but do not demand common knowledge.

[190] As Sims (2003) observes, the base to which logarithms are taken when working with entropy is immaterial, so long as one consistently uses that same base throughout. The base of 2 is used her as it greatly simplifies work with the information-theoretic concept of bits later in this paper. The base-2 formulation of entropy also has the pleasing property that, when proportions of agents are split evenly between two subintervals the entropy value then equals one.

[191] Requiring public announcement of $H_{\theta_I,t}$ is admittedly unrealistic, but is intended to underscore the fact that attainment of common knowledge (or common belief) in situations of social learning usually involves great difficulty.

Further realize that Condition K1 does not demand the higher-order iteration entailed by: all agents knowing that all agents know that all agents know their own beliefs (or other, higher-order iterations). (Also, Condition K1 is an extension to mutual knowledge of the 'positive introspection' axiom.) And Condition K2 is implied by $H_{\theta_I,t}$ being public, yet does not mandate that all agents know that all agents know that $H_{\theta_I,t}$ is mutual knowledge.[192] But, if both conditions hold at once, then common knowledge of one another's beliefs emerges for agents whenever $H_{\theta_I,t} = 0$ holds.[193]

Of course, verifying the pair of knowledge conditions may still pose a non-trivial challenge (especially when $I$ is large). Nevertheless, with finite $I$, adopting the above setup as a 'working surrogate' for common knowledge avoids an otherwise infinite amount of necessary verifications. And (as shown in Section 6.3) this total-agglomeration setup can extend to any finite number of hypotheses. Such added simplicity suggests that there may be merit in sometimes substituting it for common knowledge when the latter may otherwise be required. Yet the total-agglomeration approach's deeper value surfaces when knowledge may be imperfect and is thus replaced by belief.

Unlike knowledge (which is here assumed to be true by definition), beliefs can be false.[194] Condition K1 might therefore become weakened by requiring that: all agents believe that all agents believe their own beliefs. And likewise Condition K2 might be weakened to: all agents believe that $H_{\theta_I,t}$ is accepted by (and, by implication, made aware to) all. If these beliefs are true, then we still might expect that agents cannot agree to disagree; but, if these beliefs are false (for at least some agents), then agents may instead disagree about whether or not they even agree. And one

[192] Condition K2 extends Monderer and Samet's (1989) "evident knowledge". They claim that an event is "evident knowledge if whenever it occurs all agents know it". Condition K2 requires that when $H_{\theta_I,t}$ gets 'announced' its value becomes evident knowledge, and then that all agents know that this announced value is evident knowledge.
[193] Realize that the entropy function above is not stochastically monotone as per Bergin and Brandenburger (1990).
[194] While philosophically contestable, the assumption of knowledge as essentially true belief is one which is professed by Ramsey (1929), and seems to accord with conventional definitions for knowledge that appear widely in economics.

root of this potential disagreement as to whether or not agreement exists is cognitive diversity. That is, cognitive heterogeneity among agents may cause disagreements about whether they agree.

Recognize that if $\theta_j \neq \theta_k$ holds for some agents $j, k \in I$, then $j$ and $k$ possess distinct cognitive abilities which may cause them to hold divergent beliefs about the location of a consensus, or even on its existence (that is, its strength). For example, if belief entropy is announced publicly as being zero, but $\theta_j \neq \theta_k$, then, without seeing one another's beliefs or values for their respective discrimination parameters, $j$ and $k$ may both believe that their beliefs are in consensus, but disagree about what that consensus is.[195] Yet, if instead belief entropy must be inferred (e.g., even though no specific value gets announced publicly, some aggregate statistic is still constructible by each and every agent, despite the specific locations of every agent's beliefs not being made perfectly visible) by each agent, then, whenever $\theta_j \neq \theta_k$, $j$ and $k$ may disagree on whether they in fact agree.

For instance, if $\theta_j < \theta_k$ (i.e., $j$'s partition of the unit interval is coarser than $k$'s), then it is easy to envision a situation wherein (if $j$ and $k$ are the only two agents in $I$) that $j$ falsely perceives full consensus, yet $k$ perceives no such consensus. (Hereafter, consensus will refer only to total agglomeration – i.e., full consensus in beliefs – unless otherwise stated.) Say, e.g., $\theta_j = 3$ and $\theta_k = 6$ and that these values cause two distinct, uniform partitions of the unit interval (where $j$ faces a partition with three subintervals each of length $\frac{1}{3}$, and $k$ faces a partition with six subintervals each of length $\frac{1}{6}$). If $j$'s subjective belief falls in her first subinterval $c_1^j$, and that of $k$ falls in her second subinterval $c_2^k$, then (with only these two agents) $j$ will infer zero entropy and believe a consensus in beliefs prevails (as her first subinterval also contains $k$'s second

[195] Notice that such a public announcement of zero entropy would have to be at the level of $\max(\theta_j, \theta_k)$.

subinterval), whereas $k$'s inference will be a belief entropy of 1, and no consensus: they disagree on agreeing, and we should be entitled to say that this disagreement is due to asymmetric consensus misperception, wherein one agent falsely perceives a consensus, whilst the other perceives none.[196]

It will be useful to state more precisely when asymmetric consensus misperception could arise among any pair of agents. First (and tidying earlier notation slightly), let, for two agents, $j \neq k \in I$, $\pi_{j,t}(A)|\theta_j \in (0,1) \subset \mathbb{R}_+$ denote the discretized belief $\pi_{j,t}\left(\boldsymbol{d}_{j,t}^{A}, \boldsymbol{d}_{j,t}^{\neg A}\right)|\theta_j$ for $j$ on $A$ at $t$, and likewise allow $\pi_{k,t}(A)|\theta_k \in (0,1) \subset \mathbb{R}_+$ denote $k$'s discretized belief $\pi_{k,t}\left(\boldsymbol{d}_{k,t}^{A}, \boldsymbol{d}_{k,t}^{\neg A}\right)|\theta_k$ on $A$ at $t$. The openness of the interval $(0,1)$ indicates that neither agent holds sure beliefs (and will be justified in the next section). Let $c_{v^*}^{j}$ be the subinterval $c_v$ which contains $j$'s discretized belief such that $\pi_{j,t}(A)|\theta_j \in c_{v^*}^{j} \subset (0,1)$, and let $c_{v^-}^{j}$ be the remainder of the interval $(0,1)$ not containing $\pi_{j,t}(A)|\theta_j$ (i.e., $\pi_{j,t}(A)|\theta_j \notin c_{v^-}^{j} \subset (0,1)$) and together $c_{v^*}^{j}$ and $c_{v^-}^{j}$ exhaust the interval $(0,1)$ (i.e., $\left(c_{v^*}^{j} \cup c_{v^-}^{j}\right) = (0,1)$). This setup implies that $\left(c_{v^*}^{j} \cap c_{v^-}^{j}\right) = \emptyset$. Define $c_{v^*}^{k}$ and $c_{v^-}^{k}$ similarly for $k$, such that $\pi_{k,t}(A)|\theta_k \in c_{v^*}^{k} \subset (0,1)$ and $\pi_{k,t}(A)|\theta_k \notin c_{v^-}^{k} \subset (0,1)$, with $\left(c_{v^*}^{k} \cup c_{v^-}^{k}\right) = (0,1)$ and $\left(c_{v^*}^{k} \cap c_{v^-}^{k}\right) = \emptyset$. Moreover, do not impose any *a priori* restrictions upon the equivalency (or, alternatively, non-equivalency) of the parameter values for $\theta_j$ and $\theta_k$. Under the foregoing setup, the five conditions stated within Theorem 1 below suffice, all else being equal, to cause an

[196] The situation here is similar in spirit to that described by Billot et al. (2000) concerning shared beliefs. That paper looked at agreement (shared beliefs) from the perspective of overlap among agents' beliefs when agents entertain multiple priors (or posteriors) at once. The circumstances here are, however, slightly different because of the asymmetry of overlap from each agents' perspective: one agent perceives overlapping beliefs, while the other agent does not. The potential for cognitive heterogeneity naturally leads one to wonder whether those agents with lower cognitive capacities, such as working memory, might not be squeezed out of markets due to competitive forces. Yet, recent cognitive research suggests that some forms of learning are actually expedited by weaker working memory (see DeCaro, Thomas, and Beilock (2008), for instance). That weaker cognitive abilities could sometimes advantage particular forms of learning over others might lend a cognitive explanation for the results of Gayer and Gilboa (2014).

asymmetric consensus misperception, wherein $j$ falsely perceives a consensus between herself and $k$ on the basis of entropy $H_{\theta_j,t}$ even though agent $k$ does not perceive consensus according to $H_{\theta_k,t}$.

> **Theorem 1**: *Agent $j$ will perceive a consensus between her beliefs and those of another agent $k \neq j$, whereas $k$ will not perceive any consensus between her beliefs and those of $j$, whenever the following five conditions all hold:*
>
> (i) $c_{v^*}^j \neq c_{v^*}^k$
> (ii) $\left(c_{v^*}^j \cap c_{v^*}^k\right) \neq \emptyset$
> (iii) $\left(c_{v^*}^j \cap c_{v^-}^k\right) \neq \emptyset$
> (iv) $\pi_{k,t}(A)|\theta_k \in \left(c_{v^*}^j \cap c_{v^*}^k\right)$
> (v) $\pi_{j,t}(A)|\theta_j \in \left(c_{v^*}^j \cap c_{v^-}^k\right)$

All else being equal, the five conditions in Theorem 1 are collectively necessary and sufficient for an asymmetric consensus misperception (Appendix 6.A contains a proof).[197] Several results are immediate from this observation. First, for agents with unbounded discrimination abilities, ACMs should not exist. That is, for $\theta_j = \infty$ and $\theta_k < \infty$, agent $j$ should never falsely perceive a consensus between her beliefs and those of agent $k$ whenever $k$ perceives no consensus, because $\theta_j \to \infty$ requires that length of $c_{v^*}^j$ approach zero (via $\theta_j^{-1} \to \frac{1}{\infty}$) such that $c_{v^*}^j \subset c_{v^*}^k$ must arise and therefore $\left(c_{v^*}^j \cap c_{v^-}^k\right) = \emptyset$ must hold, which disallows agent $j$ from perceiving a consensus when $k$ does not.

Second, with $\theta_j, \theta_k < \infty$ but $\theta_j \neq \theta_k$ it should generally be possible that agent $k$ could possess a value of $\theta_k$ such that $c_{v^*}^j \not\subset c_{v^*}^k$, and that the five conditions in Theorem 1 hold so that agent $j$ misperceives a consensus between her beliefs and those of agent $k$ while $k$ perceives no consensus. Briefly, these two corollary results suggest that agents with infinite values for $\theta_i$ are not subject to ACMs, whereas any agent with $\theta_i < \infty$ could, at least hypothetically, become susceptible to ACM.

[197] Notice that, with only slight modification, these results become extensible to any finite number of agents.

Third, notice that *symmetric* consensus misperception might still arise even if Conditions (iii) and (v) are violated. That is, it could be that $\left(c_{v^*}^{j} \cap c_{v^*}^{k}\right) \neq \emptyset$ and $\pi_{k,t}(A)|\theta_k \in \left(c_{v^*}^{j} \cap c_{v^*}^{k}\right)$, but also $\left(c_{v^*}^{j} \cap c_{v^-}^{k}\right) = \emptyset$ and $\pi_{j,t}(A)|\theta_j \in \left(c_{v^*}^{j} \cap c_{v^*}^{k}\right)$, even though $\pi_{j,t} \neq \pi_{k,t}$. This possibility of both agents perceiving their belief to be in agreement when they are only approximately so stems from the fact that, under the setup here, consensus for both needs only to exist at the level of $\max(\theta_j, \theta_k)$.

Theorem 1 thus yields the possibility that consensus misperception might be a regular finding among agents with bounded cognitive abilities. One might suspect, however, that a difficulty posed by Theorem 1 is its requirement of *all else being equal*. Later sections of this paper show that other forms of cognitive diversity among agents may exist beyond simply that of $\theta_j \neq \theta_k$. The natural question then becomes: which forms of cognitive heterogeneity make Theorem 1 both necessary and sufficient for asymmetric consensus misperception to arise, and which do not? The entropy-based social learning framework introduced by Rook (2014b) suggests several dimensions along which agents might differ in cognitive capacity, and the rest of this paper explores the implications of such cognitive heterogeneity (along one or more dimensions) for misperceptions of consensus.

## **6.3**\\ Entropy-Based Social Learning

This section condenses the entropy-based social learning (EBSL) model (Rook 2014b), and quantized evidence theory (QET) (Rook 2014a), on which EBSL is based. This condensation is brief and intended mainly to highlight potential for introducing cognitive heterogeneity into the existing EBSL framework. For full motivation and theoretical underpinnings for EBSL and QET, readers are directed to the original expositions in Rook (2014a; 2014b) and the references therein.

### **6.3.1**\\ Quantizing Evidence

Continuing from the setup in Section 6.2, presume that each agent $i \in I$ seeks to construct subjective beliefs about the truthfulness of $F$ separate hypotheses, wherein each focal hypothesis $A_f$ is attended by its mutually-exclusive alternative (i.e., its negation), $\neg A_f$, and both are indexed by $f \leq F$, $f \in \mathbb{Z}_+$. For unbounded agents $F \to \infty$ might be true, but bounded agents will usually have small values for their respective parameters $F$ (e.g., $F < 10$).[198] (Recognize that, unlike $\theta_i$, the cognitive parameter $F$ is not presented with an agent-specific subscript; while $F$ plausibly varies across individuals, it cannot be fairly claimed that agents 'disagree' on hypotheses that one or more of them fails to consider, and hence this parameter is, for the present paper at least, treated as homogenous across agents, with diversity of $F$ across both agents and time left for future work.)

Furthermore, a cognitively-bounded agent $i$ may not consider all evidence and counterevidence ($\boldsymbol{D}_{i,t}^{A_f}$ and $\boldsymbol{D}_{i,t}^{\neg A_f}$, respectively) available to her at time $t$. Instead she may only be able to access subsets of the available evidence and counterevidence on hypotheses. Denote these accessible subsets as $\boldsymbol{d}_{i,t}^{A_f} \subseteq \boldsymbol{D}_{i,t}^{A_f}$ and $\boldsymbol{d}_{i,t}^{\neg A_f} \subseteq \boldsymbol{D}_{i,t}^{\neg A_f}$. Also assume that in making probability judgments each agent uses the cardinalities of accessible subsets, which are denoted by: $\widehat{\boldsymbol{d}}_{i,t}^{A_f}, \widehat{\boldsymbol{d}}_{i,t}^{\neg A_f} \in \mathbb{Z}_+$. (As both $\widehat{\boldsymbol{d}}_{i,t}^{A_f}$ and $\widehat{\boldsymbol{d}}_{i,t}^{\neg A_f}$ are strictly-positive integers, the framework applies only when agents' priors are non-dogmatic.) Using the cardinalities causes agents to effectively discretize accessible evidence.

But confining the evidential bases for agents' beliefs to being discrete and integer-valued proves to be only minimally consequential if the sizes of subsets of evidence $\boldsymbol{d}_{i,t}^{A_f}$ and counterevidence

[198] The classic study on limits to working memory by Miller (1956) would, however, put this limit at less than 10.

$\boldsymbol{d}_{i,t}^{\neg A_f}$ on $A_f$ are unrestricted. To see this fact, assume that, for each of the $F$ focal hypotheses, each agent $i$ combines $\widehat{\boldsymbol{d}}_{i,t}^{A_f}$ with $\widehat{\boldsymbol{d}}_{i,t}^{\neg A_f}$ into what is hereafter called an *intermediate evidence ratio* (IER):

**(2)** $$\sigma_{i,t}(A_f, \neg A_f) \equiv \frac{\widehat{\boldsymbol{d}}_{i,t}^{A_f}}{\widehat{\boldsymbol{d}}_{i,t}^{A_f} + \widehat{\boldsymbol{d}}_{i,t}^{\neg A_f}}$$

Notice that $\sigma_{i,t}(A_f, \neg A_f)$ can take any positive rational value inside the unit interval (except 1, because $\widehat{\boldsymbol{d}}_{i,t}^{A_f}$ and $\widehat{\boldsymbol{d}}_{i,t}^{\neg A_f}$ must both be positive), and therefore, without further restricting $\widehat{\boldsymbol{d}}_{i,t}^{A_f}$ or $\widehat{\boldsymbol{d}}_{i,t}^{\neg A_f}$, the IER may have near-infinite precision in the unit interval (because the infinite number of values that such a ratio of integers might take is ultimately countable, rather than uncountable, as it would otherwise be if $\sigma_{i,t}(A_f, \neg A_f)$ could take any positive real value inside the unit interval). Near-infinite precision nonetheless seems to represent a poor description of the capabilities of cognitively-bounded agents, and we may thus find it comfortable to limit $\widehat{\boldsymbol{d}}_{i,t}^{A_f}$ and $\widehat{\boldsymbol{d}}_{i,t}^{\neg A_f}$ as follows:

**(3)** $$\left(\widehat{\boldsymbol{d}}_{i,t}^{A_f} + \widehat{\boldsymbol{d}}_{i,t}^{\neg A_f}\right) \leq l(\theta_i, \delta_i)$$

Where:

**(4)** $$l(\theta_i, \delta_i) \equiv \theta_i (1 - \delta_i)^{-1}$$

The boundary function $l(\theta_i, \delta_i)$ is a composite cognitive limit on the working memory of the agent $i$, whereby that agent's cognitive abilities are constrained due to the interaction between her: capacity to discriminate (attention), as determined by $\theta_i \in \mathbb{Z}_+$; and volume of focus (memory), as governed by $\delta_i \in (0,1] \subset \mathbb{R}_+$.[199] (Note that the discrimination parameter $\theta_i$ controls how finely agent $i$ can distinguish between units of evidence in *generating* beliefs, and that she might be able

[199] See Rook (2014a).

to more finely distinguish between exogenous stimuli presented to her, i.e., those which she does not need to summon from memory; this particular distinction underscores that situations of thinking under uncertainty might be cognitively divergent from those under risk (Wakker 2004).)

A cognitively-unbounded agent $j \in I$ may enjoy $\theta_j \to \infty$ and $\delta_j = 1$. Yet for the most cognitively-restricted agent $i \in I$, both $\theta_i$ and $\delta_i$ will approach zero. (The original presentations of EBSL and QET (Rook 2014b; Rook 2014a) took both $\theta_i$ and $\delta_i$ – and thereby $l(\theta_i, \delta_i)$ – to be homogenous across agents, which implies no cognitive diversity.) The joint cognitive limit $l(\theta_i, \delta_i)$ for each agent $i \in I$ is, however, only a 'potential', in that less evidence might exist (or be available to any, or every, agent) than what is allowed by $l(\theta_i, \delta_i)$ (in which case there is 'spare thinking capacity'; the bearing of the amount of such spare capacity upon subjective judgment will be treated shortly).

A chief impetus behind QET is the comparability between human judgments of numerical values and quantization mechanisms from signal-processing applications. Quantization transforms continuous (analog) inputs values into discrete (digital) outputs. This transformation-by-quantization is generally compressive in nature (and thus saves on both memory and processing resources), albeit 'lossy' (that is, it tends to introduce errors proportional to the degree of compression). QET asserts that quantization techniques from signal-processing are suitable approximations for the ratio-dependent errors often observed for judgments of numerical values.[200]

The basic forward-quantization process proposed by QET to describe how IERs are transformed into subjective probability judgments under ratio-dependent limits to cognition $l(\theta_i, \delta_i)$ is detailed below. The parameter $\theta_i$ partitions the unit interval $C = [0,1] \subset \mathbb{R}$ into half-open, disjoint subintervals: *decision intervals*, $c_v^i \subset C$. Under *uniform quantization*, each $c_v^i$ has an equal length

[200] See especially the cognitive science research on 'approximate number systems'; Rook (2014a) provides a summary.

of ${\theta_i}^{-1}$. Each decision interval's endpoints are determined by a quantization index: $v \in \boldsymbol{V} = \{1, \ldots, \theta_i\} \subseteq \mathbb{Z}_+$. The lower endpoint of each $c_v^i$ is $\left((v-1){\theta_i}^{-1}\right)$, and upper endpoint is $\left(v{\theta_i}^{-1}\right)$.

An *index transmission function* then finishes the mapping $\sigma_{i,t}(A_f, \neg A_f) \rightarrow C$ (given some $l(\theta_i, \delta_i)$) to arrive at a quantized probability judgment $Q\left(\sigma_{i,t}(A_f, \neg A_f)\right)$. QET presumes the form:

$$\textbf{(5)} \qquad Q\left(\sigma_{i,t}(A_f, \neg A_f)\right) = {\theta_i}^{-1}\left(\sum_{v=1}^{\theta_i} 1_{c_v^i}\left(\sigma_{i,t}(A_f, \neg A_f)\right) S(v, \delta_i)\right)$$

Above, $1_{c_v^i}(\cdot)$ is an indicator function, where $1_{c_v^i}\left(\sigma_{i,t}(A_f, \neg A_f)\right) = 1$ if $\sigma_{i,t}(A_f, \neg A_f) \in c_v^i$, but $1_{c_v^i}\left(\sigma_{i,t}(A_f, \neg A_f)\right) = 0$ if $\sigma_{i,t}(A_f, \neg A_f) \notin c_v^i$. This indicator function picks out which decision interval corresponds to the inputted intermediate evidence ratio $\sigma_{i,t}(A_f, \neg A_f)$ (i.e., it operates as a *decision interval lookup function* would within the signal-processing versions of quantization.) After the appropriate decision interval is identified, an *encoder mapping* step selects the value $v$ from the quantization index that is suitable for that decision interval. Encoder mapping results from a value equal to the product of $1_{c_v^i}\left(\sigma_{i,t}(A_f, \neg A_f)\right)$ and a *switching function* $S(v, \delta_i)$ equaling:

$$\textbf{(6)} \qquad S(v, \delta_i) = s_{up} = v - \delta_i, \qquad \text{if } \left(\left(\widehat{\boldsymbol{d}}_{i,t}^{A_f} + \widehat{\boldsymbol{d}}_{i,t}^{\neg A_f}\right) / \, l(\theta_i, \delta_i)\right) \geq \tau$$

$$S(v, \delta_i) = s_{down} = v + \delta_i - 1, \qquad \text{if } \left(\left(\widehat{\boldsymbol{d}}_{i,t}^{A_f} + \widehat{\boldsymbol{d}}_{i,t}^{\neg A_f}\right) / \, l(\theta_i, \delta_i)\right) < \tau$$

And, furthermore:

$$\textbf{(7)} \qquad c_v^i = \left[(v-1){\theta_i}^{-1}, v{\theta_i}^{-1}\right), \qquad \text{for } s_{up}$$

$$c_v^i = \left((v-1){\theta_i}^{-1}, v{\theta_i}^{-1}\right], \qquad \text{for } s_{down}$$

Selection of $S(v, \delta_i) = s_{up}$ or $S(v, \delta_i) = s_{down}$ (i.e., whether the agent's ultimate probability judgments is 'slanted' upward or downward within the decision interval) thus becomes a matter of

how much relative weight each endpoint within $c_v^i$ receives from the focusing parameter $\delta_i$. This relative weighting is determined by the amount of 'spare cognitive capacity' (as dictated by the ratio $\left(\left(\widehat{\boldsymbol{d}}_{i,t}^{A_f} + \widehat{\boldsymbol{d}}_{i,t}^{\neg A_f}\right) / l(\theta_i, \delta_i)\right)$ that an agent has relative to a threshold $\tau \in (0,1] \subset \mathbb{R}_+$, which is itself taken to be a function of the salience of evidence which becomes accessible:[201]

$$\textbf{(8)} \qquad \tau\left(\widehat{\boldsymbol{d}}_{i,t}^{A_f}, \widehat{\boldsymbol{d}}_{i,t}^{\neg A_f}\right) = 1 - \frac{\left|\widehat{\boldsymbol{d}}_{i,t}^{A_f} - \widehat{\boldsymbol{d}}_{i,t}^{\neg A_f}\right|}{\widehat{\boldsymbol{d}}_{i,t}^{A_f} + \widehat{\boldsymbol{d}}_{i,t}^{\neg A_f}}$$

One can interpret this threshold specification as capturing the effect that an agent will tend to be more confident as the asymmetry of the balance of her evidence increases (for lower values of $\delta_i$), but face an erosion of confidence as symmetry between evidence for and against $A_f$ increases (this specification captures Griffin and Tversky's (1992) findings that agents tend to grant primary consideration to the balance of evidence, while giving only secondary consideration to its weight).

In total, the index transmission function (Eq.5) assigns the intermediate evidence ratio $\sigma_{i,t}(A_f, \neg A_f)$ to a decision interval, and selects a value from within that decision interval which reflects an agent's 'quantized' belief for each focal hypothesis $A_f$ based on her accessible evidence.

Bear in mind, however, that each agent here considers $F$ such hypotheses at once and that there are $I$ agents. Blending these two considerations in the above framework permits the EBSL model.

### **6.3.2**\\ Extension to Dynamics and Heterogeneous Agents

The value of $F$ may be taken as equaling the dimensionality of the space (*doxastic space*) inside which an agent arrays not only her own beliefs, but also her beliefs about her peers' beliefs. Take each dimension in an agent $i$'s doxastic space as corresponding to a focal hypothesis $A_f$ which she

[201] This functional specification is that suggested by salience theory due to Bordalo, Gennaioli, and Shleifer (2012).

is considering. Each one of these dimensions represents a partitioning of the probability interval $C = [0,1] \subset \mathbb{R}$ according to that agent's value of $\theta_i$ (for simplicity, assume that $\forall i \in I$, $\theta_i$ is the same across all $F$ hypotheses). As $\theta_i$ partitions each dimension into an integer-valued number of decision intervals, there are $\theta_i$ 'cells' along any dimension in an agent's doxastic space that her belief on that dimension might occupy. With $F$ dimensions in any doxastic space, there are $\theta_i^F = N$ such cells that might be occupied overall within the agent $i$'s doxastic space. Let $n \leq N$ index these cells. (Abusing notation, hereafter allow $\theta_i^F$ to denote $i$'s doxastic space, based on her value for $\theta_i$.) And let $n_{i,t}$ be the location of the cell which contains agent $i$'s beliefs in all $F$ dimensions during period $t$. This setup allows extension of Eq.1 from Section 6.2 to multiple dimensions via:

**(9)** $$H_{\theta_i^F,t} = -\sum_{n=1}^{N} p_{n,t}\left(\log_2\left(p_{n,t}\right)\right)$$

Here again $p_{n,t}$ is the proportion of the total $I$ agents whose beliefs across all $F$ hypotheses occupy cell $n$ at period $t$ in the doxastic space at the level of $\theta_i$ (i.e., from agent $i$'s perspective).[202] Note again that this entropy value is minimized if all agents' beliefs collocate in one cell. That is, $H_{\theta_i^F,t}$ is minimized at $t$ when $p_{m,t} = 1$ for some unique position $m \leq N$ with $p_{n_t} = 0$, $\forall m \neq n$ and $n \leq N$. This minimum-entropy (*total doxastic agglomeration*, or 'TDA') configuration wherein entropy is zero ($H_{\theta_i^F,t} = 0$) thus finds agent $i$ perceiving all agents to a share consensus in beliefs.

Conversely, the maximum-possible entropy at any time period $t$ (denoted by $\dot{H}_{\theta_i^F}$, with no time subscript due to time-invariance) is equal to $\log_2(N)$, which corresponds to maximum belief diversity. Within a given doxastic space, maximum entropy can be described as equal to minimum

[202] Alternatively (but equivalently), one could likewise think of each agent's beliefs as a vector combination within an $F$-dimensional doxastic space, with $p_{n,t}$ representing the proportion of such vectors that are co-terminal in a cell within that space, and realizing that the value of $\theta_i$ usually restricts the overall number of cells within that space to be finite.

*syntropy* (negative entropy), and maximum syntropy as equal to minimum entropy. The degrees of syntropy $\Gamma_{\theta_i^F,t}$ and entropy $H_{\theta_i^F,t}$ within a doxastic space $\theta_i^F$ at period $t$ are thereby related via:

**(10)** $\dot{H}_{\theta_i^F} \equiv H_{\theta_i^F,t} + \Gamma_{\theta_i^F,t}$

Both $H_{\theta_i^F,t}$ and $\Gamma_{\theta_i^F,t}$ are thus bounded from below by zero and from above by $\dot{H}_{\theta_i^F} = \log_2(N)$.

Syntropy can also be further additively decomposed into two sub-measures, *doxastic agglomeration* $\xi_{\theta_i^F,t}$, and *juncture* $\eta_{\theta_i^F,t}$, as follows:

**(11)** $\Gamma_{\theta_i^F,t} \equiv \xi_{\theta_i^F,t} + \eta_{\theta_i^F,t}$

To help clarify the definitions for $\xi_{\theta_i^F,t}$ and $\eta_{\theta_i^F,t}$, rewrite $\Gamma_{\theta_i^F,t}$ (based on Eq.10) according to:

**(12)** $\Gamma_{\theta_i^F,t} = \log_2(\mathrm{N}) + \sum_{n=1}^{N} p_{n,t}\left(\log_2\left(p_{n,t}\right)\right)$

Doxastic agglomeration can then be defined as the 'marginal sums' in each of the dimensions:

**(13)** $\xi_{\theta_i^F,t} \equiv \log_2(\mathrm{N}) + \sum_{f=1}^{F} \sum_{v=1}^{\theta_i} p_{v,t}\left(\log_2\left(p_{v,t}\right)\right)$

In Eq.13, $p_{v,t}$ denotes the proportion of agents in subinterval $c_v$ along dimension $f$ at time $t$. Juncture can thereby be defined residually as the difference between values from Eq.12 and Eq.13:

**(14)** $\eta_{\theta_i^F,t} \equiv \Gamma_{\theta_i^F,t} - \xi_{\theta_i^F,t}$

The EBSL model introduced by Rook (2014b) treats the values of $H_{\theta_i^F,t}$, $\xi_{\theta_i^F,t}$, and $\eta_{\theta_i^F,t}$ as vital for describing the dynamics of social learning when agents face conditions of uncertainty; it also makes them homogenous across agents. These three values all summarize an agent's perceptions of the belief distributions among all agents in a group (which includes herself); we may suspect

that they can be heterogeneous for agents of differing cognitive abilities (especially biases) distinct from $\theta_i$ and $\delta_i$. Section 6.3.4 treats one possibility for capturing such biases through these values.

Under EBSL, during time period $t$ each agent $i$ forms a belief about each hypothesis $A_{f \leq F}$ that may be described as her *own* belief $\mathrm{Q}\left(\sigma_{i_o,t}\left(A_f, \neg A_f\right)\right)$. Each agent also forms such a belief about what she thinks the *average* belief $\mathrm{Q}\left(\sigma_{i_{\bar{\imath}},t}\left(A_f, \neg A_f\right)\right)$ is among all (relevant) agents (including her own) at $t$.[203] Assume at $t = 0$ that all $i$ agents hold and form such 'own' and 'other-regarding' (hereafter *primary* and *secondary*, respectively) beliefs which obey the quantization process described in Section 6.3.1. *A priori*, the specific relationship between $\mathrm{Q}\left(\sigma_{i_o,t}\left(A_f, \neg A_f\right)\right)$ and $\mathrm{Q}\left(\sigma_{i_{\bar{\imath}},t}\left(A_f, \neg A_f\right)\right)$ at $t = 0$ is kept as unrestricted (they may be equal, but likewise they need not).

Furthermore, retain the *limited visibility* assumption from Section 6.2: each agent has direct visibility on only her own beliefs. This situation means that agents' secondary beliefs might be accurate (although alternatively they might be inaccurate); but, absent other conditions, agents are not perfectly sure about the accuracy of their beliefs about other agents' beliefs (they are, however, sure about their primary beliefs, which aligns with findings from psychology that people oftentimes use their own beliefs as initial reference points when forecasting the beliefs of others).

At $t = 1$ all agents receive an exogenous public signal $g \in \mathbb{Z}$ (such a signal could be made real-valued and/or continuous, but, as agents are presumed to discretize evidence anyway through

[203] Inclusion of her own beliefs in the average belief comes to have a particularly important bearing when there are few agents, and thus an agent may come to need to consider how her own beliefs might impact a larger, overall belief system. The fact that each agent also quantizes her beliefs about the average of others' beliefs roughly parallels the analogy-based treatment of agents' expectations of others in Jehiel and Koessler (2008). Furthermore, although the implications are not explored in this paper, the results of experiments by Duffy and Nagel (1997) suggest that the outcomes of EBSL processes may be sensitive to whether agents actually use the average of others' beliefs, or some other statistic (e.g., the median or maximum). It is plausible that such statistics may vary according to problem context.

quantization, the restriction to integer-valued signals is made for convenience). If $g > 0$ it supports hypothesis $A_f$; but if $g < 0$ it is then evidence instead for $\neg A_f$ (and let $g = 0$ count as 'non-evidence'). Receipt of $g$ at $t = 1$ causes all agents to update their intermediate evidence ratios via:

**(15)** $$\sigma_{i,t^1}(A_f, \neg A_f)|g = \frac{\widehat{\boldsymbol{d}}_{i,t}^{A_f} + (zg)}{\widehat{\boldsymbol{d}}_{i,t}^{A_f} + \widehat{\boldsymbol{d}}_{i,t}^{\neg A_f} + |g|}, \qquad \text{where: } z = 0 \text{ if } g < 1$$

$$z = 1 \text{ if } g \geq 1$$

Each agent is assumed to engage in this updating operation for both her primary and secondary beliefs, such that after receiving the signal $g$, her primary and secondary beliefs become (respectively) $\mathrm{Q}(\sigma_{i_o,t^1}(A_f, \neg A_f)|g)$ and $Q\left(\sigma_{i_{\bar{\imath}},t^1}(A_f, \neg A_f)|g\right)$ for all of the hypotheses $A_{f \leq F}$.[204] The assumption which propels updating of $Q\left(\sigma_{i_{\bar{\imath}},t^1}(A_f, \neg A_f)|g\right)$ is driven by the signal's publicness; assume here that all agents believe that all agents are aware of and believe the signal $g$ (as in Condition K2 from Section 6.2). Due to the limited visibility assumption, each agent may be unsure about whether all evidence that is accessible by at least some agents has become public. That is, an agent may believe that others hold information that she does not, even though she may question the accuracy or relevance of that information; alternatively, she may feel that she possesses particular information that some others lack, although she may be unsure of its validity.

EBSL therefore proposes that, in responding to the public signal $g$ at $t = 1$, each agent engages in *anticipatory updating* of her primary and secondary beliefs by forming some convex combination of both. This convex combination creates what may be called the agent's *amalgamated posterior*:

**(16)** $$\pi_{i,t^1}(A, \neg A_f)|g = \left(\frac{1}{H_{\theta_i^F,t^0} + \xi_{\theta_i^F,t^0}}\right)\left[\left(\mathrm{Q}(\sigma_{i_o,t^1}(A_f, \neg A_f)|g)\right) H_{\theta_i^F,t^0} + \left(Q\left(\sigma_{i_{\bar{\imath}},t^1}(A_f, \neg A_f)|g\right)\right) \xi_{\theta_i^F,t^0}\right]$$

[204] Understand that this updating is simultaneously (i.e., within the timeframe of a discrete time period) undertaken by all agents. Halpern and Moses (1990) claim that simultaneity is generally necessary for common knowledge to arise.

Realize that the agglomeration and entropy values in Eq.16 are functions of the distribution of primary (but not secondary) beliefs (from the perspective of agent $i$ in terms of her doxastic space) for all of the $I$ agents, and are lagged from $t = 0$: they pertain to the distribution of agents' beliefs before the signal's arrival (and before agents have had a chance to detect any changes to the consensus among their peers' beliefs). Also observe that, as functions of the belief distributions themselves, agglomeration and entropy (and juncture) change endogenously as agents' beliefs change. That is, these parameters 'self-tune' as agents learn and respond to changes in the overall distribution of beliefs. Agglomeration and entropy (made relative to one another) provide dynamic weights that indicate how much trust each agent allocates to her primary and secondary beliefs respectively. When entropy is maximal (equal to $\dot{H}_{\theta_i^F}$) an agent trusts only her primary beliefs. But when entropy is minimal (and agglomeration is maximal), then an agent trusts only her secondary beliefs (however, for the latter case her primary and secondary beliefs must anyway be equivalent).

Note, however, that the signal $g$ at $t = 1$ may not pertain to all $F$ hypotheses (i.e., it may equal zero for one or more $A_f$). If $F > 1$, and two or more hypotheses are not mutually-independent in agents' minds, then such agents may be conservative in updating those hypotheses to which $g$ does pertain (i.e., is nonzero). Juncture $\eta_{\theta_i^F,t}$ is a measure of the degree of hypothesis-interdependency suggested by agents' primary beliefs. Let $\Delta_g$ describe the dimensionality of the signal (i.e., the number of hypotheses $A_f$ for which $g$ is nonzero when it arrives at $t = 1$) with $F \geq \Delta_g \in \mathbb{N}$. Since $\Delta_g$ is a function of the public signal $g$, $\Delta_g$ is assumed to be the same for all agents. $\Delta_g$ measures how severely $\eta_{\theta_i^F,t}$ impedes updating of each agents' primary and secondary beliefs (and by extension, the amalgamated beliefs of each) from one period to the next. This impediment is best understood as the degree to which agents under-adjust beliefs in attempting to maintain coherence.

$\Delta_g$ and $F$ are related to one another and to $\eta_{\theta_i^F,t}$ through a linear penalty function, defined by:[205]

**(17)** $\Psi \equiv \frac{F-\Delta_g}{F-1}$

The penalty function $\Psi$ influences Eq. 16 and amalgamation of updated beliefs according to:

**(18)** $\tilde{\pi}_{i,t^1}(A,\neg A_f)|g = Q\left[\mathrm{Q}\left(\sigma_{i_o,t}(A_f,\neg A_f)\right) + \left[1-\left[\Psi\left(\frac{\eta_{\theta_i^F,t^0}}{\log_2(N)}\right)\right]\right]\left[\pi_{i,t^1}(A,\neg A_f)|g - \mathrm{Q}\left(\sigma_{i_o,t}(A_f,\neg A_f)\right)\right]\right]$

The belief-value emanating from Eq.18 is the quantized output of its right-hand-side components (denoted through bracketing the right-hand-side expression with a quantization function: $Q[\cdot]$). Notice that, whenever $\Delta_g = F$, juncture has no impact whatsoever upon agents' belief updating. And, importantly, the output of Eq.18 becomes agent $i$'s primary belief at the end of period $t = 1$.

As a result of the signal's arrival and anticipatory updating of agents' beliefs according to Eq.16 and Eq.18, the distribution of agents' beliefs may change from $t = 0$ to $t = 1$. EBSL assumes that when any agent changes her (quantized) belief, then her posterior contributes to an aggregate statistic on belief updating: the (publicly) visible average change (as a response to the signal $g$):[206]

**(19)** $E_{I,t^1}(A_f,\neg A_f)|g = \frac{\sum_{i=1}^{I}\left[\tilde{\pi}_{i,t^1}(A,\neg A_f)|g\right]\mathrm{I}\left(\left(\tilde{\pi}_{i,t^1}(A,\neg A_f)|g-\mathrm{Q}\left(\sigma_{i_o,t}(A_f,\neg A_f)\right)\right)\neq 0\right)}{\sum_{i=1}^{I} 1_i}$

where: $1_i = 1$ if $\left(\tilde{\pi}_{i,t^1}(A,\neg A_f)|g - \mathrm{Q}\left(\sigma_{i_o,t}(A_f,\neg A_f)\right)\right) \neq 0$

$1_i = 0$ if $\left(\tilde{\pi}_{i,t^1}(A,\neg A_f)|g - \mathrm{Q}\left(\sigma_{i_o,t}(A_f,\neg A_f)\right)\right) = 0$

[205] Differing specifications for this penalty function for various agents might feasibly give rise to forms of cognitive heterogeneity among agents. Nonetheless, Rook (2014b) finds that juncture does not often play a major role in determining belief dynamics among agents engaged in EBSL (or at least it does not play as significant a role as the agglomeration and entropy parameters when $\Delta_g$ is not significantly less than $F$). For these reasons, it is assumed throughout that this penalty function takes this same simple form described by Eq.17 for all of the agents.

[206] Rather than observing changes in average (quantized) belief among all agents, one might envision that the average of all agents' "decisions" (under Ménager's (2008) very general intention for the term) becomes visible.

Realize that $E_{I,t^1}(A_f, \neg A_f)|g$ is an un-quantized value that is observed publicly by all, and not mentally-generated.[207] This new average belief gets reintroduced into Eq.16 to produce the value:

**(20)** $$\pi_{i,t^2}(A, \neg A_f)|g = \left(\frac{1}{H_{\theta_i^F,t^1} + \xi_{\theta_i^F,t^1}}\right)\left[(\tilde{\pi}_{i,t^1}(A, \neg A_f)|g)H_{\theta_i^F,t^1} + (E_{I,t^1}(A_f, \neg A_f)|g)\xi_{\theta_i^F,t^1}\right]$$

That is, during period $t = 2$, Eq.16 is reapplied with the value from Eq.18 as agent $i$'s primary belief, the value from Eq.19 as the secondary belief for every agent, and the lagged values of agglomeration and entropy (i.e., their values for the distribution of beliefs at the end of $t = 1$). Yet, juncture may have changed as well between $t = 0$ and $t = 1$. To account for the effects of this (potentially) new juncture value, the output of Eq.20 enters Eq.18 as an input via the following (the value $\Psi = \frac{F - \Delta_g}{F-1}$ will remain unchanged from Eq.17, as no new signals are assumed to arrive):

**(21)** $$\tilde{\pi}_{i,t^2}(A, \neg A_f)|g = Q\left[\pi_{i,t^1}(A, \neg A_f)|g + \left[1 - \left[\Psi\left(\frac{\eta_{\theta_i^F,t^1}}{\log_2(N)}\right)\right]\right][\pi_{i,t^2}(A, \neg A_f)|g - \pi_{i,t^1}(A, \neg A_f)|g]\right]$$

Application of Eq.21 might well result in new quantized values of beliefs at the conclusion of period $t = 2$, which would then cause the values of agglomeration, entropy, and juncture to become different from those in $t = 1$. These changes would then drive the beginning of the next iteration of this tâtonnement process with a new value for Eq.19, and subsequent application of equations Eq.20 and Eq.21. The process eventually stops (in finite time) when no further updating occurs. Conclusion of this process can be considered a form of equilibrium, as per Jordan (1982).[208]

[207] It may seem arguable that one could generate cognitive diversity among agents through Eq.19 if some agents were to use other statistics (e.g., median or mode) rather than the average. Use of such different statistics, however, would seem to reflect more differences in thinking process than inherent differences in cognitive ability (although the two may well be linked). For this reason the choice is made to keep Eq.19 homogenous across agents for now.

[208] When understood in this way, EBSL seems to function as something akin to an 'equilibrium selection criterion'.

### **6.3.3**\\ Message Length and Collapsing the Model

Some fundamental results from information and coding theories further suggest why situations of minimum entropy (i.e., total agglomeration) among agents' beliefs could function as belief equilibria: they are the least information-intensive configurations for agents' beliefs. That is, if one were to devise a code conveying the position of every agent in some doxastic space (for simplicity assuming that all agents share an identical $\theta_I^F$) to all of the $I$ agents, then a situation of $\xi_{\theta_I,t} = \log_2(N)$ (i.e., total agglomeration in agents' beliefs at the level of $\theta_I$) would require the shortest message length for any sensible encoding (on introducing coding ideas into economics, see Sims (2003; 2010) on rational inattention theory; on applying these ideas specifically to the concept of communicating consensuses see Hanson (1998)). That is, the message to transmit the location of every agent would require the fewest 'bits' when $\xi_{\theta_I,t} = \log_2(N)$. If one assumes that longer messages carry either (or both) a higher likelihood of corruption from noise, or higher costs to receive (and decode), then situations of total agglomeration (minimal entropy) will generally be the least costly or most noise-free messages that can be either received or communicated by agents.

Furthermore, if, instead of announced messages, agents have to try to infer belief distributions, then we may expect inferences about total-agglomeration distributions to be less noisy or costly than other distributions of agents' beliefs. This interpretation of $\xi_{\theta_I,t} = \log_2(N)$ adds to the list of forces suggested by Rook (2014b) for why total-agglomeration configurations may encourage agents to join in a perceived consensus, as well as dissuade them from departing from it when they are aware of it: it represents the most credible and potentially least-costly informative arrangement.

It is perhaps also worth noting that the EBSL setup collapses into conventional Bayesian updating of quantized beliefs under either: total agglomeration (i.e., minimum entropy); or a representative

agent framework (which will necessarily entail total agglomeration). Furthermore, with infinite $\theta_i$ for all agents across all hypotheses, the above setup reduces to standard Bayesian updating. Nevertheless, as earlier mentioned, the purpose here is to study diverse agents with distinct cognitive limitations, and explore how such heterogeneity may impact their abilities to arrive at consensuses. The reduction of the above model to representative agency will therefore not be useful here, and the remainder of this paper explores several potential paths toward cognitive heterogeneity in greater detail. But first a brief aside is needed on how the foregoing setup may be leveraged to 'fracture' the infinite regress problem that common knowledge can otherwise entail.

### **6.4**\\ Total Doxastic Agglomeration as Commonality

Economic treatments of common knowledge oftentimes accentuate its infinite and hierarchical nature. Nonetheless, surprisingly little advantage appears to be taken in those treatments of the inherently *circular* nature of common knowledge. To illustrate, consider two entities (which may, but need not, be individuals; they need not even be distinct from one another). Call one entity $x$ and the other $y$. Knowledge of some event $z$ is said to be common to $x$ and $y$ if: $x$ knows $z$, and $x$ knows that $y$ knows $z$, and $y$ knows that $x$ knows that $y$ knows $z$, and $x$ knows that $y$ knows that $x$ knows that $y$ knows $z$…and so on. The "and so on" component of this construction indicates its infinite nature by virtue of intentional equivalency to a more exact expression: *ad infinitum*. Common knowledge's hierarchical nature stems from lengthening concatenations of "knows that".

Yet, while common knowledge is hierarchically infinite, it is also circular. Notice that in the (infinitely) lengthening concatenation of statements that the clause at the end ("knowledge of $z$", which serves as a direct object) never changes from one level of the hierarchy to the next, and additionally that there is an alternation in subjects ($x$ or $y$) between one level of such a hierarchy

and levels above and below it (apart from the very first level). Thus, rather than one hierarchy, there actually are two: the knowledge of $x$ (which includes knowledge of $y$'s knowledge); and the knowledge of $y$ (which includes knowledge of $x$'s knowledge). Common knowledge requires that both of these hierarchies extend infinitely, and thereby requires an infinite number of steps to fully verify, unless circularity between the two hierarchies can be somehow fractured by dissociation.[209]

### **6.4.1**\\ Fracturing Circularity

When augmented by some simple conditions, doxastic agglomeration permits such fracturing. Continuing from the above setup, let $x$ be some agent $i \in I$, and let $y$ be the 'average agent' in $I$. Let each agent $i$ form beliefs on what $x$ and $y$ believe about some $F$ hypotheses $A_{f \leq F}$ (and all agents consider $F$ identical hypotheses), as in Subsection 6.3.2. For convenience, let $\pi_i(x)$ and $\pi_i(y)$ respectively be $i$'s beliefs on what $x$ and $y$ believe about $z$. Notice that $\pi_i(y)$ is at least partially self-referential for every $i$, as the average across all of the $I$ agents' beliefs must include each agent $i$'s beliefs (although the extent to which $i$'s beliefs influence this average necessarily diminishes, in general, as the number of agents $I$ increases). Moreover, let (for now) all of the $i \in I$ agents share an identical $\theta_I$, and that, at time $t$, the measure of doxastic agglomeration $\xi_{\theta_I,t}$ across the primary beliefs $\pi_i(x)$ of all agents becomes publicly accessible (the mechanisms for such accessibility are immaterial for now; $\xi_{\theta_I,t}$ may be announced, or else all agents may infer its value at $t$). Further, substitute Condition K2* below into the pair of conditions introduced in Section 6.2:

**Condition K1**: All agents know that all agents know their own beliefs.

**Condition K2***: All agents know that $\xi_{\theta_I,t}$ is known to all agents.

[209] Nor does the alternative approach entailing embedding such infinite 'back-and-forth' iterations of knowledge in the state space (here: hypotheses themselves) reduce the infinite verifications needed to establish common knowledge.

If both Condition K1 and K2* (along with the 'homogeneity' assumptions of shared $\theta_I$ and identical hypotheses for all agents) hold, then common knowledge of one another's beliefs exists *at the level of* $\theta_I$ among the $I$ agents whenever total doxastic agglomeration occurs (i.e., when $H_{\theta_{i,t}} = 0$, or, equivalently: $\log_2(N) = \xi_{\theta_I,t}$). That is, $\pi_i(x) = \pi_i(y), \forall i \in I$ will be true, as every agent $i$ adopts the average belief among all agents, and all agents are in consensus in their beliefs. The average belief held by all of the agents thus becomes the only belief held by any of the agents. Common knowledge of, and consensus in, beliefs at the level of $\theta_I$ across all agents then prevails.

And it is plain to see that this construction of common knowledge does not require infinite verifications to establish, so long as $I$ is finite. In fact, it only requires $2I^2 + 1$ verifications ($I^2$ verifications for each of Condition K1 and K2*, plus another verification that $\log_2(N) = \xi_{\theta_I,t}$). This reduction is permitted by 'fracturing' the circularity that exists between $x$ and $y$ by splitting them apart and then utilizing a statistic ($\xi_{\theta_I,t}$) able to signal if they perfectly coincide for all agents.

The situation of perfect coincidence of all beliefs when $\log_2(N) = \xi_{\theta_I,t}$ can thus be understood as a form of belief equilibrium. Such an equilibrium is comparable to that which is described by Pearlman and Sargent (2005), wherein agents do not face any problem in forecasting the forecasts of other agents in a belief equilibrium, because all agents' forecasts are then equal to one another.

These results reinforce the pervasive findings on the impossibility of agreeing to disagree when common knowledge of posteriors exists among agents. Yet the doxastic agglomeration setup here relaxes one of the restrictions that has pervaded earlier results: the requirement of common priors. Under the present arrangement, common knowledge of beliefs is only guaranteed when all agents' beliefs agree at the level of $\theta_I$, and therefore common knowledge of posterior beliefs necessitates homogeneity of such posteriors, irrespective of whether or not agents' priors are initially all equal.

This extension may be important in the respect that it seems to accord better with observable reality: agents' beliefs oftentimes appear diverse (sometimes greatly so). Yet, given that agents' beliefs under EBSL (and QET) are founded on accessible evidence, it does not by itself require jettisoning the perspective that all differences in beliefs are traceable ultimately to asymmetries in information. Still, results from later in this paper suggest some merit might exist in evolving that perspective in order to allow for some differences in beliefs to stem from cognitive heterogeneities.

### **6.4.2**\\ Weakening to Common Belief

Importantly, this approach to common knowledge can be weakened to one of common beliefs by substituting "believe" for "know" in Conditions K1 and K2* to produce the following conditions:

**Condition B1**: All agents believe that all agents believe their own beliefs.

**Condition B2**: All agents believe that all agents are aware of and accept $\xi_{\theta_I,t}$.

When these two belief conditions hold alongside the homogeneity requirements of a shared $\theta_I$ and identical hypotheses for all agents, then common belief and consensus among the beliefs across all the $I$ agents could be said to exist whenever $\log_2(N) = \xi_{\theta_I,t}$ holds during any time period $t$.[210]

From the foregoing it should be apparent that the EBSL framework can comfortably accommodate concepts of common belief and common knowledge among agents. Nevertheless, for any large (finite) population of real-world agents, not only does the number of steps $2I^2 + 1$ in the verification process become forbiddingly large (even if finite), but also the necessary conditions and homogeneity requirements for common knowledge or belief would seemingly become harder

[210] One might wish to be less abstract and specify "p-beliefs" rather than generic 'believing' (on the distinction, see Monderer and Samet (1989) or Neeman (1996a), for example), but that possible extension is not pursued here.

to fulfill.[211] Particularly, cognitive homogeneity may be less tolerable as an assumption for sizable populations of agents, and thus larger values of $I$ (even if the value of assuming cognitive homogeneity is increasing with larger values of $I$ for the added tractability it buys). Yet it is in dealing with and exploring impacts of such cognitive diversity that the EBSL model is well-suited.

## **6.5**\\ Diverse Dimensions of Cognitive Heterogeneity

As was stated in the introduction, situations wherein agents disagree about whether or not a consensus exists (or the degree of diversity in belief even when none perceive a consensus) may stem significantly from diversity among agents' cognitive abilities. Subsections 6.2.1 and 6.2.2 highlighted several routes to exploring the effects of cognitive diversity on inter-agent disagreement within the EBSL setup. In particular, using heterogeneous values for the parameters $\theta_i$ and $\delta_i$ across agents was identified as a potentially fruitful path for such study. Recognize that these parameters pertain directly to the structure of doxastic spaces: $\theta_i$ dictates how many discrete 'cells' exist within a doxastic space from the perspective of agent $i$, whereas $\delta_i$ determines the exact location of an agent's beliefs within such cells. In terms of cognitive diversity, $\delta_i$ is therefore subordinate to $\theta_i$ in driving disagreement among agents' perceptions of consensus: agents cannot disagree about consensus if they have homogenous discrimination parameters (i.e., an identically-shared $\theta_I$) and all perceive identical values for $\xi_{\theta_i,t}$, even if they differ in their respective values for $\delta_i$. This result stems from Theorem 1, as identical discrimination parameters between agents violates the first of the five conditions in Theorem 1 (that is, subintervals cannot perfectly overlap).

[211] Arguably, the number of necessary verification steps in moving from common knowledge to common belief might be reduced to $2I + 1$, as Condition B1 and Condition B2 above might be accepted as verified even if the "believe that all agents" component of each condition is removed. This reduction might be acceptable in some circumstances, but less palatable in others. One should regardless recognize that the number of verification steps increases *at least* linearly in the number of agents, which, for large values of *I* may be problematic for agents facing severe cognitive restrictions.

Yet at least two further cognitive sources of interest for potential disagreements about consensus exist and are dealt with by this paper. First, even if all agents share an identical value for $\theta_i$, some agents might still exhibit systematic biases in their perceptions of $\xi_{\theta_i,t}$ (and for that matter $H_{\theta_i,t}$). Second, agents may differ in the degrees to which their cognitive abilities allow them to treat probabilities linearly. Those agents with high non-linearity in their perception of probability may experience strong non-uniformity in the lengths of subintervals in their doxastic spaces. The remainder of this section discusses (in turn) each of these possibilities in detail as relates to EBSL.

### **6.5.1**\\ Systematic Consensus Misperception

The prevalence of empirical findings in psychological research of distorted perceptions of consensus was discussed in the introductory section under the term *false-consensus effects* (FCEs). As was noted there, people sometimes systematically misperceive the degree of consensus among some group (or subgroup) of which they are a part. As a first approximation, then, it should be reasonable to embed these variable misperceptions as fixed scalars in the EBSL model by attaching them to the relative weights on an agents' primary and secondary beliefs. That is, one might consider some scalar parameter $\varphi_i \in (0,1) \subset \mathbb{R}_+$ which may vary across agents. Recall that in EBSL the ratio $\xi_{\theta_i,t} \,/\, H_{\theta_i,t}$ crucially determines how much relative weight each agent $i$ places on her primary and secondary beliefs (where such weights necessarily sum to unity). The original exposition of EBSL regarded this ratio as accurately perceived by each agent (at least at the level of $\theta_i$). One could, however, allow $\varphi_i$ to interfere with this perception and systematically bias perceptions of consensus according to:

**(22.1)** $$\frac{(1-\varphi_i)^{-1} H_{\theta_i,t}}{\varphi_i^{-1}\xi_{\theta_i,t}+(1-\varphi_i)^{-1} H_{\theta_i,t}}$$

**(22.2)** $\frac{\varphi_i^{-1}\xi_{\theta_i,t}}{\varphi_i^{-1}\xi_{\theta_i,t}+(1-\varphi_i)^{-1}H_{\theta_i,t}}$

The above weights replace the respective weights, $\frac{H_{\theta_i,t}}{\xi_{\theta_i,t}+H_{\theta_i,t}}$ and $\frac{\xi_{\theta_i,t}}{\xi_{\theta_i,t}+H_{\theta_i,t}}$, on an agent $i$'s primary and secondary beliefs in Eq.16 and Eq.20. With the substituted weights from (22.1) and (22.2), a value for $\varphi_i$ nearing zero means that an agent would increasingly discount the degree of consensus (agglomeration), and focus more on her primary beliefs. Likewise, with $\varphi_i$ nearing one, an agent would tend to inflate the degree of consensus and focus increasingly on her secondary beliefs. A value of $\varphi_i = 0.5$ suffices to cancel out effects from $\varphi_i$. But realize that this fixed-scalar approach by itself is limited in a key way: *it alone cannot cause disagreement about whether consensus exists*. Nor on its own is it adequate to force agreement between agents that would otherwise not agree about consensus. That is, this fixed-scalar approach influences the degree of strength agents perceive in an imperfect consensus (i.e., how much agreement agents perceive among one another's beliefs when at least some heterogeneity exists); it does not influence whether agents perceive a perfect consensus in beliefs to exist or not (see Appendix 6.B for the associated proofs).

The fixed-scalar approach therefore serves as a sort of contingent modification to (asymmetric) consensus misperception: conditional on no perfect consensus being perceived (so long as $\varphi_i \in (0,1)$ for all $i \in I$; i.e., $\varphi_i$ does not equal 0 or 1) by an agent, this fixed-scalar setup may strengthen or weaken the degree of imperfect agreement in beliefs that she detects among all agents' beliefs. This fixed-scalar setup may affect the dynamics toward or away from consensus among agents' beliefs when beliefs are updated; yet it can neither give rise to nor destroy consensuses on its own.

In the ensuing analysis, $\varphi_i$ will be taken as exogenous and time-invariant, but will be allowed in some cases to vary across agents. A more sophisticated future treatment, however, may take $\varphi_i$ to

be a function of both $\theta_i$ and $\xi_{\theta_i,t}$ and thereby potentially time-varying. The argument for doing so resides in the earlier observation on message lengths and noise-robust transmissions of the parameter $\xi_{\theta_i,t}$. If an agent's perception of $\xi_{\theta_i,t}$ is noisy, then she should trust it less the lower it is, but place greater faith in its accuracy the higher it is. Moreover, the level of discrimination $\theta_i$ at which she is operating may further influence the accuracy of her perception. Hence, it may be fruitful in future work to establish some optimal level of misperception due to $\varphi_i(\xi_{\theta_i,t}, \theta_i)$ that varies with $\xi_{\theta_i,t}$ and $\theta_i$ to serve as a more refined benchmark against which to establish to what extent systematic misperception is disadvantageous (or advantageous) to an agent. That possibility is set aside for later work. For now, recognize that $\varphi_i$, unlike $\theta_i$ and $\delta_i$ is not a parameter involved in the geometry of any doxastic space, but only influences the perceived distribution of agents' beliefs in such spaces. But next consider another force directly related to doxastic-space structure.

### **6.5.2**\\ Nonlinear Probability Scales

Nonlinear treatment of probabilities is perhaps one of the most pervasive findings in behavioral economics, and a core tenet of two of the benchmark theories that describe people's departures from what is dictated by expected utility theory: prospect theory (Kahneman and Tversky 1979) under risk, and cumulative prospect theory (Tversky and Kahneman 1992) under uncertainty. A key feature in both flavors of prospect theory is that agents distort their true beliefs by using "decisions weights". Under situations of risk, decision weights are often taken to be purely a function of preferences (see Barberis (2013a)), but under uncertainty they may be influenced as well by misperceptions of beliefs. Wakker (2004) particularly argues that cognitive interference beyond distortions due to preferences may cause decision weights to alter subjective probability judgments when probabilities are not objectively given (i.e., under uncertainty, per Knight (1921)).

Given the pervasiveness of probability weighting, perhaps the most objectionable aspect of EBSL is its linear treatment of probability scales in assuming $\theta_i$ generates a uniform partition of the probability scale, such that each dimension in a doxastic space entails $\theta_i$ disjoint subintervals (assuming that $\theta_i$ is identical across all $F$ hypotheses considered by agent $i$), each of equivalent length $\theta_i^{-1}$. Such uniformity, while simplifying, is unsatisfying for being contradicted by so much empirical evidence. One remedy is to allow the subintervals generated by $\theta_i$ to instead be compressed or stretched relative to the uniform 'mesh size' dictated by $\theta_i^{-1}$. A simple way to achieve such distortions is described by Rook (2014a), and involves a transformation of each subinterval according to $\theta_i$ in order to capture the "diminishing sensitivity" that is argued by Kahneman and Tversky (1979) to partly drive nonlinear probabilities. More specifically, the diminishing-sensitivity parameter $\gamma$ (appearing in both versions of prospect theory) relates to $\theta_i$ in as follows (owing to to its dependence upon $\theta_i$, let $\gamma_i$ hereafter carry an agent-specific subscript):

(**23**) $\gamma_i \equiv 1 - \zeta_i \theta_i^{-1}$

In Eq.23, $\zeta_i \in [0,1] \subset \mathbb{R}$ serves as a scalar influencing the direct correspondence between $\theta_i$ and $\gamma_i$ by potentially dampening it. For the sake of focusing discussion, this paper will only be concerned with the extreme values of $\zeta_i$, such that $\zeta_i \in \{0,1\}$ here. Restricting $\zeta_i$ in this way allows us to easily alternate between: agents that have finite $\theta_i$ but treat probabilities linearly (when $\zeta_i = 0$ and thereby $\gamma_i$ is forced to equal 1); and those who treat probabilities nonlinearly (with $\gamma_i \neq 1$).

For values of $\theta_i$ between 2 and 5, $\gamma_i$ as defined by Eq.23 takes on values that have been frequently observed in empirical investigations of (both cumulative and non-cumulative) prospect theory (Rook 2014a). Let $\theta_i^{-1}$ and $\gamma_i$ then define an endpoint transformation function:

(**24**) $$\lambda(b\theta_i^{-1},\gamma_i) \equiv \frac{(b\theta_i^{-1})^{\gamma_i}}{\left[(b\theta_i^{-1})^{\gamma_i}+(1-b\theta_i^{-1})^{\gamma_i}\right]^{1/\gamma_i}}$$

Notice that, given Eq.24, there is only one independent parameter in Eq.24 (i.e., $\theta_i^{-1}$), but that its influence on the nonlinearity of probability can be switched 'on' and 'off' through the impact of $\zeta_i$ in decoupling $\gamma_i$ from $\theta_i$ (as described above in Eq.23). Also recognize that the functional form for Eq.24 is that from cumulative prospect theory (Tversky and Kahneman 1992). While other forms are surely plausible, the one above is used out of deference to the fact that it has been so widely-adopted and seems to (roughly) capture the relevant perceptual phenomenon of probability discrimination which is the focus here. Apart from the parameters $\theta_i^{-1}$ and $\gamma_i$, $\lambda(b\theta_i^{-1},\gamma_i)$ also takes as input $b$: an endpoint of the relevant subinterval generated by $\theta_i^{-1}$. The non-uniformity of probability subintervals caused by $\lambda(b\theta_i^{-1},\gamma_i)$ for finite values of $\theta_i$ requires some notational changes in representing the quantization process from Subsection 6.2.1. (To save space this updated notation is omitted, and readers are directed to Rook (2014a) for a map of notational conversion in using non-uniform decision intervals.) Observe that the departures from a uniform partitioning of any dimension of the doxastic space become more severe as $\theta_i$ decreases in its size.

To assist this observation, note that two agents sharing the same $\theta_i$ will partition the unit interval into an equal number of disjoint subintervals, regardless of how the lengths of these subintervals might differ between the agents. Maintaining the indexing convention for these subintervals from before of $v \in \boldsymbol{V} = \{1, \ldots, \theta_i\}$, we can measure the cumulative absolute differences between the lengths of paired subintervals for some agent $x$ for whom $\gamma_x = 1$ (i.e., her subintervals are of uniform length because, assuming finite $\theta_x$, and $\zeta_x = 0$), and some agent $y$ for whom $0 < \gamma_y < 1$, and $\zeta_y = 1$, causing her subintervals to be of non-uniform lengths), with $\theta_x = \theta_y$, according to the

following formula for 'cumulative absolute length difference (ALD)', where $\left\|c_v^i\right\|$ is the length of subinterval $c_v$ for agent $i$ (notice: this setup permits Condition 1 of Theorem 1, $c_{v^*}^j \neq c_{v^*}^k,$ to hold):

(**25**) Cumulative ALD $\left(\theta_x, \theta_y\right) = \sum_{v=1}^{\theta_x} \left|\left(\left\|c_v^x\right\| - \left\|c_v^y\right\|\right)\right|$

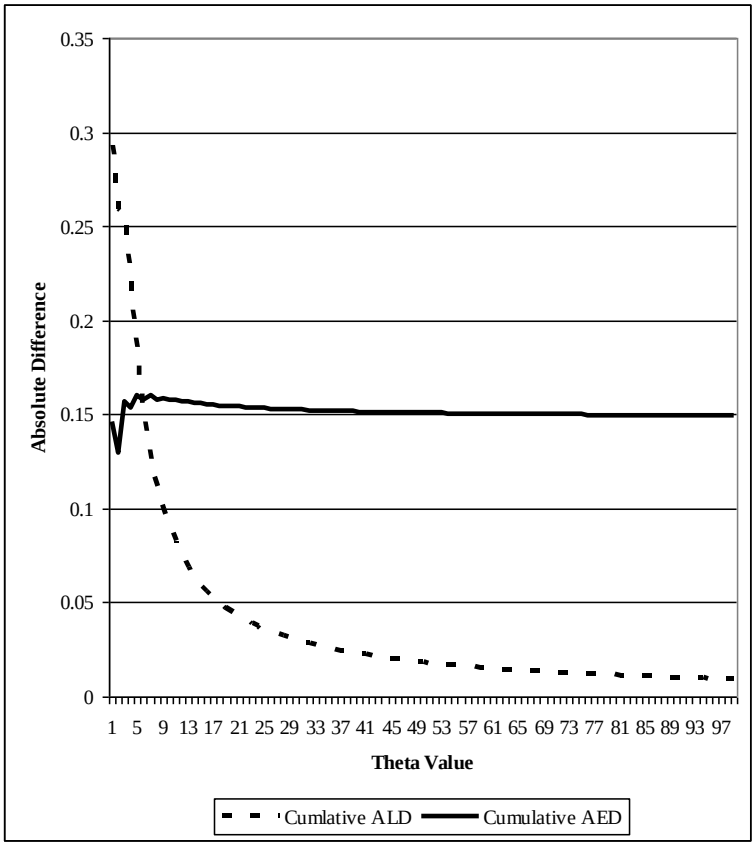
Figure 6.1 shows that the value of Cumulative ALD drops sharply as $\theta_i$ increases, but does not disappear for finite $\theta_i$ when one agent has $\zeta_i = 1$. Yet, increasing $\theta_i$ also means that both agents' capacities for discriminating between probabilities increase as well. To see that two agents $x$ and $y$ may still be vulnerable to disagree about consensus as $\theta_{\dot{x}} = \theta_y$ increases (and $\zeta_x \neq \zeta_y$), consider the 'cumulative absolute endpoint difference (AED)' rather than cumulative ALD as in Eq.25. That is, rather than considering the absolute differences in the lengths of the respective paired subintervals of the probability scale used by $x$ and $y$, consider the absolute differences between the (either upper or lower) endpoints of such subintervals. Figure 5.1 shows that such cumulative AED decreases slowly compared to increases in the shared discrimination parameter, $\theta_x = \theta_y$.[212]

As a quick illustration of the influence of nonlinearity on disagreements about consensus, consider two agents, $j$ and $k$, with: $\theta_j = \theta_k = 5$ and $\zeta_j = 0$, $\zeta_k = 1$ (i.e., the pair share a discrimination parameter, but $k$ uses subintervals of non-uniform length, according to Eq.24, whereas the five probability subintervals used by *j* are all of equal length $\theta_j^{-1} = 0.2$). If one agent holds a subjective probability of $24\%$, and the other holds a probability of $20\%$, for the hypothesis in question, then, *ceteris paribus*, $k$ would perceive consensus (their beliefs would both occupy her lowest

[212] Even with a shared value of $\theta_x = \theta_y = 10{,}000$ cumulative AED still exceeds 0.148; a value for the discrimination parameter at this level is equivalent to an agent possessing precision in probabilities to a level of $1 / 100^{\text{th}}$ of a percent!

probability subinterval, generating entropy of 0), whereas *j* would detect no such consensus (she would perceive entropy of 1, if base-2 logarithms are used): they would disagree about agreement.

*Figure 6.1: Cumulative Differences in Subintervals*




Hence, this section shows that heterogeneous degrees of nonlinear distortion of the probability scale may be more influential for asymmetric misperceptions of consensus than are systematic misperceptions of consensus (at least when implemented through a fixed scalar as was done here). Furthermore, both of these additional forms of cognitive diversity supported the necessity and sufficiency of the five conditions stated in Theorem 1 for ACMs to arise. One might nevertheless wonder what conditions might be necessary for an ACM to become 'dislodged'. That is, under the EBSL setup, what sort of information might be required to arrive that should cause all agents to

realize that no consensus exists among their beliefs, even when an ACM initially exists among them? The next section addresses this question and sketches implications for rational expectations.

## **6.6**\\ Dynamics of Cognitive Heterogeneity

The aim of this section is to demonstrate, through a series of three indicative cases (which shall serve as existence proofs), that situations can arise wherein dislodging an ACM requires arrival of an informational signal that is neither too large nor too small in size. This requirement of relative signal sizes that are 'just right' to dislodge an ACM is called the 'Goldilocks Condition' of EBSL:

> **Conjecture** ("Goldilocks Condition"): *When, for two agents $j \neq k$, it holds that $c^{j}_{v^*} \neq c^{k}_{v^*}$, and these two agents are initially in a state of asymmetric consensus misperception (according to Theorem 1 stated earlier), then there can exist a strength for an exogenous signal below which the asymmetric consensus misperception persists, and above which a new asymmetric consensus misperception occurs, and intermediate signal values falling between these two thresholds suffice to eliminate any asymmetric consensus misperception.*

The Goldilocks Condition can sometimes be needed to dislodge an ACM: both agents must receive an exogenous signal that is 'just right' in strength relative to their respective accessible evidences: it cannot be too small, or else the misperceived consensus will persist; and it cannot be too large, or else the agents will merely be pushed from one state of asymmetric consensus misperception into another. Precisely characterizing this Goldilocks range relative to accessible evidence lies beyond the scope of this paper, but some demonstrative cases that follow may aid with intuition. It is shown that these demonstrative cases arise from various configurations of cognitive diversity.

### **6.6.1**\\ Dislodging Misperceived Consensuses

The basic situation which we shall consider here is one of two agents, $j$ and $k$, who at time $t = 0$ are in a situation of asymmetric consensus misperception. At $t = 1$ they receive an exogenous signal and update their beliefs as described in Section 6.3 (specifically, they both witness the average of those quantized beliefs which change at any point in time; unchanged beliefs are not

publicly visible, and for the belief of an agent to qualify as 'changed' between two given periods it must move into another subinterval at the specific level of $\theta_i$ between one period and the next). As in Section 6.3, time passes in discrete periods, and agents are able to update their beliefs simultaneously, although they can only see where the average changed belief between them ends up at each period's finish. Both are able to observe the entropy and agglomeration values for the pair, but only as influenced and allowed by their respective cognitive parameters: $\theta_i$, $\delta_i$, $\varphi_i$, and $\zeta_i$ (in particular, agents $j$ and $k$ can perceive consensus only at the levels of $\theta_j$ and $\theta_k$, respectively). For simplicity, let both agents consider only a single hypotheses ($F = 1$) in the three ensuing cases.

**6.6.1.1**\\ Demonstrative Case #1

To begin, let $\theta_j = 3$ and $\theta_k = 5$, with $\delta_j = \delta_k = 0.5$, $\varphi_j = \varphi_k = 0.5$, and $\zeta_j = \zeta_k = 0$, and $\widehat{\boldsymbol{d}}_{j,t}^{A_f} = 1$, $\widehat{\boldsymbol{d}}_{j,t}^{\neg A_f} = 5$, $\widehat{\boldsymbol{d}}_{k,t}^{A_f} = 2$, $\widehat{\boldsymbol{d}}_{k,t}^{\neg A_f} = 8$. That is, both agents are operating at their cognitive limits in terms of evidence summoned to mind. Furthermore, let neither agent have clear visibility on the beliefs of the other agent, and let each suppose that the average belief between them is the same as their own beliefs. Then the pair will disagree about whether their beliefs are in consensus. Before receiving the exogenous signal, agent $j$ will formulate a (quantized) belief as a probability for the truth of hypothesis $A_f$ based upon her accessible evidence ($\widehat{\boldsymbol{d}}_{j,t}^{A_f} = 1$, $\widehat{\boldsymbol{d}}_{j,t}^{\neg A_f} = 5$) that is equal to $\frac{1}{6}$. Meanwhile, $k$ will form beliefs based upon here accessible evidence equal to a quantized value of $\frac{3}{10}$ (due to influence of the parameter value $\delta_k = 0.5$). Thus, from the perspective of $j$, both agents' initial probability estimates fall within her first decision interval; whereas from the perspective of $k$, $j$'s belief falls into $k$'s second decision interval while $k$'s own belief falls into her second decision interval. Agent $j$ will detect perfect consensus, however, as she will perceive (if, as

previously, logarithms of base 2 are used) an entropy of 0, whereas $k$ will detect an entropy of 1, and recognize no such consensus. What conditions could then dislodge the ACM in this situation?

To start, consider a signal of size $g = 1$ arriving at the start of period $t = 1$. This signal is too weak to dislodge the ACM (if both agents follow the EBSL updating rules described in Section 6.3), as neither agent will change her quantized beliefs. A signal $g = 2$, however, does suffice to remove the misperceived consensus, as both agents come to agree that consensus exists at the end of period $t = 2$ (i.e., there is one period of anticipatory updating and another period of reactive updating; Rook (2014b) calls the latter type of updating belief "echoing") with $j$ believing both agents' beliefs occupy her second decision interval, and $k$ believing that both beliefs occupy her third decision interval: both agree that there is consensus (i.e., consensus perception is symmetric).

Likewise, with positive signal values of $g = 3$ through $g = 9$, the misperceived consensus will be dislodged (although for signal sizes $g = 4$ through $g = 8$ no reactive updating occurs). Yet, for signal sizes $g = 10$ through $g = 29$, the agents are forced into another situation of ACM: $j$ will perceive entropy of 0 for both agents' updated beliefs (i.e., total agglomeration and perfect consensus), whereas $k$ will detect an entropy of 1 and see that no consensus exists. Positive signal values of $g = 30$ and larger suffice to fully-eradicate asymmetric consensus misperceptions for those conditions (i.e., parameter values, initial evidence, and number of agents) that are used here.

#### **6.6.1.2**\\ Demonstrative Case #2

As a second demonstration of the Goldilocks Condition, consider again the setup in Case #1, but with the alterations that $\theta_j = 9$ and $\theta_k = 10$, and the initial evidences accessible to the agents is: $\widehat{\boldsymbol{d}}_{j,t}^{A_f} = 1$, $\widehat{\boldsymbol{d}}_{j,t}^{\neg A_f} = 17$, $\widehat{\boldsymbol{d}}_{k,t}^{A_f} = 2$, and $\widehat{\boldsymbol{d}}_{k,t}^{\neg A_f} = 18$. This setup once more finds $j$ initially perceiving a consensus, and $k$ (correctly) discerning that no consensus actually exists. Under these new

conditions we find again that $g = 1$ cannot dislodge this misperceived consensus. And $g = 2$ only serves to alter the perpetrator of the ACM: at the end of period $t = 1$ (after the signal's arrival and both agents have engaged in anticipatory updating), agent $k$ perceives a consensus, whereas $j$ does not. From $j$'s perspective, $k$'s prior belief fell into $j$'s first decision interval (as did $j$'s own prior). Given $j$'s initially-accessible evidence, $g = 2$ is adequate to move her amalgamated posterior belief into her second decision interval. But the evidence initially accessible to $k$ causes her not to change her quantized belief in response to $g = 2$. The result is that with $g = 2$ both agents still disagree about consensus, but their positions are reversed. A value of $g = 3$, however, convinces both agents that no consensus exists after they finish their belief-updating in response to the signal's arrival (i.e., no reactive updating occurs thereafter). Likewise for positive signal values of $g = 4$ through $g = 12$, both agents agree that their beliefs are not in consensus. But values of $g = 13$ through $g = 15$ serve to push $j$ to believe that a consensus persists, even though $k$ disagrees. Signal values of $g = 16$ and higher prevent asymmetric consensus misperception from happening.

**6.6.1.3**\\ Demonstrative Case #3

Finally, consider a third case of dislodging asymmetric consensus misperceptions. Here initially let: $\theta_j = 5,\ \theta_k = 9,\ \delta_j = \delta_k = 0.5,\ \zeta_j = 1,\ \zeta_k = 0,\ \varphi_j = \varphi_k = 0.5,$ and $\widehat{\boldsymbol{d}}_{j,t}^{A_f} = 1,\ \widehat{\boldsymbol{d}}_{j,t}^{\neg A_f} = 9,$ $\widehat{\boldsymbol{d}}_{k,t}^{A_f} = 3,$ and $\widehat{\boldsymbol{d}}_{k,t}^{\neg A_f} = 15.$ As in the previous two cases, agent $j$ misperceives (at the onset, and before receiving the exogenous signal) a consensus, whereas agent $k$ does not. And, as previously, $g = 1$ at $t = 1$ is too weak to dislodge the misperceived consensus, whereas values of $g = 2$, $g = 3$, and $g = 4$ serve to disabuse both agents, and cause them to agree that consensus does in fact exist after $t = 1$. Yet, for $g = 5$, $g = 11$ through $g = 23$, and $g = 27$ through $g = 116$, an ACM will persist, as agent $j$ will perceive a consensus while $k$ recognizes that no consensus actually

exists. (Signal values of $g = 6$ through $g = 10$, $g = 24$ through $g = 26$, and greater than $g = 116$ find both agents agreeing on whether consensus exists or not among their respective beliefs.)

Moreover, notice from the three cases presented here that the Goldilocks Condition does not always pertain to the entire possible range of signal values, but instead to only subsets of it. That is, for some interval of signal values, $[g = \omega, g = x]$, the asymmetric misperception of consensus will be dislodged, but for some values of $g < \omega$ or $g > x$ the asymmetric consensus misperception will persist (or else a new one will arise). But observe that this statement about misperceptions outside of the interval does not apply to all signal values $g < \omega$ or $g > x$; for example, for some value $g = y > x$ both agents may come to agree on whether consensus exists between their beliefs.

Furthermore, it is interesting to realize that the Goldilocks Condition harbors some intriguing implications for economic study of the existence and durability of rational expectations equilibria.

### **6.6.2**\\ Implications of Misperceived Consensus for Rational Expectations Equilibria

Economic concepts of rational expectations equilibria (REE) generally involve some form of common knowledge and agreement among participating agents (Tirole 1982). Indeed, given the endogenous properties of REE, assumptions of (some form of) common knowledge among agents is practically indispensable in dealing with either the attainment or stability of any REE. In practice, however, common knowledge (and its weaker sibling: common belief) is extraordinarily demanding of even cognitively unbounded agents. We could readily expect, then, that agents with cognitive constraints (i.e., all actual agents) might struggle to achieve any sort of common knowledge (or belief) unless they are privileged enough to have some kind of summary statistic(s) about their worlds. The doxastic agglomeration measure that was discussed in this paper represents one such statistic. But in the real world it seems likely that the degree of doxastic agglomeration

must be inferred, instead of being given. And, as has been shown, unless agents are on equal cognitive footings, and do not significantly differ from one another along important dimensions of cognitive ability (especially those represented in this paper by the cognitive parameters $\theta_i$, $\varphi_i$, and $\zeta_i$), genuine problems can arise in their proper perceptions of whether or not common knowledge and consensus actually exist among themselves. And these same difficulties hold equally true for researchers who analyze the agents. One contribution of this paper has been to show the rather large number of situations in which agents may hold errant views about the degree of consensus among themselves due to only relatively modest (and even small) differences in cognitive abilities.

Thus, even if agents are able to occasionally converge on a rational expectations equilibrium (REE), it may often be something of a misnomer to consider such an arrangement as an 'equilibrium' (or at least a stable one), unless all agents involved equally understand it as such. The results from this paper cast doubt that such understanding will generally be universal among agents when cognitive diversity exists among them. Nevertheless, there remains the possibility that if the cognitive dimensions treated here are time-variable, and that these may converge among agents toward homogeneity through some form of learning process, then there is hope that attainment and stability of REE may be as prevalent and durable as conventionally supposed. Results from laboratory experiments make this possibility doubtful. While many findings do show that cognitive parameters can be strongly time-varying, there is as yet no strong evidence that seems to support a conclusion that, with practice, there is strong tendency toward homogenization.

One consequence of the alternative to prevalent rational expectations equilibria might, however, be that mentioned in the introduction (in Subsection 6.1.2): substantial and irregular 'surprises' emanating from compositional uncertainties among market agents as their suppositions about the level and location of consensus in a market become suddenly disproven. Such a world of oscillating

uncertainty about agreement may nonetheless prove important in the pursuit of resolving various asset-pricing puzzles which have intrigued financial economists for decades. Such curiosities include, for example, many open questions surrounding the time-variability of measured levels of risk aversion. If, as suggested in the introduction, some of this aversion is actually distaste for uncertainty, and if compositional uncertainty is part of the uncertainty which agents dislike, then surprises due to dislodging of misperceived consensuses may be a significant contributor to this time-variability conundrum. Also, time-varying cognitive abilities among market participants may only serve to strengthen such effects. Investigating the truthfulness of these possibilities seems to merit future research, and ideas in this paper may provide a framework for just such investigation.

## **6.7**\\ Conclusion and Extensions

This paper has sought to investigate conditions under which agents might disagree about whether their beliefs are in consensus. The main focus has been conditions under which cognitive heterogeneity among agents drives asymmetric misperceptions about the levels of consensus that various agents detect with respect to the beliefs which they and other agents possess. A main contribution of this paper has been to formulate a new operational surrogate for the ideas of common knowledge and common belief which underpin many economic theories. The operational replacement suggested in this paper involves a measure of doxastic agglomeration derived from the distributional entropy of the beliefs of all agents in a group. It was shown that, given some basic knowledge (or else belief) conditions, mutual acceptance of this doxastic agglomeration measure could suffice to achieve common knowledge (or else common beliefs). It was nonetheless demonstrated that agents' interpretations of this agglomeration measure could differ according to their cognitive abilities. Hence, cognitive diversity can jeopardize appropriate detection of common knowledge or belief among agents who are subject to limits on their cognitive resources.

Despite the ground covered by this paper, a substantial amount of work remains in several respects. Firstly, EBSL is still in its early days as a theoretical device, and, despite grounding in substantial economic and psychological precedent, requires additional refinement and empirical testing to further verify its credibility. Secondly, this paper has explored only a few of the enormous number of combinations of cognitive diversity that might plausibly underlie not only consensus misperceptions, but also other interesting economic phenomena. Thirdly, while some may find ideas in this paper worthwhile for their own sakes, others might find them more valuable for their bearing on extant theories and findings elsewhere in economic research. To conclude this paper, therefore, some possible routes are traced for marrying elements of ESBL and consensus misperception with two other strands of recent research concerning boundedly-rational behavior.

### **6.7.1**\\ Possible Syntheses with Other Research

The first strand could be called "inferential naivety" (Eyster and Rabin's 2010), and involves difficulties people have in making accurate inferences about others' information, beliefs, and strategic thinking from observing others' behaviors. In two papers, Eyster and Rabin (2005; 2010) describe twin poles of inferential naivety: one where agents' systematically overweight their own beliefs and evidence; and another where they overweight the beliefs and evidence of (some) others (see also Weizsäcker (2010)). While neither pole as described by Eyster and Rabin explicitly deals with consensus detection by agents, the latter pole seems to involve elements of the fundamental attribution error (FAE) described by Jones and Nisbett (1972). Agents who commit FAEs see their

own judgments and behaviors as influenced by contextual forces, but see others' behaviors as less driven by situational factors than their own (and more revealing others' information and beliefs).[213]

Which pole of inferential naïvety an agent occupies therefore may be linked to whether or not she commits errors of fundamental attribution. An unanswered question is whether or not both inferential naivety and attribution errors are influenced by the strength of perceived consensuses. The EBSL model suggests that when perceived consensus is low, agents tend to under-infer about others' evidence and information (as in Eyster and Rabin's (2005) "cursed equilibrium"), but when perceived consensus is high they instead tend to over-infer (which may cause naïve herding, as in Eyster and Rabin (2010)). A better understanding of how cognitive diversity impacts consensus misperceptions by agents may help to improve models of behavior with inferentially-naïve agents.

The second wave of research that might benefit from incorporating elements of misperceived consensuses is that conducted under the banners of "cognitive hierarchies" and "level-k thinking" (e.g. Camerer, Ho, and Chong (2004), Costa-Gomes and Crawford (2006), Crawford and Iriberri (2007); see Crawford, Costa-Gomes, and Iriberri (2013) for a comprehensive review of both early and recent work in the area). These approaches to behavior in non-cooperative settings prioritize cognitive diversity by respecting two dimensions of heterogeneity in strategic thinking noted by Thaler (2000): 1) degree of cognitive sophistication; and 2) degree of other-regard. The former is usually understood as the level of agents' logical skill, and the latter as "social intelligence", or degree to which they comprehend how skilled in logic others are (Camerer, Ho, and Chong 2003).

[213] A curious corollary of the fundamental attribution error would be that in the presence of framing effects people should interpret the behavior of others as less impacted by frames than would be their own judgments and choices. Notably, Kahneman and Tversky (1973) identify a similar concept in what they designate as "fallacy of initiative".

Furthermore, like EBSL, cognitive hierarchy/level-k (CH/L-k) approaches are novel for explicitly modeling the relationship(s) between cognitive diversity and heterogeneous beliefs among agents. Rather than being competing approaches to synthesizing diversity among agents' beliefs and cognitive abilities, however, EBSL and CH/L-k approaches appear to exhibit both compatibility as well as the potential to remedy each other's weaker points. Firstly, CH/L-k approaches necessarily assume that agents utilize only a limited number of 'decision rules'. Specifying these rules *a priori* can present the theorist or modeler with myriad difficulties (Costa-Gomes and Crawford 2006), and also may considerably restrict the degrees of diversity that CH/L-k approaches can encompass (this restricted diversity may be problematic for realistic settings wherein the incompleteness of information is severe, and the structure of problems not readily apparent to all agents). Furthermore, CH/L-k approaches do not typically deal with learning, but usually are concerned instead with first responses; generating accurate predictions on these first responses actually seems a main *raison d'être* for the CH/L-k models of non-cooperative behavior.

Additionally, CH/L-k approaches essentially embed only a single dimension along which agents can be cognitively diverse: their cognitive level, or the number of steps of reasoning which they employ to form their beliefs and beliefs about others. Thus, under CH/L-k, agents only differ in cognitive ability in one way and are assumed to correctly perceive the level of consensus for all agents that are less cognitively able than themselves. Such a single dimension of cognitive heterogeneity and partial accuracy of consensus detection may be suitable for many game settings, but, for many real-world applications (e.g., financial markets), broadening the hierarchy by allowing agents to be cognitively diverse along multiple dimensions (as in EBSL) may be helpful.

EBSL, meanwhile, leaves open the question of how agents select among the evidence available to them in formulating their initial beliefs. It does specify how agents form initial beliefs once such

evidence has been selected and discretized, and stipulates how agents make use of their perceptions of consensus to update their beliefs in light of potentially incomplete information. Thus, whereas CH/L-k approaches may be overly restrictive on diversity for particular applications, EBSL may be too unrestrictive on diversity. While not undertaken here, a future marriage of EBSL and CH/L-k approaches might be fruitful for achieving some happy middle-ground not only for levels of cognitive and belief heterogeneity, but also in describing how agents with diverse constraints upon their thinking abilities behave in non-cooperative settings *both* initially and over time as they learn.

Moreover, a union with EBSL may allow CH/L-k to more directly address strategic behavior under ambiguity. At present, CH/L-k approaches do not expressly accommodate agents who entertain multiple priors (as in, e.g.: Gilboa and Schmeidler (1989); Hansen and Sargent (2001; 2008)), whereas EBSL implicitly embeds multiple priors through usage of decision intervals. Future work on CH/L-k may embrace ambiguity by adopting EBSL's method for representing multiple priors.

The foregoing discussion of possible fusions between inferential naivety, CH/L-k approaches, and EBSL hint that a future combination of these ideas may help to further Simon's (1974) aim of parameterizing more accurately that ultimate substrate of all economic behavior: the human mind.

## **APPENDIX 6.A**\\ Proof of Theorem 1

In proving Theorem 1, one aims to show that, for two agents $j \neq k$: $H_{\theta_j} = 0$, whereas $H_{\theta_k} \neq 0$ (in the proof, the perceptions of consensus and beliefs of both agents are taken to be drawn from the exact same time period; time subscripts are suppressed for notational tidiness). Further simplifying notation, we can take the beliefs of *j* and *k* about some hypothesis (which is assumed to be identical for both agents) as $\pi_j$ and $\pi_k$, respectively. From the main text, all of the following are then true by definition: $\pi_j \in c^j_{v^*}$; $\pi_k \in c^k_{v^*}$; $\pi_j \notin c^j_{v^-}$; $\pi_k \notin c^k_{v^-}$. Additionally, it should be evident from Equation 1 (contained in the main text) that, with only two agents, $H_{\theta_j} = 0$ requires that $\pi_j \in c^j_{v^*}$ and $\pi_k \in c^j_{v^*}$ both hold (and likewise $H_{\theta_k} = 0$ requires that $\pi_j \in c^k_{v^*}$ and $\pi_k \in c^k_{v^*}$ both hold true).

Demonstrating that $H_{\theta_j} = 0$ while $H_{\theta_k} \neq 0$ therefore reduces to a task of showing that $\pi_k \in c^j_{v^*}$ and $\pi_j \notin c^k_{v^*}$, since $\pi_j \in c^j_{v^*}$ and $\pi_k \in c^k_{v^*}$ are true by definition. Theorem 1 is then essentially a claim about the five necessary and sufficient conditions to allow $\pi_k \in c^j_{v^*}$ and $\pi_j \notin c^k_{v^*}$ such that $H_{\theta_j} = 0$ and $H_{\theta_k} \neq 0$. Once again, these conditions are:

(i) $c^j_{v^*} \neq c^k_{v^*}$
(ii) $\left(c^j_{v^*} \cap c^k_{v^*}\right) \neq \emptyset$
(iii) $\left(c^j_{v^*} \cap c^k_{v^-}\right) \neq \emptyset$
(iv) $\pi_{k,t}(A)|\theta_k \in \left(c^j_{v^*} \cap c^k_{v^*}\right)$
(v) $\pi_{j,t}(A)|\theta_j \in \left(c^j_{v^*} \cap c^k_{v^-}\right)$

Condition (i) is required because if $c^j_{v^*} = c^k_{v^*}$, then both $\pi_j \in c^k_{v^*}$ and $\pi_k \in c^j_{v^*}$ would hold, and therefore $H_{\theta_j} = H_{\theta_k} = 0$. Thus $c^j_{v^*} \neq c^k_{v^*}$ is necessary. Condition (ii) is required because if $\left(c^j_{v^*} \cap c^k_{v^*}\right) = \emptyset$ held, then $\pi_k \notin c^j_{v^*}$ would be true and hence $H_{\theta_j} \neq 0$. $\left(c^j_{v^*} \cap c^k_{v^*}\right) \neq \emptyset$ is thus needed for $H_{\theta_j} = 0$ and $H_{\theta_k} \neq 0$. Similarly, Condition (iii) is necessary because if $\left(c^j_{v^*} \cap c^k_{v^-}\right) =$

$\emptyset$ then it could be true that $c_{v^*}^{j} \subset c_{v^*}^{k}$, and thus $\pi_j \in c_{v^*}^{k}$ such that $H_{\theta_k} = 0$. Hence, Condition (iii) is needed for Proposition 1. Condition (iv) simply imposes the requirement that $\pi_k \in c_{v^*}^{j}$, and Condition (v) demands $\pi_j \in c_{v}^{k}$- which thereby means $\pi_j \notin c_{v^*}^{k}$. Collectively these five conditions cause $\pi_k \in c_{v^*}^{j}$ and $\pi_j \notin c_{v^*}^{k}$, such that $H_{\theta_j} = 0$ and $H_{\theta_k} \neq 0$, which therefore proves Theorem 1.

□

## **APPENDIX 6.B**\\ Proofs on Fixed-Scalar Approaches to Consensus Misperception

The aim of this appendix is to prove the truth of two propositions about how fixed scalars $\varphi_i \in (0,1) \subset \mathbb{R}_+$ impact asymmetric consensus misperceptions through altering the weights that agents place on their primary and secondary beliefs. This alteration was proposed (in Subsection 6.5.1) to influence those weights through the following pair of expressions (which together sum to unity):

$$\frac{(1-\varphi_i)^{-1} H_{\theta_i,t}}{\varphi_i^{-1}\xi_{\theta_i,t} + (1-\varphi_i)^{-1} H_{\theta_i,t}} \quad \text{and} \quad \frac{\varphi_i^{-1}\xi_{\theta_i,t}}{\varphi_i^{-1}\xi_{\theta_i,t} + (1-\varphi_i)^{-1} H_{\theta_i,t}}$$

The first proposition to be proved is:

**Proposition A1**: *$\varphi_i$ alone is insufficient to force disagreement about consensus.*

To see that Proposition A1 is true, consider two agents $j \neq k \in I$. Let both agents perceive total consensus among all considered hypotheses such that $H_{\theta_{j,t}} = H_{\theta_{k,t}} = 0$ at time $t$ (the generality of the argument allows for $\theta_j \neq \theta_k$ such that $\xi_{\theta_{j,t}}$ need not equal $\xi_{\theta_{k,t}}$ at the time $t$). The relevant proof will be by contradiction. What is needed to be shown is that, for any $\varphi_j, \varphi_k \in (0,1) \subset \mathbb{R}_+$, the following inequality fails to hold whenever $H_{\theta_{j,t}} = H_{\theta_{k,t}} = 0$:

$$\frac{\varphi_j^{-1}\xi_{\theta_j,t}}{\varphi_j^{-1}\xi_{\theta_j,t} + \left(1-\varphi_j\right)^{-1}H_{\theta_j,t}} \neq \frac{\varphi_k^{-1}\xi_{\theta_k,t}}{\varphi_k^{-1}\xi_{\theta_k,t} + (1-\varphi_k)^{-1}H_{\theta_k,t}}$$

Because $H_{\theta_j,t} = H_{\theta_k,t} = 0$, the above inequality reduces to:

$$\frac{\varphi_j^{-1}\xi_{\theta_j,t}}{\varphi_j^{-1}\xi_{\theta_j,t}} \neq \frac{\varphi_k^{-1}\xi_{\theta_k,t}}{\varphi_k^{-1}\xi_{\theta_k,t}}$$

Which itself reduces to $1 \neq 1$, which is false and therefore proves Proposition A1. Likewise one could demonstrate the same result using:

$$\frac{\left(1-\varphi_j\right)^{-1}H_{\theta_j,t}}{\varphi_j^{-1}\xi_{\theta_j,t} + \left(1-\varphi_j\right)^{-1}H_{\theta_j,t}} \neq \frac{(1-\varphi_k)^{-1}H_{\theta_k,t}}{\varphi_k^{-1}\xi_{\theta_k,t} + (1-\varphi_k)^{-1}H_{\theta_k,t}}$$

As the inequality above reduces to $0 \neq 0$ and is false, this path likewise proves Proposition A1.

The second proposition to be proved is:

**Proposition A2**: *$\varphi_i$ alone is insufficient to force agreement about consensus.*

As with Proposition A1, the proof of A2 will proceed by contradiction. Once again, consider two agents $j \neq k \in I$, but now assume that $H_{\theta_j,t} = 0$ whereas $H_{\theta_k,t} \neq 0$. Consider the following equality for any $\varphi_j, \varphi_k \in (0,1) \subset \mathbb{R}_+$:

$$\frac{\varphi_j^{-1}\xi_{\theta_j,t}}{\varphi_j^{-1}\xi_{\theta_j,t} + \left(1-\varphi_j\right)^{-1}H_{\theta_j,t}} = \frac{\varphi_k^{-1}\xi_{\theta_k,t}}{\varphi_k^{-1}\xi_{\theta_k,t} + (1-\varphi_k)^{-1}H_{\theta_k,t}}$$

Because $H_{\theta_j,t} = 0$ but $H_{\theta_k,t} \neq 0$, the above equality reduces to:

$$1 = \frac{\varphi_k^{-1}\xi_{\theta_k,t}}{\varphi_k^{-1}\xi_{\theta_k,t} + (1-\varphi_k)^{-1}H_{\theta_k,t}}$$

Which requires $H_{\theta_k,t} = 0$ in order for the above equality to be true. But since $H_{\theta_k,t} \neq 0$, the above equality fails to hold, which proves Proposition A2. Realize that the same result arises if instead one works with the following equality:

$$\frac{\left(1-\varphi_j\right)^{-1} H_{\theta_j,t}}{\varphi_j^{-1}\xi_{\theta_j,t} + \left(1-\varphi_j\right)^{-1} H_{\theta_j,t}} = \frac{(1-\varphi_k)^{-1} H_{\theta_k,t}}{\varphi_k^{-1}\xi_{\theta_k,t} + (1-\varphi_k)^{-1} H_{\theta_k,t}}$$

Here again, the fact that $H_{\theta_j,t} = 0$ causes the above equality to reduce to:

$$0 = \frac{(1-\varphi_k)^{-1} H_{\theta_k,t}}{\varphi_k^{-1}\xi_{\theta_k,t} + (1-\varphi_k)^{-1} H_{\theta_k,t}}$$

As the equality above demands that $H_{\theta_k,t} = 0$ in order to be true, but $H_{\theta_k,t} \neq 0$, Proposition A2 is yet again proven true. The proof of both Proposition A1 and Proposition A2 is therefore completed.

□

# CHAPTER 7

# CONCLUSION AND EXTENSIONS

## **7.1**\\ Taking Stock

The introduction to this thesis opened with a question on how human beliefs should be studied. There it was asserted that the proper study of human beliefs relies jointly upon: the particular definition for beliefs used; and the purpose of any such study. A claim undergirding this thesis has been that, when beliefs are defined as subjective expectations on the truthfulness of propositions about the world, then the doxastic-space framework presented in this thesis is generally a useful approach to studying human beliefs. Yet, to judge the fitness of this approach, the intention behind its use must also be considered. And, over the course of this project, that intention has remained the desire to understand the formation of, and changes to, human beliefs in ways that accentuate the three key properties attributed to them in the opening: *relativity*, *relationality*, and *reflexivity*. In closing, it therefore seems suitable to briefly reflect on how the doxastic-space framework, in its present state of development, achieves this purpose, as well as how it might be further improved.

First, the doxastic-space framework has exhibited an ability to facilitate research on beliefs in a manner underscoring their relative (and often unstable) nature. Particularly, the model of quantized evidence theory (QET) developed in Chapter 4, and extended through the model of entropy-based social learning (EBSL) in Chapters 5 and 6, demonstrated how the belief-formation process might be enduringly relative. Specifically, exploring QET through the doxastic-space approach showed how the relative accessibility of evidence to a decision-maker, and the relative cognitive abilities of the decision-maker herself, might fuse to drive the beliefs she holds about the world which surrounds her. Chapter 6 especially demonstrated how perceptions of agreement between decision-makers might be relative to their respective working memory and attentional (and other cognitive) capacities. Furthermore, Chapter 4 showed how the relativity of context (in terms of the confluence of the judgment problem facing a decision-maker, her available cognitive resources, and accessible

strength and weight of evidence) can be crucial for determining how decision-makers distort their beliefs, and may be a key ingredient in explanations for why beliefs behave as they are seen to do.

Second, the doxastic-space framework has evinced the capacity to reflect the relational nature of human beliefs in two main ways, especially through implementation of the EBSL model within it. For one, such a setup has shown how doxastic spaces can capture the relational characters of beliefs in an intra-personal sense, by appreciating connections between distinct beliefs held by the same decision-maker (e.g., through forces such as juncture; see Chapter 6). And this approach to intra-personal relationality in beliefs seems to harbour more potential for faithfully describing relationships among a single decision-maker's beliefs than conventional state-space approaches, because of the selective and syntactic natures that doxastic spaces can exhibit; that is, rather than restricting relationships between an individual decision-maker's beliefs to exist primarily at some global level of the universe of all beliefs she might possibly hold, the doxastic-space approach (with EBSL implemented in it) focuses on intra-personal belief relationships at a more realistic level: that of those beliefs (and hypotheses) which she is capable of entertaining in mind at once. This accentuation of active belief relationships seems to accord better with reality, in that people facing cognitive limitations are not likely to continuously and perfectly curate a universe of beliefs.

For another thing, the doxastic-space framework here has also proven that it can handle intersubjective relationships among decision-makers' beliefs. Again, this intersubjective connectivity among people's beliefs was perhaps most pointedly conveyed through explorations with the EBSL model, wherein decision makers' beliefs and learning was influenced by their awareness of the beliefs of their peers. Yet this other-regarding characterisation of beliefs in EBSL also demonstrated that doxastic spaces, and the entropy-based methods which they enable, are adept at reflecting some of the reflexive characteristics of beliefs. Particularly, this reflexivity was

embedded through the fact that decision-makers using EBSL are concerned with what a 'typical' member of their peer group believes, and since their own beliefs are involved in their estimations of typicality, there is a natural reflexivity which arises in the belief dynamics stemming from EBSL (i.e., if a decision-maker adjusts her beliefs to account for the average belief among all relevant agents, then she may effectively change the average, which may prompt more changes, and so on). In essence, the entropy-based approaches to studying human beliefs that the doxastic-space framework facilitates (but are far more difficult if not impossible to utilise in conventional state-space setups) allow for perhaps the closest union between relationality, relativity, and reflexivity. Moreover, while Chapter 3 was less formal in its mode of inquiry than its three successors, and dealt with doxastic spaces less directly (opting instead to treat neural spaces – which ultimately must underpin and restrict doxastic spaces), it nevertheless hinted at how frameworks capable of describing the interplay between belief reflexivity, relationality, and relativity accord with the inherently reflexive, relational, and relative nature of neural (and neuronal) functioning. More specifically, the main findings from Chapter 3 – that two of the main features characterising the brain and underpinning human thought, dynamic reflexivity and mult-scalar interactivity (especially between the 'more local' and 'more global' chains of neural activity) – starkly call into question whether it can be ever really assumed that people's beliefs (which must be driven by neural activity) can behave *as if* they are exogenous, hermetic, and universal (i.e., decidedly *not* reflexive, relational, or relative, respectively), as suggested by Savage's (1954) (normative) proposals for them. Even more concretely, Chapter 3 underscored the clear roles that context and neglect have to play in the formation and change of people's beliefs, but that context and neglect are themselves strongly endogenous: they are practically never simply 'imposed' on a thinker, but instead are nearly always (at least partly) *constructed* by the thinker.

With direct application to the future study of beliefs, such a situation stands strongly at loggerheads with ideas about a decision maker simply being endowed with a set of beliefs, or that it is ever reasonable to assert that 'common priors' (or even mutal priors) exist about any aspect of the world other than the most trivial and basic. Indeed, the findings from this chapter about reflexivity, relationality, and relativity in the human brain both parallel and reinforce the ideas in Chapter 1 that people do not likely merely 'carry around' a set of priors with them that they smoothly update, but continually reconstruct their beliefs time and again in reflexive, relational, and relative ways. This parallelism suggests that future extensions of the doxastic-space framework might possibly be melded with further findings from cognitive and neurosciences in order to better represent and more penetratingly understand the genesis and dynamics among human beliefs. And, although that union ostensibly remains some distance in the future, the potential for such strikes as encouraging.

Nonetheless, several more immediately-improvable components of the doxastic-space approach exist, and will be discussed shortly. But first it seems prudent to revisit whether ideas cultivated thus far can meet the five desiderata culled from the literature review which comprised Chapter 2.

To remind readers, these five desirable features of frameworks for belief study in social science were deemed (from a selective review of pertinent literature) to include an ability to capture:

1. Significant belief diversity, but with allowance for consensus;
2. Dynamics that reflect various modes of belief change;
3. Various forms of (cognitively reconcilable) belief distortion;
4. Social influences in the formation and updating of beliefs;
5. The role which accessible evidence plays in belief dynamics

Without reviewing all of the specifics on how each of these desiderata were met in exhaustive detail, it may be comfortably claimed that each of these was achieved (to varying degrees) within the four body chapters of this thesis. To help assure the reader, however, an abridged listing of some examples of how the doxastic-space framework (and QET and EBSL) satisfy the above requirements includes the following. Desideratum 1 is met by the EBSL model's ability to range from maximum to minimum entropy (the latter equivalent to perfect consensus); and the expandability of the doxastic space is what permits extensive degrees of belief diversity. Desideratum 2 is fulfilled by the several forms of learning captured under EBSL (e.g., anticipatory and reactive updating). Desideratum 3 is met by both QET's ability to exhibit super- *and* subadditivity in probabilities, as well as the various potential distortions of consensus judgments which appear in Chapter 5. Desideratum 4 is fulfilled by the regard for others' beliefs which is embedded into the EBSL model. And Desideratum 5 is met by QET's direct reliance on evidence.

In total, it therefore appears that, in the space allotted for this thesis, the doxastic-space framework (and the new models of beliefs which it permits) has been adequately developed to meet all of the objectives set for it in this project. Moreover, it also succeeds in achieving this aims in a manner that is wholly-deterministic, which was prioritised as one further goal within the opening chapters. Nevertheless, science (including social science) should always have its eye toward future enhancements and further research. And it is that future-gazing that will now conclude this project.

## **7.2**\\ Future Extensions

While each of the four previous chapters presents its own suggestions for future research, which are mostly specific to the particular content of each, this subsection articulates a 'wish list' for subsequent investigations about, and extensions for, the doxastic-space framework more generally. To begin, perhaps one obvious deficiency of the framework as thus far presented is its divorce

from preferences. As was stated in the introduction, this hermetic approach that separates beliefs from utility concepts was intentional and has served a purpose in isolating beliefs as focal objects. Yet, as was previously remarked, a viable future for the doxastic-space framework may hinge upon its successful marriage with preference specifications (or even new modes of representing and analysing preferences) which are faithful to reality and respect the context-sensitivity of judgment.

One such approach to uniting the doxastic-space framework with more realistic preference specifications is to make use of the general approach to reference-dependent preferences developed by Koszegi and Rabin (2006; 2007). Their setup is generally applicable to most formulations of beliefs, and is notable for both its flexibility and its allowance for decision-makers' preferences to be partially influenced by their beliefs. Moreover, Koszegi and Rabin's setup permits not only decision makers' pure preferences to rely on beliefs (as reference points), but also permits the preference-based components of decision weights to be influenced by the belief-based components, which seems an interesting potential progression for the ideas due to Wakker (2004).

Apart from injecting a more solidified role for preferences into the doxastic-space framework, another apparent and imperative next-step for the framework's evolution can be seen in its treatment of hypothesis and evidence selection. At present, the theories of QET and EBSL take hypotheses as essentially exogenous, and do not articulate the process(es) by which evidence becomes selected. Making either or both of these steps semi-endogenous (in at least some circumstances) could greatly help to amplify the power of both QET and EBSL in becoming more generalised models of human cognition.[214] Further (and perchance more importantly), such

[214] Furthermore, such extensions might permit exploring doxastic spaces wherein the relevant dimensions (i.e., considered hypotheses) are not given equal weight to one another, and are somehow 'prioritised' in the thinker's mind.

advancements towards making these components endogenous could assist in allowing QET and EBSL to deliver deeper insights on the specific roles beliefs have in constructing judgment context.

To achieve this end of making hypothesis and evidential selection within the doxastic-space framework more (if only semi-) endogenous, exploration of incorporating similarity judgments (particularly of the asymmetric variety studied by Tversky (1977)) into models like QET and EBSL would seemingly pose a fruitful first action. Moreover, hybridising those models (or others that can be implemented in the doxastic-space framework) with theories of constrained direction of cognitive attention, such as the local-thinking model of Gennaioli and Shleifer (2010) and the sparsity-based method of representing bounded rationality by Gabaix (2014) seem to hold promise.

Finally, a far more ambitious objective for the framework of doxastic spaces is to help the social sciences communicate more fluidly with one another (as well as potentially with the physical sciences). In this vein a near-term aim should be not only to continue the mission of progress for the doxastic-space concept itself as a sandbox for developing models of human cognition, but also to function as a translational tool to convert models and constructs from one discipline into another, and to potentially serve as a 'racetrack' for not only comparison but also competition between ideas on beliefs from across the social sciences which otherwise may not be placed on even footing.

# REFERENCES

**Abdellaoui, M., O. L'Haridon, and C. Paraschiv** (2011) "Experienced vs. described uncertainty: Do we need two prospect theory specifications?", *Management Science*, 57(10):1879-1895.

**Abreu, D., and M. Brunnermeier** (2003) "Bubbles and crashes", *Econometrica*, 71(1):173-204.

**Agranov, M., E. Potamites, A. Schotter, and C. Tergiman** (2012) "Beliefs and endogenous cognitive levels: An experimental study", *Games and Economic Behavior*, 75(2):449-463.

**Akerlof, G., and R. Kranton** (2000) "Economics and identity", *Quarterly Journal of Economics*, 115(3):715-753.

**Allais, M.** (1953) "Le comportement de l'homme rationnel devant le risque: Critique des postulats et axioms de l'ecole americane", *Econometrica*, 21(4):503-546.

**Allen, F., S. Morris, and A. Postlewaite** (1993) "Finite bubbles with short sale constraints and asymmetric information", *Journal of Economic Theory*, 61(2):206-229.

**Anscombe, F., and R. Aumann** (1962) "A definition of subjective probability", *Annals of Mathematical Statistics*, 34(1):199-205.

**Aragones, E., I. Gilboa, A. Postlewaite, and D. Schmeidler** (2005) "Fact-free learning", *American Economic Review*, 95(5):1355-1368.

**Arbib, M., P. Érdi., and J. Szentágothai** (1999) *Neural Organization: Structure, Function, and Dynamics*. London: MIT Press.

**Ariely, D.** (2001) "Seeing sets: Representation by statistical properties" *Psychological Science*, 12(2):157-162.

**Ariely, D., and Z. Carmon** (2000) "Gestalt characteristics of experiences: The defining features of summarized events" *Journal of Behavioral Decision Making*, 13:191-201.

**Ariely, D., G. Loewenstein, and D. Prelec** (2003) "Coherent arbitrariness: Stable demand curves without stable preferences", *Quarterly Journal of Economics*, 118(1):73-105.

**Ariely, D., G. Loewenstein, and D. Prelec** (2006) "Tom Sawyer and the construction of value", *Journal of Economic Behavior & Organization*, 60(1):1-10.

**Asch, S**. (1951) "Effects of group pressure on the modification and distortion of judgments" in H. Guetzkow (ed.) *Groups, Leadership and Men*, Pittsburgh: Carnegie Mellon Press.

**Asch, S.** (1955) "Opinions and social pressure", *Scientific American*, 193:31-35.

**Asch, S**. (1956) "Studies of independence and conformity", *Psychological Monographs*, 70 (9; 416).

**Aumann, R.** (1976) "Agreeing to disagree", *Annals of Statistics*, 4(6):1236-1239.

**Aumann, R.** (1987) "Correlated equilibrium as an expression of Bayesian rationality", *Econometrica*, 55(1):1-18.

**Aumann, R.** (1998) "Common priors: A reply to Gul", *Econometrica*, 66(4):929-938.

**Avery, C., and P. Zemsky** (1998) "Multidimensional uncertainty and herd behavior in financial markets", *American Economic Review*, 88(4):724-748.

**Bacharach, M.** (1985) "Some extensions of a claim of Aumann in an axiomatic model of knowledge", *Journal of Economic Theory*, 37(1):167-190.

**Bachelier, L.** (1900) translated by J. Boness, "Theory of speculation", in P. Cootner (ed.) (1964) *The Random Character of Stock Market Prices*. Cambridge, MA: MIT Press, p. 17-78,

**Backus, D., M. Chernov, and S. Zin** (2014) "Sources of entropy in representative agent models", *Journal of Finance*, 69(1):51-99.

**Banerjee, A.** (1992) "A simple model of herding behavior", *Quarterly Journal of Economics*, 107(3):797-817.

**Barber, B., and T. Odean** (2001) "Boys will be boys: Gender, overconfidence, and common stock investment", *Quarterly Journal of Economics*, 116(1):261-292.

**Barberis, N.** (2013a) "The psychology of tail events: Progress and challenges", *American Economic Review*, 103(3):611-616.

**Barberis, N.** (2013b) "Thirty years of prospect theory in economics: A review and assessment", *Journal of Economic Perspectives*, 27(1):173-196.

**Barberis, N., R. Greenwood, L. Jin, and A. Shleifer** (2014) "X-CAPM: An extrapolative capital asset pricing model", forthcoming, *Journal of Finance*.

**Barberis, N., and M. Huang** (2008) "Stocks as lotteries: The implications of probability weighting for security prices", *American Economic Review*, 98(5):2006-2100.

**Barberis, N., A. Shleifer, and R. Vishny** (1998) "A model of investor sentiment", *Journal of Financial Economics*, 49(3):307-343.

**Barro, R.** (2006) "Rare disasters and asset markets in the twentieth century", *Quarterly Journal of Economics*, 121(3):823-866.

**Barron, G., and I. Erev** (2003) "Small feedback-based decisions and their limited correspondence to description-based decisions", *Journal of Behavioral Decision Making*, 16(3):215-233.

**Barron, G., and G. Ursino** (2013) "Underweighting rare events in experienced based decisions: Beyond sampling error", *Journal of Economic Psychology*, 39:278-286.

**Barron, G., and E. Yechiam** (2009) "The coexistence of overestimation and underweighting of rare events and the contingent recency effect", *Judgment and Decision Making*, 4(6):447-460.

**Barseghyan, L., F. Molinari, T. O'Donoghue, and J. Teitelbaum** (2013a) "Distinguishing probability weighting from risk misperceptions in field data", *American Economic Review*, 103(3):580-585.

**Barseghyan, L., F. Molinari, T. O'Donoghue, and J. Teitelbaum** (2013b) "The nature of risk preferences: Evidence from insurance choices", *American Economic Review*, 103(6):2499-2529.

**Bathelt, H., and J. Glückler** (2003) "Toward a relational economic geography", *Journal of Economic Geography*, 3(2):117-144.

**Bathelt, H., and J. Glückler** (2005) "Resources in economic geography: From substantive concepts towards a relational perspective", *Environment and Planning A*, 37(9):1545-1563.

**Bathelt, H., and J. Glückler** (2011) *The Relational Economy: Geographies of Knowing and Learning*. Oxford: Oxford University Press.

**Battigalli, P., and G. Bonanno** (1999) "Recent results on belief, knowledge and the epistemic foundations of game theory", *Research in Economics*, 53:149-225.

**Beard, J., T. Wallsten, and C. Fox** (2007) "Contrasting stochastic and support theory accounts of subadditivity", *Journal of Mathematical Psychology*, 51(4):229-241.

**Bergin, J., and A. Brandenburger** (1990) "A simple characterization of stochastically monotone functions", *Econometrica*, 58(5):1241-1243.

**Bernheim, B.** (1986) "Axiomatic characterizations of rational choice in strategic environments", *Scandinavian Journal of Economics*, 88(3):473-488.

**Bernheim, B.** (1994) "A theory of conformity", *Journal of Political Economy*, 102(5):841-877.

**Bernheim, B.** (2009) "On the potential of neuroeconomics: A critical (but hopeful) appraisal", *American Economic Journal: Microeconomics*, 1(2):1-41.

**Bernoulli, D.** (1738) "Specimen theoriae novae de mensura sortis", *Commentarri Academiae Scientarum Imperialis Petropolitanae*, Tomus V, p. 175-192.

**Bi, G-Q., and M-M. Poo** (1999) "Distributed synaptic modification in neural networks induced by patterned stimulation", *Nature*, 401:792-796.

**Bikhchandani, S., D. Hirshleifer, and I. Welch** (1992) "A theory of fads, fashion, custom, and cultural change as informational cascades", *Journal of Political Economy*, 100(5):992-1026.

**Bikhchandani, S., D. Hirshleifer, and I. Welch** (1998) "Learning from the behavior of others: Conformity, fads, and informational cascades", *Journal of Economic Perspectives*, 12(3):151-170.

**Billot, A., A. Chateauneuf, I. Gilboa, and J-M. Tallon** (2000) "Sharing beliefs: Between agreeing and disagreeing", *Econometrica*, 68(3):685-694.

**Billot, A., I. Gilboa, D. Samet, and D. Schmeidler** (2005) "Probabilities as similarity-weighted frequencies", *Econometrica*, 73(4):1125-1136.

**Billot, A., I. Gilboa, and D. Schmeidler** (2008) "Axiomatization of an exponential similarity function", *Mathematical Social Sciences*, 55(2):107-115.

**Black, F.** (1986) "Noise", *Journal of Finance*, 41(3):529-543.

**Black, F., and M. Scholes** (1973) "The pricing of options and corporate liabilities", *Journal of Political Economy*, 81(3):637-654.

**Blanco, M., D. Engelmann, and H. Normann** (2008) "A within-subject analysis of other-regarding preferences", *Games and Economic Behavior*, 72(2):321-338.

**Bordalo, P., N. Gennaioli, and A. Shleifer** (2012) "Salience theory of choice under risk", *Quarterly Journal of Economics*, 127(3):1243-1285.

**Bordalo, P., N. Gennaioli, and, A. Shleifer** (2013) "Salience and asset prices", *American Economic Review*, 103(3):623-628.

**Braitenberg, V., and A. Schüz** (1998): *Cortex: Statistics and Geometry of Neuronal Connectivity*. Berlin: Springer.

**Brandenburger, A.** (1992) "Knowledge and equilibrium in games", *Journal of Economic Perspectives*, 6(4):83-101.

**Brandenburger, A., and E. Dekel** (1987) "Common knowledge with probability 1", *Journal of Mathematical Economics*, 16(3):237-245.

**Brandenburger, A., and E. Dekel** (1993) "Hierarchies of beliefs and common knowledge", *Journal of Economic* Theory, 59(1):189-198.

**Bray, M.** (1982) "Learning, estimation, and the stability of rational expectations", *Journal of Economic Theory*, 26(2):318-339.

**Brenner, L.** (2003) "A random support model of the calibration of subjective probabilities", *Organizational Behavior and Human Decision Processes*, 90(1):87-110.

**Brenner, L., D. Griffin, and D. Koehler** (2005) "Modeling patterns of probability calibration with random support theory: Diagnosing case-based judgement", *Organizational Behavior and Human Decision Processes*, 97(1):64-81.

**Brenner, L., D. Koehler, and Y. Rottenstreich** (2002) "Remarks on support theory: Advances and future directions", in T. Gilovich, D., D. Griffin, and D. Kahneman (2002) *Heuristics and Biases: The Psychology of Intuitive Judgment*, p.489-509. Cambridge: Cambridge University Press.

**Brenner, L., and Y. Rottenstreich** (1999) "Focus, repacking, and the judgment of grouped hypotheses", *Journal of Behavioral Decision Making*, 12:141-148.

**Brewer, K., and W. Fellner** (1965) "The slanting of subjective probabilities – agreement on some essentials", *Quarterly Journal of Economics*, 79(4):657-663.

**Brock, W., and C. Hommes** (1998) "Heterogeneous beliefs and routes to chaos in a simple asset pricing model", *Journal of Economic Dynamics and Control*, 22:1235-1274.

**Bruhin, A., H. Fehr-Duda, and T. Epper** (2010) "Risk and rationality: Uncovering heterogeneity in probability distortion", *Econometrica*, 78(4):1375-1412.

**Bruner, J.** (1973) *Beyond the Information Given: Studies in the Psychology of Knowing*. London: George Allen & Unwin.

**Brunnermeier, M., and J. Parker** (2005) "Optimal expectations", *American Economic Review*, 95(4):1092-1118.

**Brunermeier, M., C. Gollier, and J. Parker** (2007) "Optimal beliefs, asset prices, and the preference for skewed returns", *American Economic Review*, 97(2):159-165.

**Budescu, D., I. Erev, and T. Wallsten** (1997) "On the importance of random error in the study of probability judgment. Part I: New theoretical developments", *Journal of Behavioral Decision Making*, 10(3):157-171.

**Budescu, D., T. Wallsten, and W. Au** (1997) "On the importance of random error in the study of probability judgment. Part II: Applying the stochastic judgment model to detect systematic trends", *Journal of Behavioral Decision Making*, 10(3):173-188.

**Cacioppo, J., P. Visser, and C. Pickett** (eds.) (2006) *Social Neuroscience: People Thinking about Thinking People*. Cambridge, MA: MIT Press.

**Camerer, C.** (2007) "Neuroeconomics: Using neuroscience to make economic predictions", *Economic Journal*, 117(519):C26-C24.

**Camerer, C.** (2008) "Neuroeconomics: Opening the gray box", *Neuron*, 60(3):416-419.

**Camerer, C.** (2013) "A review essay about 'Foundations of Neuroeconomic Analysis' by Paul Glimcher", *Journal of Economic Literature*, 51(4):1155-1182.

**Camerer, C., T-H. Ho, and J-K. Chong** (2003) "Models of thinking, learning, and teaching in games", *American Economic Review*, 93(2):192-195.

**Camerer, C., T-H. Ho, and J-K. Chong** (2004) "A cognitive hierarchy model of games", *Quarterly Journal of Economics*, 119(3):861-898.

**Camerer, C., G. Loewenstein, and D. Prelec** (2004) "Neuroeconomics: Why economics needs brains", *Scandinavian Journal of Economics*, 106(3):555-579.

**Camerer, C., G. Loewenstein, and D. Prelec** (2005) "Neuroeconomics: how neuroscience can inform economics", *Journal of Economic Literature*, 43(1):9-64.

**Camerer, C., and M. Weber** (1992) "Recent developments in modeling preferences: Uncertainty and ambiguity", *Journal of Risk and Uncertainty*, 5:325-370.

**Camilleri, A., and B. Newell** (2009) "The role of representation in experience-based choice", *Judgment and Decision Making*, 4(7):518-529.

**Camilleri, A., and B. Newell** (2011a) "When and why rare events are underweighted: A direct comparison of the sampling, partial feedback, full feedback and description choice paradigms", *Psychonomic Bulletin & Review*, 18:377-384.

**Camilleri, A., and B. Newell** (2011b) "Description- and experience-based choice: Does equivalent information equal equivalent choice?", *Acta Psychologica*, 136(3):276-284.

**Camilleri, A., and B. Newell** (2013a) "Mind the gap? Description, experience, and the continuum of uncertainty in risky choice", in V. Pammi, and N. Srinivasan (eds.) *Progress in Brain Research, Volume 202*. Amsterdam: Elsevier.

**Camilleri, A., and B. Newell** (2013b) "The long and the short of it: Closing the description-experience "gap" by taking the long-run view", *Cognition*, 126(1):54-71.

**Campello, M., J. Graham, and C. Harvey** (2010) "The real effects of financial constraints: Evidence from a financial crisis", *Journal of Financial Economics*, 97(3):470-487.

**Cantlon, J., M. Platt, and E. Brannon** (2009) "Beyond the number domain", *Trends in Cognitive Sciences*, 13(2):83-91.

**Cao, H., J. Coval, and D. Hirshleifer** (2002) "Sidelined investors, trading-generated news, and security returns", *Review of Financial Studies*, 15(2):615-648.

**Caplin, A., and M. Dean** (2008) "Dopamine, reward prediction error, and economics", *Quarterly Journal of Economics*, 123(2):663-701.

**Caplin, A., and M. Dean** (2011) "Search, choice, and revealed preference", *Theoretical Economics*, 6(1):19-48.

**Caplin, A., and M. Dean** (2013) "The behavioral implications of rational inattention with Shannon Entropy", working paper, available at: http://www.econ.brown.edu/fac/mark_dean/Working_Paper_10.pdf.

**Caplin, A., M. Dean, P. Glimcher, and R. Rutledge** (2010) "Measuring beliefs and rewards: A neuroeconomic approach", *Quarterly Journal of Economics*, 125(3):923-960.

**Caplin, A., M. Dean, and D. Martin** (2011) "Search and satisficing", *American Economic Review*, 101(7):2899-2922.

**Caplin, A., and J. Leahy** (1994) "Business as usual, market crashes, and wisdom after the fact", *American Economic Review*, 84(3):548-565.

**Caporale, N., and Y. Dan** (2008) "Spike timing-dependent plasticity: A Hebbian learning rule", *Annual Review of Neuroscience*, 31:25-46.

**Cave, J.** (1983) "Learning to agree", *Economics Letters*, 12(2):147-152.

**Chaitin, G.** (2005) *Meta Maths: The Quest for Omega*. London: Atlantic Books.

**Chaitin, G.** (2006) "The limits of reason", *Scientific American*, 294:74-81.

**Chakrabarti, R., and R. Roll** (1999) "Learning from others, reacting, and market quality", *Journal of Financial Markets*, 2:153-178.

**Chamley, C.** (2004) "Delays and equilibria with large and small information in social learning", *European Economic Review*, 48:477-501.

**Charness, G., and U. Gneezy** (2012) "Strong evidence for gender in risk taking", *Journal of Economic Behavior & Organization*, 83(1):50-58.

**Chari, V., and P. Kehoe** (2004) "Financial crises as herds: Overturning the critiques", *Journal of Economic Theory*, 119:128-150.

**Chase, W., and H. Simon** (1973) "Perception in chess", *Cognitive Psychology*, 4(1):55-81.

**Choquet, G.** (1955) "Theory of capacities", *Annales de l'Institut Fourier*, 5:131-295.

**Church, A.** (1936) "An unsolvable problem of elementary number theory", *American Journal of Mathematics*, 58:345-363.

**Clark, G.** (1998) "Stylized facts and close dialogue: Methodology in economic geography", *Annals of the Association of American Geographers*, 88(1):73-87.

**Clark, G.** (2005) "Money flows like mercury: The geography of global finance", *Geografiska Annaler: B, Human Geography*, 87(2):99-112.

**Clark, G.** (2010) "Human nature, the environment, and behaviour: Explaining the scope and geographical scale of financial decision-making", *Geografiska Annaler: Series B, Human Geography*, 92(2):159-173.

**Clark, G.** (2011) "Myopia and the global financial crisis: Context-specific reasoning, market structure, and institutional governance", *Dialogues in Human Geography*, 1(1):4-25.

**Clark, G., A. Dixon, and A. Monk** (eds.) (2009) *Managing Financial Risks: From Global to Local*. Oxford: Oxford University Press.

**Clark, G., M. Feldman, and M. Gertler** (eds.) (2000) *The Oxford Handbook of Economic Geography*. Oxford: Oxford University Press.

**Clark, G., and A. Monk** (2013) "Financial institutions, information, and investing-at-a-distance", *Environment and Planning A*, 45(6):1318-1336.

**Clark, G., K. Strauss, and J. Knox-Hayes** (eds.) (2012) *Saving for Retirement*. Oxford: Oxford University Press.

**Clark, G., and D. Wojcik** (2007) *The Geography of Finance: Corporate Governance in the Global Marketplace*. Oxford: Oxford University Press.

**Cohen, M., and S. Grossberg** (1983) "Absolute stability of global pattern formation and parallel memory storage by competitive neural networks", *IEEE Transactions on Systems, Man, and Cybernetics*, 13(5):815-826.

**Copeland, B., C. Posy, and O. Shagrir** (eds.) (2013) *Computability: Turing, Gödel, Church, and Beyond*. Cambridge, MA: MIT Press.

**Costa-Gomes, M., and V. Crawford** (2006) "Cognition and behavior in two-person guessing games: An experimental study", *American Economic Review*, 96(5):1737-1768.

**Coval, J., and T. Moskowitz** (1999) "Home bias at home: Local equity preference in domestic portfolios", *Journal of Finance*, 54(6):2045-2073.

**Coval, J., and T. Moskowitz** (2001) "The geography of investment: Informed trading and asset prices", *Journal of Political Economy*, 109(4):811-841.

**Crawford, V.** (1995) "Adaptive dynamics in coordination games", *Econometrica*, 63(1):103-143.

**Crawford, V., M. Costa-Gomes, and N. Iriberri** (2013) "Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications", *Journal of Economic Literature*, 51:5-62.

**Crawford, V., and N. Iriberri** (2007) "Level-k auctions: Can a nonequilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions?", *Econometrica,* 75(6):1721-1770.

**Cripps, M., J. Ely, G. Mailath, and L. Samuelson** (2008) "Common learning", *Econometrica,* 76(4):909-933.

**Crutchfield, R.** (1955) "Conformity and character", *American Psychologist,* 10(5):191-198.

**Cutler, D., J. Poterba, and L. Summers** (1993) "What moves stock prices?" in R. Thaler (ed.) *Advances in Behavioral Finance,* New York: Russell Sage.

**Dan, Y., and M-M. Poo** (2006) "Spike timing-dependent plasticity: From synapse to perception", *Physiological Review,* 86:1033-1048.

**Daw, N., and J. O'Doherty** (2014) "Multiple systems for value learning", in P. Glimcher, and E. Fehr (eds.) *Neuroeconomics: Decision Making and the Brain*. London: Academic Press, (p. 393-410).

**Dawes, R., and M. Mulford** (1996) "The false consensus effect and overconfidence: Flaws in judgment or flaws in how we study judgment?", *Organizational Behavior and Human Decision Processes,* 65(3):201-211.

**Dayan, P., and L. Abbott (**2001) *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Cambridge, MA: MIT Press.

**De Giorgi, E., and S. Reimann** (2008) "The α-beauty contest: Choosing numbers, thinking intervals", *Games and Economic Behavior,* 64(2):470-486.

**De Long, B., A. Shleifer, L. Summers, and R. Waldmann** (1990a) "Positive feedback investment strategies and destabilizing rational speculation", *Journal of Finance,*45(2):379-395.

**De Long, B., A. Shleifer, L. Summers, and R. Waldmann** (1990b) "Noise trader risk in financial markets", *Journal of Political Economy*, 98(4):703-738.

**DeCaro, M., R. Thomas, and S. Beilock** (2008) "Individual differences in category learning: Sometimes less working memory capacity is better than more", *Cognition,* 107(1):284-294.

**Dehaene, S. (**1997) *The Number Sense: How the Mind Creates Mathematics*. New York: Oxford University Press.

**Delbaen, F., and W. Schachermayer** (1994) "A general version of the fundamental theorem of asset pricing", *Mathematische Annalen,* 300:463-520.

**DellaVigna, S., and E. Kaplan** (2007) "The Fox News effect: Media bias and voting", *Quarterly Journal of Economics,* 122(3):1187-1234.

**Dempster, A.** (1967) "Upper and lower probabilities induced by a multivalued mapping", *Annals of Mathematical Statistics,* 38:325-339.

**Diamond, D., and R. Verrechia** (1981) "Information aggregation in a noisy rational expectations economy", *Journal of Financial Economics,* 9:221-235.

**Dimson, E., and M. Mussavian** (1998) "A brief history of market efficiency", *European Financial Management*, 4(1):91-103.

**Dominitz, J., and A. Hung** (2009) "Empirical models of discrete choice and belief updating in observational learning experiments", *Journal of Economic Behavior and Organization*, 69(2):94-109.

**Dougherty, M., C. Gettys, and E. Ogden** (1999) "MINERVA-DM: A memory processes model for judgments of likelihood", *Psychological Review*, 106(1):180-209.

**Dougherty, M., and J. Hunter** (2003a) "Hypothesis generation, probability judgment, and individual differences in working memory capacity", *Acta Psychologica*, 113:263-282.

**Dougherty, M., and J. Hunter** (2003b) "Probability judgment and subadditivity: The role of working memory capacity and constraining retrieval", *Memory & Cognition*, 31(6):968-982.

**Duffy, J., and R. Nagel** (1997) "On the robustness of behaviour in experimental 'beauty contest' games", *Economic Journal*, 107(445):1684-1700.

**Eliaz, K., A. Rubinstein** (2014) "A model of boundedly rational "neuro" agents", *Economic Theory*, 57:515-528.

**Ellsberg, D.** (1961) "Risk, ambiguity, and the Savage axioms", *Quarterly Journal of Economics*, 75(4):643-669.

**Epstein, L., and M. Schneider** (2003) "Recursive multiple-priors", *Journal of Economic Theory*, 113(1):1-31.

**Epstein, L., and J. Zhang** (1999) "Least convex capacities", *Economic Theory*, 13:263-286.

**Epstein, L., and S. Zin** (1989) "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework", *Econometrica*,57(4):937-969.

**Erev, I., T. Wallsten, D. Budescu** (1994) "Simultaneous over- and underconfidence: The role of error in judgment processes", *Psychological Review*, 101(3):519-527.

**Ericsson, K., N. Charness, P. Feltovich, and R. Hoffman** (eds.) (2006) *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge: Cambridge University Press.

**Etchart-Vincent, N.** (2004) "Is probability weighting sensitive to the magnitude of consequences? An experimental investigation on losses", *Journal of Risk and Uncertainty*, 28(3):217-235.

**Etchart-Vincent, N**. (2009) "Probability weighting and the 'level' and 'spacing' of outcomes: An experimental study over losses", *Journal of Risk and Uncertainty*, 39:45-63.

**Eyster, E., and M. Rabin** (2005) "Cursed equilibrium", *Econometrica*, 73(5):1623-1672.

**Eyster, E., and M. Rabin** (2010) "Naïve herding in information-rich settings", *American Economic Journal: Microeconomics*, 2(4):221-243.

**Eyster, E., and M. Rabin** (2012) "The limits to imitation in rational observational learning", unpublished paper, available at: http://eml.berkeley.edu/~rabin/wind-up.pdf.

**Fama, E.** (1965) "The behavior of stock market prices", *Journal of Business*, 38:34-105.

**Fama, E.** (1981) "Stock returns, real activity, inflation, and money", *American Economic Review*, 71:545-565.

**Feher, E., and C. Camerer** (2007) "Social neuroeconomics: The neural circuitry of social preferences", *Trends in Cognitive Sciences*, 11(10):419-427.

**Fehr, E., and A. Rangel** (2011) "Neuroeconomic foundations of economic choice: Recent advances", *Journal of Economic Perspectives*, 25(4):3-30.

**Feigenson, L., S. Dehaene, and E. Spelke** (2004) "Core systems of number", *Trends in Cognitive Sciences*, 8(7):307-314.

**Fellner, W.** (1961) "Distortion of subjective probabilities as a reaction to uncertainty", *Quarterly Journal of Economics*, 75(4):670-689.

**Fernandez, R., and D. Rodrik** (1991) "Resistance to reform: Status quo bias in the presence of individual-specific uncertainty", *American Economic Review*, 81(5):1146-1155.

**Fiedler, K.** (2000) "Beware of samples! A cognitive-ecological sampling approach to judgment biases", *Psychological Review*, 107(4):659-676.

**Fiedler, K., and P. Juslin** (eds.) (2006) *Information Sampling and Adaptive Cognition*. Cambridge: Cambridge University Press.

**Fingleton, B.** (ed.) (2007) *New Directions in Economic Geography*. Cheltenham, UK: Edward Elgar.

**Fisher, R.** (1925) "Theory of statistical estimation", *Proceedings of the Cambridge Philosophical Society*, 22:700-725.

**Fox, C., and L. Hadar** (2006) "Decisions from experience = sampling error + prospect theory: Reconsidering Hertwig, Barron, Weber & Erev (2004)", *Judgment and Decision Making*, 1(2):159-161.

**Fox, C., and J. Levav** (2000) "Familiarity bias and belief reversal in relative likelihood judgment", *Organizational Behavior and Human Decision Processes*, 82(2):268-292.

**Fox, C., and D. Tannenbaum** (2011) "The elusive search for stable risk preferences", *Frontiers in Psychology*, 2(298):1-4.

**Fox, C., and A. Tversky** (1995) "Ambiguity aversion and comparative ignorance", *Quarterly Journal of Economics*, 110(3):585-603.

**Fox, C., and A. Tversky** (1998) "A belief-based account of decision under uncertainty", *Management Science*, 44(7):879-895.

**Fox, C., and M. Weber** (2002) "Ambiguity aversion, comparative ignorance, and decision context", *Organizational Behavior and Human Decision Processes*, 88(1):476-498.

**Friedman, M.** (1953) *Essays in Positive Economics*. Chicago: Chicago University Press.

**Friedman, M., and L. Savage** (1948) "The utility analysis of choices involving risk", *Journal of Political Economy*, 56(4):279-304.

**Fujita, M., P. Krugman, and T. Mori** (1999) "On the evolution of hierarchical urban system", *European Economic Review*, 43:209-251.

**Fujita, M., P. Krugman, and A. Venables** (1999) *The Spatial Economy: Cities, Regions, and International Trade*. London: MIT Press.

**Fumagalli, R.** (2010) "The disunity of neuroeconomics: A methodological appraisal", *Journal of Economic Methodology*, 17(2):119-131.

**Gabaix, X.** (2012) "Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance", *Quarterly Journal of Economics*, 127(2):645-700.

**Gabaix, X.** (2014) "A sparsity-based model of bounded rationality", forthcoming, *Quarterly Journal of Economics*.

**Gabaix, X., D. Laibson, G. Moloche, and S. Weinberg** (2006) "Costly information acquisition: Experimental analysis of a boundedly rational model", *American Economic Review*, 96(4):1043-1068.

**Gayer, G., and I. Gilboa** (2014) "Analogies and theories: The role of simplicity and the emergence of norms", *Games and Economic Behavior*, 83:267-283.

**Gale, D.** (1996) "What have we learned from social learning?", *European Economic Review*, 40:617-628.

**Gallistel, C., and R. Gelman** (2000) "Non-verbal numerical cognition: From reals to integers", *Trends in Cognitive Sciences*, 4(2):59-65.

**Gayer, G.** (2010) "Perception of probabilities in situations of risk: A case-based approach", *Games and Economic Behavior*, 68(1):130-143.

**Geanakoplos, J.** (1992) "Common knowledge", *Journal of Economic Perspectives*, 6(4):53-82.

**Geanakoplos, J., and H. Polemarkchakis** (1982) "We can't disagree forever", *Journal of Economic Theory*, 28(1):192-200.

**Gennaioli, N., and S. Shleifer** (2010) "What comes to mind", *Quarterly Journal of Economics*, 125(4):1399-1433.

**Gennaioli, N., A., Shleifer, and R. Vishny** (2012) "Neglected risks, financial innovation, and financial fragility", *Journal of Financial Economics*, 104(3):452-468.

**Gennaioli, N., A. Shleifer, and R. Vishny** (2013) "A model of shadow banking", *Journal of Finance*, 68(4):1331-1363.

**Gentzkow, M., and J. Shapiro** (2010) "What drives media slant? Evidence from U.S. daily newspapers", *Econometrica*, 78(1):35-71.

**Gibbon, J.** (1992) "Ubiquity of scalar timing with a Poisson clock", *Journal of Mathematical Psychology*, 35:283-293.

**Gigerenzer, G., and U. Hoffrage** (1995) "How to improve Bayesian reasoning without instruction: Frequency formats", *Psychological Review*, 102(4):684-704.

**Gigerenzer, G., Z. Swijtink, T. Porter, L. Daston, J. Beatty, and L. Krüger** (1989) *The Empire of Chance: How Probability Changed Science and Everyday Life*. Cambridge: Cambridge University Press.

**Gigerenzer, G., and P. Todd** (eds.) (1999) *Fast and Frugal Heuristics: The Adaptive Toolbox*. New York: Oxford University Press.

**Gilboa, I.** (2009) *Theory of Decision under Uncertainty*. Cambridge: Cambridge University Press.

**Gilboa, I.** (2010) "Questions in decision theory", *Annual Review of Economics*, 2:1-19.

**Gilboa, I., O. Lieberman, and D. Schmeidler** (2006) "Empirical similarity", *Review of Economics and Statistics*, 88(3):433-444.

**Gilboa, I., O. Lieberman, and D. Schmeidler** (2010) "On the definition of objective probabilities by empirical similarity", *Synthese*, 172:79-95.

**Gilboa, I., O. Lieberman, and D. Schmeidler** (2011) "A similarity-based approach to prediction", *Journal of Econometrics*, 162(1):124-131.

**Gilboa, I., A. Postlewaite, L. Samuelson, and D. Schmeidler** (forthcoming) "Economic models as analogies", *Economic Journal*.

**Gilboa, I., L. Samuelson, and D. Schmeidler** (2013) "Dynamics of inductive inference in a unified framework", *Journal of Economic Theory*, 148(4):1399-1432.

**Gilboa, I., and D. Schmeidler** (1989) "Maxmin expected utility with non-unique prior", *Journal of Mathematical Economics*, 18:141-153.

**Gilboa, I., and D. Schmeidler** (1995) "Case-based decision theory", *Quarterly Journal of Economics*, 110(3):605-639.

**Gilboa, I., and D. Schmeidler** (2001) *A Theory of Case-Based Decisions*. Cambridge: Cambridge University Press.

**Gillies, D.** (2000) *Philosophical Theories of Probability*. New York: Routledge.

**Gilovich, T., D. Griffin, and D. Kahneman** (eds.) (2002) *Heuristics and Biases: The Psychology of Intuitive Judgment*. New York: Cambridge University Press.

**Glimcher, P.** (2003) *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics*. Cambridge, MA: MIT Press.

**Glimcher, P.** (2014) "Value-based decision making", in P. Glimcher, and E. Fehr (eds.) *Neuroeconomics: Decision Making and the Brain*. London: Academic Press, (p. 373-392).

**Glimcher, P., M. Dorris, and H. Bayer** (2005) "Physiological utility theory and the neuroeconomics of choice", *Games and Economic Behavior*, 52(2):213-256.

**Glimcher, P., and E. Fehr** (eds.) (2014) *Neuroeconomics: Decision Making and the Brain*. London: Academic Press.

**Glimcher, P., and A. Rustichini** (2004) "Neuroeconomics: The consilience of brain and decision", *Science*, 306:447-452.

**Gobet, F., and G. Clarkson** (2004) "Chunks in expert memory: Evidence for the magical number four…or is it two?" *Memory*, 12(6):732-747.

**Gobet, F., P. Lane, S. Croker, P. Cheng, G. Jones, I. Oliver, and J. Pine** (2001) "Chunking mechanisms in human learning", *Trends in Cognitive Sciences*, 5(6):236-243.

**Gobet, F., and H. Simon** (1996a) "The roles of recognition processes and look-ahead search in time-constrained expert problem solving: Evidence from grand-master-level chess", *Psychological Science*, 7(1):52-55.

**Gobet, F., and H. Simon** (1996b) "Templates in chess memory: A mechanism for recalling several boards", *Cognitive Psychology*, 31(1):1-40.

**Gödel, K.** (1931) "Über formal unentscheidbare sätze der Principia Mathematica un verwandter Syteme, I", *Monatshefte für Mathematik und Physik*, 38:173-178.

**Goeree, J., T. Palfrey, B. Rogers, and R. McKelvey** (2007) "Self-correcting information cascades", *Review of Economic Studies*, 74(3):733-762.

**Goldberg, A.** (2011) "Mapping shared understandings using relational class analysis: The case of the cultural omnivore reexamined", *American Journal of Sociology*, 116(5):1397-1436.

**Goldstein, W., and H. Einhorn** (1987) "Expression theory and the preference reversal phenomenon", *Psychological Review*, 94:236-254.

**Gollier, C.** (2005) "Optimal illusions and decisions under risk", working paper, available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=648602.

**Gonzalez, C., and V. Dutt** (2011) "Instance-based learning: Integrating sampling and repeated decisions from experience", *Psychological Review*, 118(4):523-551.

**Gonzalez, R., and G. Wu** (1999) "On the shape of the probability weighting function", *Cognitive Psychology*, 38(1):129-166.

**Good, I.** (1979) "Studies in the history of probability and statistics. XXXVII A. M. Turing's statistical work in World War II", *Biometrika*, 66(2):393-396.

**Granovetter, M.** (1985) "Economic action and social structure: The problem of embeddedness", *American Journal of Sociology*, 91(3):481-510.

**Granovetter, M.** (2005) "The impact of social structure on economic outcomes", *Journal of Economic Perspectives*, 19(1):33-50.

**Grant, S.** (1995) "Subjective probability without monotonicity: Or how Machina's mom may also be probabilistically sophisticated", *Econometrica*, 63(1):159-189.

**Greenwood, R., and S. Nagel** (2009) "Inexperienced investors and bubbles", *Journal of Financial Economics*, 93(2):239-258.

**Greenwood, R., and. A. Shleifer** (2014) "Expectations of returns and expected returns", *Review of Financial Studies*, 27(3):714-746.

**Griffin, D., and A. Tversky** (1992) "The weighing of evidence and the determinants of confidence", *Cognitive Psychology*, 24(3):411-435.

**Grossman, S.** (1976) "On the efficiency of competitive stock markets where trades have diverse information", *Journal of Finance*, 31(2):573-585.

**Grossman, S., and J. Stiglitz** (1976) "Information and competitive price systems", *American Economic Review*, 66(2):246-253.

**Grossman, S., and J. Stiglitz** (1980) "On the impossibility of informationally efficient markets", *American Economic Review*, 70(3):393-408.

**Gul, F.** (1998) "A comment on Aumann's Bayesian view", *Econometrica*, 66(4):923-927.

**Gul, F., and R. Lundholm** (1995) "Endogenous timing and the clustering of agents' decisions", *Journal of Political Economy*, 103(5):1039-1066.

**Gul, F., and W. Pesendorfer** (2008) "The case for mindless economics", in A. Caplin, and A. Schotter (eds.) *The Foundations of Positive and Normative Economics: A Handbook*. New York: Oxford University Press, p. 3-41.

**Gul, F., W. Pesendorfer, and T. Strzalecki** (2014) "Coarse competitive equilibrium and extreme prices", working paper, available at: http://www.econ.upf.edu/docs/seminars/strzalecki.pdf.

**Güney, Ş., and R. Newll** (2014) "Overcoming ambiguity aversion through experience", *Journal of Behavioural Decision Making*, early-availability online edition.

**Hadar, L., and C. Fox** (2009) "Information asymmetry in decision from description versus decision from experience", *Judgment and Decision Making*, 4(4):317-325.

**Hadjichristidis, C., S. Sloman, and E. Wisniewski** (2001) "Judging the probability of representative and unrepresentative unpackings", in J. Moore (ed.) *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*, p. 376-380. Mahwah, NJ: Erlbaum.

**Haile, P., A. Hortaçsu, and G. Kosenok** (2008) "On the empirical content of quantal response equilibrium", *American Economic Review*, 98(1):180-200.

**Hájek, A.** (2007) "The reference class problem is your problem too", *Synthese*, 156:563-585.

**Halpern, J., and Y. Moses** (1990) "Knowledge and common knowledge in a distributed environment", *Journal of the ACM*, 37(3):549-587.

**Handa, J.** (1977) "Risk, probabilities, and a new theory of cardinal utility", *Journal of Political Economy*, 85(1):97-122.

**Hanna, R., S. Mulainathan, and J. Schwartzstein** (2014) "Learning through noticing: Theory and evidence from a field experiment", *Quarterly Journal of Economics*, 129(3):1311-1353.

**Hansen, L., and T. Sargent** (2001) "Robust control and model uncertainty", *American Economic Review*, 91(2):60-66.

**Hansen, L., and T. Sargent** (2008) *Robustness*. Princeton: Princeton University Press.

**Hanson, R.** (1998) "Consensus by identifying extremists", *Theory and Decision*, 44(3):293-301.

**Hanson, R.** (2002) "Disagreement is unpredictable", *Economics Letters*, 77(3):365-369.

**Harless, D.** (1993) "Experimental tests of prospective reference theory", *Economics Letters*, 43(1):71-76.

**Harris, M., and A. Raviv** (1993) "Differences of opinion make a horse race", *Review of Financial Studies*, 6(3):473-506.

**Harrison, G.** (2008) "Neuroeconomics: A critical reconsideration", *Economics and Philosophy*, 24(3):303-344.

**Harrison, J., and D. Kreps** (1978) "Speculative investor behavior in a stock market with heterogeneous expectations", *Quarterly Journal of Economics*, 92(2):323-336.

**Harrison, J., and D. Kreps** (1979) "Martingales and arbitrage in multiperiod securities markets", *Journal of Economic Theory*, 20(3):381-408.

**Harrison, J., and S. Pliska** (1981) "Martingales and stochastic integrals in the theory of continuous trading", *Stochastic Processes and their Applications*, 11(3):215-260.

**Harrison, G., and D. Ross** (2010) "The methodologies of neuroeconomics", *Journal of Economic Methodology*, 17(2):185-196.

**Harsanyi, J.** (1967) "Games with incomplete information played by "Bayesian" players. Part I: The basic model", *Management Science*, 14:159-182.

**Harsanyi, J.** (1968a) "Games with incomplete information played by "Bayesian" players. Part II: Bayesian equilibrium points", *Management Science*, 14:320-334.

**Harsanyi, J.** (1968b) "Games with incomplete information played by "Bayesian" players. Part III: The basic probability distribution of the game", *Management Science*, 14:486-502.

**Haug, E., and N. Taleb** (2011) "Option traders use (very) sophisticated heuristics, never the Black-Scholes-Merton formula", *Journal of Economic Behavior & Organization*, 77(2):97-106.

**Hayek, F.** (1945) "The use of knowledge in society", *American Economic Review*, 35(4):519-530.

**Hayek, F.** (1952) *The Sensory Order: An Inquiry into the Foundations of Theoretical Psychology*. Chicago: University of Chicago Press.

**Heath, C., and A. Tversky** (1991) "Preference and belief: Ambiguity and competence in choice under uncertainty", *Journal of Risk and Uncertainty*, 4:5-28.

**Heifetz, A.** (1996) "Comments on consensus without common knowledge", *Journal of Economic Theory*, 70(1):273-277.

**Hertwig, R., G. Barron, E. Weber, and I. Erev** (2004) "Decisions from experience and the effect of rare events in risky choice", *Psychological Science*, 15(8):534-539.

**Hertwig, R., and I. Erev** (2009) "The description-experience gap in risky choice", *Trends in Cognitive Sciences*, 13(12):517-523.

**Hertwig, R., and T. Pleskac** (2010) "Decisions from experience: Why small samples?", *Cognition*, 115(2):225-237.

**Hilgetag, C., M. Müller-Linow, M-T. Hütt** (2011) "Dynamics of hierarchical neural networks", in R. Wang and F. Gu (eds.) *Advances in Cognitive Neurodynamics (II)*, p. 215-219.

**Hills, T., and R. Hertwig** (2010) "Information search in decisions from experience: Do our patterns of sampling foreshadow our decisions?", *Psychological Science*, 21(12):1787-1792.

**Hirshleifer, D., and S. Teoh** (2003) "Herding behaviour and cascading in capital markets: A review and synthesis", *European Financial Management*, 9(1):25-66.

**Ho, T-H., C. Camerer, and J-K. Chong** (2007) "Self-tuning experience weighted attraction learning in games", *Journal of Economic Theory*, 133(1):177-198.

**Ho, T-H., C. Camerer, and K. Weigelt** (1998) "Iterated dominance and iterated best response in experimental "p-beauty contests"", *American Economic Review*, 88(4):947-969.

**Hoch, S.** (1987) "Perceived consensus and predictive accuracy: The pros and cons of projection", *Journal of Personality and Social Psychology*, 53(2):221-234.

**Hong, H., and J. Stein** (2003) "Differences of opinion, short-sale constraints, and market crashes", *Review of Financial Studies*, 16(2):487-525.

**Hong, H., and J. Stein** (2007) "Disagreement and the stock market", *Journal of Economic Perspectives*, 21(2):109-128.

**Huberman, G.** (2001) "Familiarity breeds investment", *Review of Financial Studies*, 14(3):659-680.

**Idson, L., D. Krantz, D. Osherson, and N. Bonini** (2001) "The relation between probability and evidence judgment: An extension of Support Theory", *Journal of Risk and Uncertainty*, 22(3):227-249.

**Jehiel, P.** (2005) "Analogy-based expectation equilibrium", *Journal of Economic Theory*, 123:81-104.

**Jehiel, P., and F. Koessler** (2008) "Revisiting games of incomplete information with analogy-based expectations", *Games and Economic Behavior*, 62:533-557.

**Jones, E., and R. Nisbett** (1972) "The actor and the observer: Divergent perceptions of the causes of behavior", in E. Jones, D. Kanouse, H. Kelley, R. Nisbett, S. Valins, and B. Weiner (eds.) *Attribution: Perceiving the Causes of Behavior*, p. 79-94, Morristown, NJ: General Learning Press.

**Jordan, J.** (1982) "A dynamic model of expectations equilibrium", *Journal of Economic Theory*, 28(2):235-254.

**Kable, J., and P. Glimcher** (2009) "The neurobiology of decision: Consensus and controversy", *Neuron*, 63(6):733-745.

**Kahneman, D., J. Knetsch, and R. Thaler** (1991) "Anomalies: The endowment effect, loss aversion, and status quo bias", *Journal of Economic Perspectives*, 5(1):193-206.

**Kahneman, D., and D. Miller** (1986) "Norm theory: Comparing reality to its alternatives", *Psychological Review*, 93(2):136-153.

**Kahneman, D., P. Slovic, and A. Tversky** (eds.) (1982) *Judgment under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.

**Kahneman, D., and A. Tversky** (1972) "Subjective probability: A judgment of representativeness", *Cognitive Psychology*, 3(3):430-454.

**Kahneman, D., and A. Tversky** (1974) "Judgment under uncertainty: Heuristics and biases", *Science*, 185:1124-1131.

**Kahneman, D., and A. Tversky** (1979) "Prospect theory: An analysis of decision under risk", *Econometrica*, 47(2):263-291.

**Kahneman, D., and A. Tversky** (2000) *Choices, Values, and Frames*. Cambridge: Cambridge University Press.

**Kandel, E.** (2006) *In Search of Memory: The Emergence of a New Science of Mind*. New York: W.W. Norton & Company.

**Kandel, E., and N. Pearson** (1995) "Differential interpretation of public signals and trade in speculative markets", *Journal of Political Economy*, 103(4):831-872.

**Kandel, E., J. Schwartz, and T. Jessell** (eds.) (2000) *Principles of Neural Science*. New York: McGraw-Hill.

**Kareev, Y.** (2000) "Seven (indeed, plus or minus two) and the detection of correlations", *Psychological Review*, 107(2):397-402.

**Karmarker, U.** (1978) "Subjectively weighted utility: A descriptive extension of the expected utility model", *Organizational Behavior and Human Performance*, 21(1):61-72.

**Karni, E.** (2010) "Red herrings: Some thoughts on the meaning of zero-probability events and mathematical modeling", *Economics Letters*, 107(2):134-135.

**Keynes, J.** (1921) *Treatise on Probability*. London: Macmillan & Co.

**Keynes, J.** (1936) *The General Theory of Employment, Interest and Money*. London: Macmillan.

**Kilka, M., and M. Weber** (2001) "What determines the shape of the probability weighting function?", *Management Science*, 47(12):1712-1726.

**Kirman, A.** (1992) "Whom or what does the representative individual represent?", *Journal of Economic Perspectives*, 6(2):117-136.

**Knight, F.** (1921) *Risk, Uncertainty and Profit*. New York: Houghton Mifflin.

**Koch, C.** (1999) *Biophysics of Computation: Information Processing in Single Neurons*. Oxford: Oxford University Press.

**Koehler, D., C. White, and R. Grondin** (2003) "An evidential support accumulation model of subjective probability", *Cognitive Psychology*, 46:152-197.

**Koszegi, B., and M. Rabin** (2006) "A model of reference-dependent preferences", *Quarterly Journal of Economics*, 121(4):1133-1165.

**Koszegi, B., and M. Rabin** (2007) "Reference-dependent risk attitudes", *American Economic Review*, 97(4):1047-1073.

**Krasucki, P.** (1996) "Protocols forcing consensus", *Journal of Economic Theory*, 70(1):266-272.

**Kreps, D.** (1988) *Notes on the Theory of Choice*. Boulder, CO: Westview Press.

**Krugman, P.** (1994) "Complex landscapes in economic geography", *American Economic Review*, 84(2):412-416.

**Krugman, P.** (1996) *The Self-Organizing Economy*. Oxford: Blackwell.

**Krugman, P.** (1998a) "What's new about the new economic geography?", *Oxford Review of Economic Policy*, 14(2):7-17.

**Krugman, P.** (1998b) "Space: The final frontier", *Journal of Economic Perspectives*, 12(2):161-174.

**Kullback, S.** (1959) *Information Theory and Statistics*. New York: John Wiley & Sons.

**Kurz, M.** (1994a) "On rational belief equilibria", *Economic Theory*, 4:859-876.

**Kurz, M.** (1994b) "On the structure and diversity of rational beliefs", *Economic Theory*, 4:877-900.

**Kurz, M., H. Jin, and M. Motolese** (2005) "Endogenous fluctuations and the role of monetary policy", in A. Frydman, J. Stiglitz, and M. Woodford (eds.) *Knowledge, Information and Expectations in Modern Macroeconomics: Essays in Honor of Edmund S. Phelps*. p. 188-227. Princeton: Princeton University Press.

**Kyburg, H., and H. Smokler** (eds.) (1964) *Studies in Subjective Probability*. New York: John Wiley & Sons.

**Lee, I.** (1993) "On the convergence of informational cascades", *Journal of Economic Theory*, 61:395-411.

**Lee, I.** (1998) "Market crashes and informational avalanches", *Review of Economic Studies*, 65:741-759.

**Lee, R., D. Greene, and P. House** (1977) "The "false consensus effect": An egocentric bias in social perception and attribution processes", *Journal of Experimental Social Psychology*, 13(3):279-301.

**Lejarraga, T., R. Hertwig, and C. Gonzalez** (2012) "How choice ecology influences search in decisions from experience", *Cognition*, 124(3):334-342.

**Leslie, A., R. Gelman, and C. Gallistel** (2008) "The generative basis of natural number concepts", *Trends in Cognitive Sciences*, 12(6):213-218.

**Li, C-Y., J-T. Lu, C-P. Wu, S-M. Duan, and M-M. Poo** (2004) "Bidirectional modification of presynaptic neuronal excitability accompanying spike timing-dependent synaptic plasticity", *Neuron*, 41:257-268.

**Lintner, J.** (1965a) "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets", *Review of Economics and Statistics*, 47:13-37.

**Lintner, J.** (1965b) "Securities prices, risk, and maximal gains from diversification", *Journal of Finance*, 20(4):587-615.

**Lo, A., and A. MacKinley** (1999) *A Non-Random Walk Down Wall Street*. Princeton: Princeton University Press.

**Loewenstein, G., S. Rick, and J. Cohen** (2008) "Neuroeconomics", *Annual Review of Psychology*, 59:647-672.

**Lubenov, E., and A. Siapas** (2008) "Decoupling through synchrony in neuronal circuits with propagation delays", *Neuron*, 58:118-131.

**Lucas, R.** (1972) "Expectations and the neutrality of money", *Journal of Economic Theory*, 4(2):103-124.

**Lucas, R.** (1975) "An equilibrium model of the business cycle", *Journal of Political Economy*, 83:1113-1144.

**Luce, R.** (1986) *Response Times: Their Role in Inferring Elementary Mental Organization*. Oxford: Oxford University Press.

**Luce, R.** (2003) "Whatever happened to information theory in psychology?", *Review of General Psychology*, 7(2):183-188.

**Luce, R., and P. Fishburn** (1991) "Rank- and sign-dependent linear utility models for finite first-order gambles", *Journal of Risk and Uncertainty*, 4:29-59.

**Luria, A.** (1973) *The Working Brain: An Introduction to Neuropsychology*. New York: Basic Books.

**Maass, W., T. Natschläger, and H. Markram** (2002) "Real-time computing without stable states: A new framework for neural computation based on perturbations", *Neural Computation*, 14(11):2531-2560.

**Macchi, L., D. Osherson, and D. Krantz** (1999) "A note on superadditive probability judgment", *Psychological Review*, 106(1):210-214.

**Machina, M., and D. Schmeidler** (1992) "A more robust definition of subjective probability", *Econometrica*, 60(4):745-780.

**Machina, M., and D. Schmeidler** (1995) "Bayes without Bernoulli: Simple conditions for probabilistically sophisticated choice", *Journal of Economic Theory*, 67(1):106-128.

**MacKenzie, D.** (2006) *An Engine, Not a Camera: How Financial Models Shape Markets*. Cambridge, MA: MIT Press.

**Magoun, R.** (1958) *The Waking Brain.* Springfield, IL: Charles C. Thomas Publishing.

**Malmendier, U., and S. Nagel** (2011) "Depression babies: Do macroeconomic experiences affect risk-taking?", *Quarterly Journal of Economics*, 126(1):373-416.

**Malmendier, U., and S. Nagel** (forthcoming) "Learning from inflation experiences", *Quarterly Journal of Economics*.

**Mandelbrot, B., and R. Hudson** (2004) *The (Mis)Behavior of Markets: A Fractal View of Financial Turbulence*, New York: Basic Books.

**Mandler, G.** (1985) *Cognitive Psychology: An Essay in Cognitive Science*. Hillsdale, NJ: Lawrence Earlbaum Associates.

**Mankiw, N., and R. Reis** (2002) "Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve", *Quarterly Journal of Economics*, 117(4):1295-1328.

**Mankiw, N., R. Reis, and J. Wolfers** (2003) "Disagreement about inflation expectations", NBER Working Paper no. 9796, available at: http://www.nber.org/papers/w9796.

**Manski, C.** (2004) "Measuring expectations", *Econometrica*, 72(5):1329-1376.

**Markowitz, H.** (1952a) "Portfolio selection", *Journal of Finance*, 7(1):77-91.

**Markowitz, H.** (1952b) "The utility of wealth", *Journal of Political Economy*, 60(2):151-158.

**Markowitz, H.** (1959) *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley & Sons.

**Markowitz, H.** (1991) "Foundations of portfolio theory", *Journal of Finance*, 46(2):469-477.

**Markowitz, H.** (1999) "The early history of portfolio theory: 1600-1960", *Financial Analysts Journal*, 55(4):5-16.

**Marks, G., and N. Miller** (1987) "Ten years of research on the false-consensus effect: An empirical and theoretical review", *Psychological Bulletin*, 102(1):72-90.

**Martin, J.** (1999a) "Entropic measures of belief system constraint", *Social Science Research*, 28(1):111-134.

**Martin, J.** (2002) "Power, authority, and the constraint of belief systems", *American Journal of Sociology*, 107(4):861-904.

**Martin, R.** (1999b) "The new 'geographical turn' in economics: Some critical reflections", *Cambridge Journal of Economics*, 23:65-91.

**Martin, R., and P. Sunley** (1996) "Paul Krugman's geographical economics and its implications for regional development theory: A critical assessment", *Economic Geography*, 72(3):259-292.

**Martin, R., and P. Sunley** (2007) "Complexity thinking and evolutionary economic geography", *Journal of Economic Geography*, 7(5):573-601.

**Masatlioglu, Y., D. Nakajima, and E. Ozbay** (2012) "Revealed attention", *American Economic Review*, 102(5):2183-2205.

**Masatlioglu, Y., and E. Ok** (2005) "Rational choice with status quo bias", *Journal of Economic Theory*, 121(1):1-29.

**McDowell, L.** (2004) "Masculinity, identity and labour market change: Some reflections on the implications of thinking relationally about difference and the politics of inclusion", *Geografiska Annaler: Series B, Human Geography*, 86(1):45-56.

**McDowell, L.** (2006) "Reconfigurations of gender and class Relations: Class differences, class condescension and the changing place of class relations", *Antipode*, 38(4):825-850.

**McFadden, D.** (1999) "Rationality for economists?", *Journal of Risk and Uncertainty*, 19(1-3):73-105.

**McKelvey, R., and T. Palfrey** (1995) "Quantal response equilibrium for normal form games", *Games and Economic Behavior*, 10(1):6-38.

**McKelvey, R., and T. Palfrey** (1998) "Quantal response equilibrium for extensive form games", *Experimental Economics*, 1:9-41.

**McKelvey, R., and T. Page** (1986) "Common knowledge, consensus, and aggregate information", *Econometrica*, 54(1):109-127.

**Ménager, L.** (2006) "Consensus, communication and knowledge: An extension with Bayesian agents", *Mathematical Social Sciences*, 51(3):274-279.

**Ménager, L.** (2008) "Consensus and common knowledge of an aggregate of decisions", *Games and Economic Behavior*, 62(2):722-731.

**Mertens, J-F., and S. Zamir** (1985) "Formulation of Bayesian analysis for games with incomplete information", *International Journal of Game Theory*, 14(1):1-29.

**Merton, R.** (1973) "The theory of rational option pricing", *Bell Journal of Economics and Management Science*, 4:141-183.

**Milgrom, P.** (1981) "An axiomatic characterization of common knowledge", *Econometrica*, 49(1):219-222.

**Milgrom, P., and N. Stokey** (1982) "Information, trade and common knowledge", *Journal of Economic Theory*, 26(1):17-27.

**Miller, G.** (1956) "The magical number seven, plus or minus two: Some limits on our capacity for processing information", *Psychological Review*, 63(2):81-97.

**Miller, G., and W. Madow** (1963) "On the maximum likelihood estimate of the Shannon-Wiener measure of information" in R. Luce, R. Bush, and E. Galanter (eds.) *Readings in Mathematical Psychology, Volume 1*, p. 448-469, New York: John Wiley & Sons.

**Minsky, M.** (1986) *The Society of Mind*. New York: Simon & Schuster.

**Minksy, M., and S. Papert** (1969) *Perceptrons: An Introduction to Computational Geography*. Cambridge, MA: MIT Press.

**Monderer, D., and D. Samet** (1989) "Approximating common knowledge with common beliefs", *Games and Economic Behavior*, 1(2):170-190.

**Morris, S.** (1994) "Trade with heterogeneous prior beliefs and asymmetric information", *Econometrica*, 62(6):1327-1347.

**Morris, S.** (1995) "The common priors assumption in economic theory", *Economics and Philosophy*, 11(2):227-253.

**Morris, S.** (1999) "Approximate common knowledge revisited", *International Journal of Game Theory*, 28:385-408.

**Mossin, J.** (1966) "Equilibrium in a capital asset market", *Econometrica*, 34(4):768-783.

**Mosteller, F., and P. Nogee** (1951) "An experimental measurement of utility", *Journal of Political Economy*, 59:371-404.

**Moyer, R., and T. Landauer** (1967) "Time required for judgments of numerical inequality", *Nature*, 215:1519-1520.

**Mullen, B., J. Atkins, D. Champion, C. Edwards, D. Hardy, J. Story, and M. Vanderklok** (1985) "The false consensus effect: A meta-analysis of 115 hypothesis tests", *Journal of Experimental Social Psychology*, 21(3):262-283.

**Mullainathan, S.** (2002) "A memory-based model of bounded rationality", *Quarterly Journal of Economics*, 117(3):735-774.

**Mullainathan, S., J. Schwartzstein, and A. Shleifer** (2008) "Coarse thinking and Persuasion", *Quarterly Journal of Economics*, 123(2):577-619.

**Mullen, B., J. Atkins, D. Champion, C. Edwards, D. Hardy, J. Story, and M. Vanderklok** (1985) "The false consensus effect: A meta-analysis of 115 hypothesis tests", *Journal of Experimental Social Psychology*, 21:262-283.

**Muth, J.** (1961) "Rational expectations and the theory of price movements", *Econometrica*, 29(3):315-335.

**Nagel, R.** (1995) "Unraveling in guessing games: An experimental study", *American Economic Review*, 85(5):1313-1326.

**Nagel, S.** (2012) "Evaporating liquidity", *Review of Financial Studies*, 25(7):2005-2039.

**Nagel, R., A. Bosch-Domènech, A. Satorra, and J. García-Montalvo** (2002) "One, two, (three), infinity: Newspaper and lab beauty-contest experiments", *American Economic Review*, 92(5):1687-1701.

**Neeman, Z.** (1996a) "Approximating agreeing to disagree results with common p-beliefs", *Games and Economic Behavior*, 12(1):162-164.

**Neeman, Z.** (1996b) "Common beliefs and the existence of speculative trade", *Games and Economic Behavior*, 16(1):77-96.

**Newell, B., T. Rakow, N. Weston, and D. Shanks** (2004) "Search strategies in decision making: The success of "success"", *Journal of Behavioral Decision Making*, 17(2):117-137.

**Nielsen, L.** (1995) "Common knowledge of a multivariate aggregate statistic", *International Economic Review*, 36(1):207-216.

**Nielsen, L., A. Brandenburger, J. Geanakoplos, R. McKelvey, and T. Page** (1990) "Common knowledge of an aggregate of expectations", *Econometrica*, 58(5):1235-1239.

**Nistbett, R., and Z. Kunda** (1985) "Perception of social distributions", *Journal of Personality and Social Psychology*,48(2):297-311.

**Nyarko, Y., and A. Schotter** (2002) "An experimental study of belief learning using elicited beliefs", *Econometrica*, 70(3):971-1005.

**Ottaviani, M., and P. Sørenson** (2000) "Herd behavior and investment: Comment", *American Economic Review*, 90(3):695-704.

**Padoa-Schioppa, C.** (2008) "The syllogism of neuro-economics", *Economics and Philosophy*, 24(3):449-457.

**Page, S.** (2011) *Diversity and Complexity*, Princeton: Princeton University Press.

**Parikh, R., and P. Krasucki** (1990) "Communication, consensus, and knowledge", *Journal of Economic Theory*, 52(1):178-189.

**Pástor, L., and P. Veronesi** (2012) "Uncertainty about government policy and stock prices", *Journal of Finance*, 67(4):1219-1264.

**Pettit, P.** (2002) *Rules, Reasons, and Norms*. Oxford: Clarendon Press.

**Piazessi, M., and M. Schneider** (2010) "Interest rate risk in credit markets", *American Economic Review*, 100(2):579-584.

**Pearlman, J., and T. Sargent** (2005) "Knowing the forecasts of others", *Review of Economic Dynamics*, 8(2):480-497.

**Pickel, A.** (2012) "Between Homo Sociologicus and Homo Biologicus: The reflexive self in the age of social neuroscience", *Science and Education*, 21:1507-1526.

**Poldrack, R.** (2006) "Can cognitive processes be inferred from neuroimaging data?", *Trends in Cognitive Sciences*, 10(2):59-63.

**Poldrack, R., and A. Wagner** (2004) "What can neuroimaging tell us about the mind? Insights from the prefrontal cortex", *Current Directions in Psychological Science*, 13(5):177-181.

**Popper, K.** (1935) *Logik der Forschung*, Vienna: Verlag.

**Prelec, D.** (1998) "The probability weighting function", *Econometrica*, 66(3):497-527.

**Preston, M., and P. Baratta** (1948) "An experimental study of the auction-value of an uncertain outcome", *American Journal of Psychology*, 61(2):183-193.

**Purves, D., E. Brannon, R. Cabeza, S. Huettel, K. LaBar, M. Platt, and M. Woldorff** (2008) *Principles of Cognitive Neuroscience.* Sunderland, MA: Sinauer Associates.

**Pykett, J.** (2013) "Neurocapitalism and the new neuros: Using neuroeconomics, behavioural economics and picoeconomics for public policy", *Journal of Economic Geography*, 13:845-869.

**Quiggin, J.** (1982) "A theory of anticipated utility", *Journal of Economic Behavior & Organization*, 3(4):323-343.

**Rabin, M.** (2013a) "Incorporating limited rationality into economics", *Journal of Economic Literature*, 51(2):528-543.

**Rabin, M.** (2013b) "An approach to incorporating psychology into economics", *American Economic Review*, 103(3):617-622.

**Rabin, M., and J. Schrag** (1999) "First impressions matter: A model of confirmatory bias", *Quarterly Journal of Economics*, 114(1):37-82.

**Rabinovich, M., P. Varona, A. Selverston, and H. Abarbanel** (2006) "Dynamical principles in neuroscience", *Reviews of Modern Physics*, 78(4):1213-1265.

**Radner, R.** (1972) "Existence of equilibrium of plans, prices, and price expectations in a sequence of markets", *Econometrica*, 40(2):289-303.

**Radner, R.** (1979) "Rational expectations equilibrium: General existence and the information revealed by prices, *Econometrica*, 47(3):655-678.

**Rakow, T., K. Demes, and B. Newell** (2008) "Biased samples not mode of presentation: Re-examining the apparent underweighting of rare events in experience-based choice", *Organizational Behavior and Human Decision Processes*, 106(2):168-179.

**Rakow, T., and B. Newell** (2010) "Degrees of uncertainty: An overview and framework for future research on experience-based choice", *Journal of Behavioral Decision Making*¸23(1):1-14.

**Ramsey, F.** (1926) "Truth and probability", in D. Mellor (ed.) (1990) *Philosophical Papers*, Cambridge: Cambridge University Press, p. 52-104.

**Ramsey, F.** (1929) "Knowledge", in D. Mellor (ed.) (1990) *Philosophical Papers*, Cambridge: Cambridge University Press, p. 110-111.

**Ramsey, J.** (2010) "Six problems for causal inference from fMRI", *NeurogImage*, 49(2):1545-1558.

**Riley, M., K. Shockley, and G. Van Orden** (2012) "Learning from the body about the mind", *Topics in Cognitive Science*, 4(1):21-34.

**Rodrik, D.** (2013) "Who needs the nation state?", *Economic Geography*, 89(1):1-19.

**Rogers, B., T. Palfrey, and C. Camerer** (2009) "Heterogeneous quantal response equilibrium and cognitive hierarchies", *Journal of Economic Theory*, 144(4):1440-1467.

**Roll, R.** (1988) "$R^2$", *Journal of Finance*, 43:541-566.

**Romer, D.** (1993) "Rational asset-price movements without news", *American Economic Review*, 83(5):1112-1130.

**Rook, D.** (2012) "How can we know if investors are coherently linking sustainability concepts?", *Journal of Sustainable Finance and Investment*, 2(3-4):198-221.

**Rook, D.** (2014a) "Cognitive limitations on belief approximations: A theory of evidential quantization", working paper, available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2485419.

**Rook, D.** (2014b) "Entropy-based social learning", working paper, available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2509909.

**Ross, D.** (2008) "Two styles of neuroeconomics", *Economics and Philosophy*, 24(3):473-483.

**Rottenstreich, Y., and A. Tversky** (1997) "Unpacking, repacking, and anchoring: Advances in support theory", *Psychological Review*, 104(2):406-415.

**Roy, A.** (1952) "Safety first and the holding of assets", *Econometrica*, 20(3):431-449.

**Rubinstein, A.** (1988a) *Modeling Bounded Rationality*. Cambridge, MA: MIT Press.

**Rubinstein, A.** (1988b) "Similarity and decision-making under risk (is there a utility theory resolution to the Allais paradox?)", *Journal of Economic Theory*, 46(1):145-153.

**Rubinstein, A.** (1989) "The electronic mail game: Strategic behavior under "almost common knowledge"", *American Economic Review*, 79(3):385-391.

**Rubinstein, A.** (2007) "Instinctive and cognitive reasoning: A study of response times", *Economic Journal*, 117:883-896.

**Rubinstein, A.** (2013) "Response time and decision making: A "free" experimental study", *Judgment and Decision Making*, 8:540-551.

**Rubinstein, A., and Y. Salant** (2006) "A model of choice from lists", *Theoretical Economics*, 1(1):3-17.

**Rubinstein, A., and A. Wolinsky** (1990) "On the logic of "agreeing to disagree" type results", *Journal of Economic Theory*, 51(1):184-193.

**Rubinstein, M.** (2006) *A History of the Theory of Investments: My Annotated Bibliography*. Hoboken, NJ: John Wiley & Sons.

**Rutledge, R., M. Dean, A. Caplin, and P. Glimcher** (2010) "Testing the reward prediction error hypothesis with an axiomatic model", *Journal of Neuroscience*, 30(40):13525-13536.

**Samet, D.** (1990) "Ignoring ignorance and agreeing to disagree", *Journal of Economic Theory*, 52(1):190-207.

**Samet, D.** (2010) "Agreeing to disagree: The non-probabilistic case", *Games and Economic Behavior*, 69(1):169-174.

**Samuelson, P.** (1938) "A note on the pure theory of consumer's behaviour", *Economica*, 5(17):61-71.

**Samuelson, P.** (1965) "Proof that properly anticipated prices fluctuate randomly", *Industrial Management Review*, 6(2):41-49.

**Samuelson, W., and R. Zeckhauser** (1988) "Status quo bias in decision making", *Journal of Risk and Uncertainty*, 1:7-59.

**Sanfey, A., G. Loewenstein, S. McClure, and J. Cohen** (2006) "Neuroeconomics: Cross-currents in research on decision-making", *Trends in Cognitive Sciences*, 10(3):108-116.

**Sapolsky, R.** (2014) "Heights and lengths and areas of rectangles", *Edge*, available at: http://www.edge.org/response-detail/25536 (accessed 18 March 2014).

**Savage, L.** (1954) *The Foundations of Statistics*. New York: John Wiley & Sons.

**Savage, L.** (1964) "The elicitation of personal probabilities" in H. Kyburg and H. Smokler, *Studies in Subjective Probability*, New York: John Wiley & Sons.

**Schacter, D., K. Norman, and W. Koustaal** (1998) "The cognitive neuroscience of constructive memory", *Annual Reviews of Psychology*, 49:289-318.

**Scharfstein, D., and J. Stein** (1990) "Herd behavior and investment", *American Economic Review*, 80(3):465-479.

**Schelling, T.** (1978) *Micromotives and Macrobehavior*. London: W.W. Norton & Co.

**Schemaker, P.** (1982) "The expected utility model: Its variants, purposes, evidence and limitations", *Journal of Economic Literature*, 20(2):529-563.

**Schmeidler, D.** (1989) "Subjective probability and expected utility without additivity", *Econometrica*, 57(3):571-587.

**Schwanen, T., and M-P. Kwan** (2012) "Critical space-time geographies", *Environment and Planning A*, 44(9):2043-2048.

**Schwanen, T., and D. Wang** (2014) "Well-being, context and everyday activities in space and time", *Annals of the American Association of Geographers*, 104(4):833-851.

**Sedlmeier, P., and G. Gigerenzer** (1997) "Intuitions about sample size: The empirical law of large numbers", *Journal of Behavioral Decision Making*, 10(1):33-51.

**Segal, U.** (1987) "The Ellsberg paradox and risk aversion: An anticipated utility approach", *International Economic Review*, 28(1):175-202.

**Segal, U.** (1989) "Anticipated utility: A measure representation approach", *Annals of Operations Research*, 19:359-373.

**Selten, R., and R. Stoecker** (1986) "End behavior in sequences of finite Prisoner's Dilemma supergames: A learning theory approach", *Journal of Economic Behavior & Organization*, 7(1):47-70.

**Shackle, G.** (1952) *Expectation in Economics*. Cambridge: Cambridge University Press.

**Shafer, G.** (1976) *A Mathematical Theory of Evidence*. Princeton: Princeton University Press.

**Shannon, C.** (1948) "A mathematical theory of communication", *Bell System Technical Journal*, 27(3):379-423.

**Sharp, W.** (1964) "Capital asset prices: A theory of market equilibrium under conditions of risk", *Journal of Finance*, 19(3):425-442.

**Sheppard, E.** (2000) "Geography or economics? Conceptions of space, time, interdependence, and agency", in G. Clark, M. Feldman, and M. Gertler (eds.) *The Oxford Handbook of Economic Geography*. Oxford: Oxford University Press, (p. 99-119).

**Shiller, R.** (1982) "The use of volatility measures in assessing market efficiency", *Journal of Finance*, 36(2):291-304.

**Shiller, R.** (1990) "Investor behavior in the October 1987 stock market crash: Survey evidence", in R. Shiller, *Market Volatility*, Cambridge, MA: MIT Press.

**Shiller, R.** (1995) "Conversation, information, and herd behavior", *American Economic Review*, 85(2):181-185.

**Shiller, R.** (2011) "The neuroeconomics revolution", available at: http://www.project-syndicate.org/commentary/the-neuroeconomics-revolution (accessed: 18 March 2014).

**Shiller, R., F. Kon-Ya, and Y. Tsutsui** (1991) "Investor behavior in the October 1987 stock market crash: The case of Japan", *Journal of the Japanese and International Economies*, 5(1):1-13.

**Shleifer, A.** (2012) "Psychologists at the gate: A review of Daniel Kahneman's *Thinking, Fast and Slow*", *Journal of Economic Literature*, 50(4):1-12.

**Shleifer, A., and L. Summers** (1990) "The noise trader approach to finance", *Journal of Economic Perspectives*, 4(2):19-33.

**Shleifer, A., and R. Vishny** (1997) "The limits of arbitrage", *Journal of Finance*, 52(1):35-55.

**Simon, H.** (1955) "A behavioral model of rational choice", *Quarterly Journal of Economics*, 69(1):99-118.

**Simon, H.** (1956) "Rational choice and the structure of the environment", *Psychological Review*, 63(2):129-138.

**Simon, H.** (1973) "The organization of complex systems", in H. Pattee (ed.) *Hierarchy Theory: The Challenge of Complex Systems*. New York: Braziller, (p. 1-27).

**Simon, H.** (1974) "How big is a chunk?", *Science*, 8:482-488.

**Simon, H.** (1979) *Models of Thought*. New Haven: Yale University Press.

**Simon, H.** (1990) "Invariants of human behavior", *Annual Review of Psychology*, 41:1-19.

**Simon, H.** (1996) *The Sciences of the Artificial*. Cambridge, MA: MIT Press.

**Simon, H., and A. Ando** (1961) "Aggregation of variables in dynamic systems", *Econometrica*, 29:111-138.

**Simon, H., and W. Chase** (1973) "Skill in chess: Experiments with chess-playing tasks and computer simulation of skilled performance throw light on some human perceptual and memory processes", *American Scientist*, 61(4):394-403.

**Simon, H., and K. Gilmartin** (1973) "A simulation of memory for chess positions", *Cognitive Psychology*, 5(1):29-46.

**Sims, C.** (2003) "Implications of rational inattention", *Journal of Monetary Economics*, 50(3):665-690.

**Sims, C.** (2010) "Rational inattention and monetary economics", in B. Friedman, and M. Woodford (eds.) *Handbook of Monetary Economics*, p. 155-180, San Diego: North Holland.

**Sloman, S., Y. Rottenstreich, E. Wisniewski, C. Hadjichristidis, and C. Fox** (2004) "Typical versus atypical unpacking and superadditive probability judgment", *Journal of Experimental Psychology*, 30(3):573-582.

**Smith, V.** (2008) *Rationality in Economics: Constructivist and Ecological Forms*. Cambridge: Cambridge University Press.

**Smith, L., and P. Sørenson** (2000) "Pathological outcomes of observational learning", *Econometrica*, 68(2):371-398.

**Snowberg, E., and J. Wolfers** (2010) "Explaining the favorite-longshot bias: Is it risk-love, or misperceptions?", *Journal of Political Economy*, 118(4):723-746.

**Sobel, J.** (2009) "Neuroeconomics: A comment on Bernheim", *American Economic Journal: Microeconomics*, 1(2):60-67.

**Soros, G.** (2013) "Fallibility, reflexivity, and the human uncertainty principle", *Journal of Economic Methodology*, 20(4):309-329.

**Sonsino, D.** (1995) ""Impossibility of speculation" theorems with noisy information", *Games and Economic Behavior*, 8:406-423.

**Squire, L., D. Berg, F. Bloom, S. du Lac, A. Ghosh, N. Spitzer** (2013) *Fundamental Neuroscience*. London: Academic Press.

**Squire, L., and E. Kandel** (2009) *Memory: From Mind to Molecules*: Greenwood Village, CO: Roberts & Company.

**Stahl, D., and P. Wilson** (1994) "Experimental evidence on players' models of other players", *Journal of Economic Behavior and Organizations*, 25(3):309-327.

**Stahl, D., and P. Wilson** (1995) "On players' models of other players: Theory and experimental evidence", *Games and Economic Behavior*, 10(1):218-254.

**Stephen, D., and G. Van Orden** (2012) "Searching for general principles in cognitive performance: Reply to commentators", *Topics in Cognitive Science*, 4(1)94-102.

**Stewart, N.** (2009) "Decision by sampling: The role of the decision environment in risky choice", *Quarterly Journal of Experimental Psychology*, 62(6):1041-1062.

**Stewart, N., N. Chater, and G. Brown** (2006) "Decision by sampling", *Cognitive Psychology*, 53(1):1-26.

**Stewart, N., S. Reimers, and A. Harris** (2014) "On the origin of utility, weighting, and discounting functions: How they get their shapes and how to change their shapes", *Management Science*, advanced online publication.

**Stiglitz, J.** (1961) "The economics of information", *Journal of Political Economy*, 69(3):213-225.

**Stiglitz, J.** (2000) "The contributions of the economics of information to twentieth century economics", *Quarterly Journal of Economics*, 115(4):1441-1478.

**Storper, M., and A. Venables** (2004) "Buss: Face-to-face contact and the urban economy", *Journal of Economic Geography*, 4(4):351-370.

**Strauss, K.** (2008) "Re-engaging with rationality in economic geography: Behavioural approaches and the importance of context in decision-making", *Journal of Economic Geography*, 8(2):137-156.

**Suppes, P.** (1974) "The measurement of belief", *Journal of the Royal Statistical Society, B*, 36:160-191.

**Sydnor, J.** (2010) "(Over)insuring modest risks", *American Economic Journal: Applied Economics*, 2(4):177-199.

**Szentágothai, J.** (1993) "Self-organization: The basic principle of neural functions", *Theoretical Medicine*, 14:101-116.

**Taleb, N.** (2007) *The Black Swan: The Impact of the Highly Improbable*. London: Allen Lane.

**Taleb, N.** (2009) "Errors, robustness, and the fourth quadrant", *International Journal of Forecasting*, 25(4):744-759.

**Tetlock, P.** (2005) *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton: Princeton University Press.

**Thaler, R.** (ed.) (1993) *Advances in Behavioral Finance*. New York: Russell Sage Foundation.

**Thaler, R.** (2000) "From homo economicus to homo sapiens", *Journal of Economic Perspectives*, 14(1):133-141.

**Thaler, R.** (ed.) (2005) *Advances in Behavioral Finance, Volume II.* Princeton: Princeton University Press.

**Timmerman, A.** (1993) "How learning in financial markets generates excess volatility and predictability in stock prices", *Quarterly Journal of Economics*, 108(4):1135-1145.

**Tirole, J.** (1982) "On the possibility of speculation under rational expectations", *Econometrica*, 50(5):1163-1181.

**Townsend, R.** (1983) "Forecasting the forecasts of others", *Journal of Political Economy*, 91(4):546-588.

**Trappenberg, T.** (2002) *Fundamentals of Computational Neuroscience*. Oxford: Oxford University Press.

**Treynor, J.** (1961) "Market value, time, and risk", unpublished manuscript.

**Treynor, J.** (1962) "Toward a theory of market value of risky assets", unpublished manuscript.

**Tulving, E., and W. Donaldson** (eds.) (1972) *Organization of Memory*. New York: Academic Press.

**Turing, A.** (1936) "On computable numbers, with an application to the Entsheidungsproblem", *Proceedings of the London Mathematical Society*, 2(42): 230-265.

**Tversky, A.** (1977) "Features of similarity", *Psychological Review*, 84:327-352.

**Tversky, A., and D. Kahneman** (1971) "Belief in the law of small numbers", *Psychological Bulletin*, 76(2):105-110.

**Tversky, A., and D. Kahneman** (1973) "Availability: A heuristic for judging frequency and probability", *Cognitive Psychology*, 5(2):207-232.

**Tversky, A., and D. Kahneman** (1992) "Advances in prospect theory: Cumulative representation of uncertainty", *Journal of Risk and Uncertainty*, 5:297-323.

**Tversky, A., and D. Koehler** (1994) "Support theory: A nonextensional representation of subjective probability", *Psychological Review*, 101(4):547-567.

**Tversky, A., and P. Wakker** (1995) "Risk attitudes and decision weights", *Econometrica*, 63(6):1255-1280.

**Ungemach, C., N. Chater, and N. Stewart** (2009) "Are probabilities overweighted or underweighted when rare outcomes are experienced (rarely)?", *Psychological Science*, 20(4):473-479.

**Ungemach, C., N. Stewart, and S. Reimers** (2011) "How incidental values from the environment affect decisions about money, risk, and delay", *Psychological Science*, 22(2):253-260.

**Van Huyck, J., R. Battalio, and R. Beil** (1990) "Tacit coordination games, strategic uncertainty, and coordination failure", *American Economic Review*, 80(1):234-248.

**Van Huyck, J., R. Battalio, and R. Beil** (1991) "Strategic uncertainty, equilibrium selection, and coordination failure in average opinion games", *Quarterly Journal of Economics*, 106(3):885-910.

**van Rooij, M., and J. Holden** (2013) "Correspondence: Are cognitive functions localizable? Colin Camerer et al. versus Marieke van Rooij and John. G. Holden", *Journal of Economic Perspectives*, 27(2):247-252.

**van Rooij, M., and G. Van Orden** (2011) "It's about space, it's about time, Neuroeconomics and the brain sublime", *Journal of Economic Perspectives*, 25(4):31-55.

**Vassilakis, S., and S. Zamir** (1993) "Common belief and common knowledge", *Journal of Mathematical Economics*, 22(5): 495-505.

**Varian, H.** (1985) "Divergence of opinion in complete markets: A note", *Journal of Finance*, 40(1):309-317.

**Veronesi, P.** (2004) "The Peso problem hypothesis and stock market returns", *Journal of Economic Dynamics and Control*, 28(4):707-725.

**Viscusi, K.** (1989) "Prospective reference theory: Toward an explanation of the paradoxes", *Journal of Risk and Uncertainty*, 2:235-264.

**Vives, X.** (1995) "The speed of information revelation in a financial market mechanism", *Journal of Economic Theory*, 67:178-204.

**von Neuman, J. and O. Morgenstern** (1944) *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.

**von Neuman, J. and O. Morgenstern** (1947) *Theory of Games and Economic Behavior – 2nd, Revised Edition*. Princeton: Princeton University Press.

**Wachter, J.** (2013) "Can time-varying risk of rare disasters explain aggregate stock market volatility?", *Journal of Finance*, 68(3):987-1035.

**Wagner, J., and S. Johnson** (2011) "An association between understanding cardinality and analog magnitude representations in preschoolers", *Cognition*, 119:10-22.

**Wakker, P.** (2004) "On the composition of risk preference and belief", *Psychological Review*, 111(1):236-241.

**Wakker, P., and A. Tversky** (1993) "An axiomatization of cumulative prospect theory", *Journal of Risk and Uncertainty*, 7(7):147-176.

**Walley, P.** (1991) *Statistical Reasoning with Imprecise Probabilities*. London: Chapman & Hall.

**Walliser, B.** (2008) *Cognitive Economics*. Berlin: Springer-Verlag.

**Wallsten, T., and C. Gonzalez-Vallejo** (1994) "Statement verification: A stochastic model of judgment and response", *Psychological Review*, 101(3):490-504.

**Weitzman, M.** (2000) "Economic profitability versus ecological entropy", *Quarterly Journal of Economics*, 115(1):237-263.

**Weitzman, M.** (2007) "Subjective expectations and asset-return puzzles", *American Economic Review*, 97(4):1102-1130.

**Weizsäcker, G.** (2010) "Do we follow others when we should? A simple test of rational expectations", *American Economic Review*, 100(5):2340-2360.

**Whalen, J., C. Gallistel, and R. Gelman** (1999) "Nonverbal counting in humans: The psychophysics of number representation", *Psychological Science*, 10(2):130-137.

**Wieland, V., and M. Wolters** (2011) "The diversity of forecasts from macroeconomic models of the US economy", *Economic Theory*, 47:247-292.

**Wolfram, S.** (2002) *A New Kind of Science*. Champaign, IL: Wolfram Media.

**Wu, G., and R. Gonzalez** (1996) "Curvature of the probability weighting function", *Management Science*, 42(12):1676-1690.

**Wu, G., and R. Gonzalez** (1999) "Nonlinear decision weights in choice under uncertainty", *Management Science*, 45(1):74-85.

**Yaari, M.** (1987) "The dual theory of choice under risk", *Econometrica*, 55(1):95-115.

**Zhang, L., and M-M. Poo** (2001) "Electrical activity and development of neural circuits", *Nature Neuroscience*, 4:1207-1214.

**Zimmer, C.** (2014) "Secrets of the brain", *National Geographic*, 225(2):28-57.

**Zwiebel, J.** (1995) "Corporate conservatism and relative compensation", *Journal of Political Economy*, 103(1):1-25.