

Extremes in Heterogeneous Economic Data

Julian Winkler

Trinity 2024

Extremes in Heterogeneous Economic Data

Julian Winkler

University of Oxford, Department of Economics

Thesis submitted for the degree of Doctor of Philosophy

Large economic surprises occur with greater frequency than previously assumed. Popular examples include the severity of the 2008 financial crisis, the wages paid to top employees, and the damage from climate change. Their frequencies are growing more obvious, not just thanks to the passing of time, but also the greater availability of large scale, granular data. On one hand, these datasets improve the statistical power of traditional, econometric models developed for small sample sizes. On the other hand, an obsession with central moments, or ‘average effects’, obscures the role of higher moments in both theory as well as statistical forecasting. This thesis studies this systematic bias in economic models for the distribution of economic quantities.

The role of granular data is increasingly important for economic research. The first part of this thesis studies the impact of extreme values in two applications. The first application documents the impact of social contagion on stock returns using text data scraped from the WallStreetBets online forum. The sentiments of users who attract thousands of comments and follow-up conversations generate an outsized impact on stock prices relative to average sentiments, due to the heavy-tailed nature of viral social media content. The second application documents the heavy tails in the productivity of European firms. This study implies that a nontrivial number of firms generate millions of euros per employee in value added, while productivity for the typical firm remains in line with national averages, at around tens of thousands of euros. This dispersion is difficult to reconcile with the usual assumption in econometric studies that asserts a finite second moment for labour productivity at the firm level.

I formalise the relationship between the size of economic variables and their frequency in the second half of this thesis, by introducing entropy as a key measure for variety. Entropy is a useful measure of variety when constructing aggregates from heterogeneous data, such as economic output. In one example, the degree to which output is more or less concentrated is consequential to what economic growth truly represents. In finance, heterogeneity plays a crucial role in diversification. Banks decide on the maturity of their investments, and mutual funds decide on what mandate to construct a portfolio on. Entropy regulates the frequency of wealth allocated to different bonds or stocks and thus governs interest rates and stock prices both in and out of equilibrium.

Acknowledgements

I am indebted to Rick Van der Ploeg and J. Doyne Farmer for their guidance and support. The value of their time is not lost on me, so their generosity in spending it on our countless meetings leaves me humbled.

I am grateful to Valentina Semenova for her valuable co-authorship and collegiality. It is not possible to have someone better to share a desk with.

The various chapters included in this thesis received crucial comments from a list of people too long to fit on one page. However, I will give special attention to a few who helped me more than I could have ever expected, in no particular order: François Lafond, Jangho Yang, Torsten Heinrich, Farshad Ravasan, Kieran Marray, Oren Sussman, Ian Goldin, Xavier Gabaix, Andrei Shleifer, Erwin Diewert, Steven Durlauf, Francis DiTraglia, Kevin Sheppard, Piers August, and Matthias Winkler.

Over the years, this thesis developed out of many interactions with scholars at the Institute for New Economic Thinking at the Oxford Martin School and the Department of Economics at the University of Oxford. I particularly benefited from my many presentations, good and bad, at the department's Macroeconomics Workshop led by Michael McMahon and Federica Romei. I thank Xavier Gabaix, and the economics department at Harvard University, for hosting me as a visiting student in Fall 2022.

I thank Baillie Gifford and the Institute for New Economic Thinking at the Oxford Martin School for funding my stay at the University of Oxford, as well as The David Walton Memorial Fund for awarding me the David Walton Distinguished Doctoral Student Scholarship. Parts of this thesis are part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 870245.

For Valentina

Contents

1	Introduction	1
1.1	Granular patterns in stock price and firm data	2
1.2	Distributions in economics	3
1.3	Overview	6
2	Social Contagion and asset prices:	9
2.1	Introduction	10
2.2	What is <i>WallStreetBets</i> ?	14
2.3	Social dynamics on WSB	16
2.3.1	Identifying peer influence: Frequent Posters	18
2.3.2	Identifying peer influence – Commenter Network	23
2.3.3	Further insights	27
2.4	Has WSB destabilised markets?	28
2.4.1	Evidence of trading	29
2.4.2	Evidence of granularity	31
2.4.3	Evidence of bubbles	37
2.5	Conclusion	40
	Appendices	45
2.A	Data appendix	45
2.A.1	Extended description of WSB	45
2.A.2	Tickers mentioned on WSB	48
2.A.3	Sentiment modelling in WSB posts	49
2.A.4	Market variables	50
2.A.5	How prevalent are hype traders?	51
2.A.6	Preliminary Data Exploration	52
2.B	Model for social dynamics and asset prices	56
2.B.1	Static model with investment complementarities	57
2.B.2	Equilibrium price dynamics with peer effects	58
2.B.3	Persistent fluctuations	61
2.B.4	Modeling predictions	61
2.B.5	Discussion on Assumptions 2.1, 2.2	62
2.B.6	Additional modelling frameworks: bubbles with peer effects	62
2.C	Extra results for peer effects	64

2.C.1	Target independent variable	64
2.C.2	First stage regression estimates	65
2.C.3	Normalisation procedure and non-normalised coefficient estimates	65
2.C.4	Evidence of identification strategy – Frequent Posters	66
2.C.5	Evidence of identification strategy – Commenter Network	70
2.D	Additional results for market impact	71
2.D.1	GIV	71
2.D.2	Additional evidence of price impact	73
3	Is social media a game-changer for financial stability?	79
3.1	Introduction	80
3.2	Model setup	81
3.3	Results	84
3.4	Discussion	90
3.5	Methods	90
3.5.1	Finding the steady states	90
3.5.2	Estimating Parameters from WSB	91
	Appendices	97
3.A	Model appendix	97
3.A.1	Deriving aggregate buying intensity and market impact	97
3.A.2	Stability Regions	99
3.B	Hype investor sentiment spillovers	100
3.C	Hype investor interest	101
3.C.1	Results	102
4	The heavy tails in labour productivity	105
4.1	Introduction	106
4.2	Data and descriptive statistics	109
4.2.1	Data sources	109
4.2.2	Construction of variables	109
4.2.3	Characteristic patterns of the distributions	111
4.2.4	Testing for power law tails and infinite variance	112
4.3	Models and estimation	119
4.3.1	Models	119
4.3.2	Distribution fit and model comparison	120
4.4	Measuring productivity dispersion	121
4.5	Discussion and conclusion	124
	Appendices	131
4.A	Data appendix	131
4.A.1	Raw data	131
4.A.2	Data cleaning	131

4.A.3	Deflation	132
4.A.4	Size proportion	134
4.B	The Lévy alpha-stable distribution	134
4.B.1	Characteristics of the Lévy alpha-stable distribution	134
4.C	Fitting procedure	136
4.C.1	Fitting methods	136
4.C.2	Estimated parameters	138
4.D	Reference model and comparison of the fit	139
4.D.1	AST and AEP: models, fitting procedure, and estimated parameters	139
4.E	Derivation of the scaling of the sample standard deviation with sample size	140
5	Accounting for the slowdown in labour productivity	143
5.1	Introduction	144
5.2	Accounting for the slowdown	145
5.2.1	Contributions of inputs growth and TFP	145
5.2.2	Industry-specific contributions and structural change	149
	Appendices	155
5.A	Slowdown and missing GDP	155
5.B	Contributions of TFP and capital deepening using OECD Productivity database	157
5.C	Contributions of industries and reallocation using OECD's STAN	157
6	Accounting for variety	159
6.1	Introduction	160
6.2	Indexing variety	162
6.2.1	Entropy to measure variety	163
6.2.2	Intermezzo: what is entropy?	165
6.2.3	Theoretical motivation	168
6.2.4	Choosing weights	170
6.2.5	Summary	172
6.3	Indices, revisited	172
6.3.1	Real output: Hulten's Paradox in KLEMS data	172
6.3.2	Labour productivity: technology or allocation?	176
6.4	Conclusion	178
	Appendices	181
6.A	Proofs	181
6.A.1	Proposition 6.2	181
6.A.2	Proposition 6.3	182
6.B	Additional results	182
6.B.1	Does the level of aggregation matter?	182

7	Volatility in forward rates	185
7.1	Introduction	186
7.2	Data	188
7.2.1	Overview of forward rates and futures	188
7.2.2	Forward Rate Curve	189
7.2.3	Fluctuations along the FRC	190
7.2.4	Level, slope and curvature	192
7.3	Model	192
7.3.1	Notation and setup	192
7.3.2	Dynamics of the FRC	195
7.3.3	FRC equilibrium	198
7.3.4	Extreme fluctuations	202
7.3.5	Model predictions	202
7.4	Results	203
7.4.1	Explaining FRC fluctuations	203
7.4.2	Gamma pattern in spreads	206
7.4.3	Maximal volatility	206
7.5	Conclusion	208
	Appendices	211
7.A	Data appendix	211
7.B	Model appendix	211
7.B.1	Structure of the model's banking system	211
7.B.2	Lemma 7.1	213
7.B.3	Proposition 7.1	214
7.B.4	Corollary 7.1.1	216
7.C	Daily fluctuations in future prices	216
8	Changing preferences as a source of stock return variation	219
8.1	Introduction	220
8.2	Modelling a portfolio's evolution	223
8.2.1	Model with one factor	225
8.2.2	Demand as a robustness criterion	228
8.2.3	Dynamic asset demand	233
8.2.4	Discussion	234
8.3	Decomposing price fluctuations	236
8.3.1	Data	236
8.3.2	Empirical framework	238
8.3.3	How do portfolios grow?	242
8.3.4	Aggregating stock returns	245
8.3.5	Discussion and limitations	249
8.4	Conclusion	249

Appendices	253
8.A Proofs	253
8.A.1 Proof to Proposition 8.1	253
8.A.2 Proof to Proposition 8.2	253
8.B Data appendix	255
8.B.1 CRSP stock variables	255
8.B.2 Compustat variables	255
9 Lessons and future direction	257
9.1 Making sense of distributions	257
9.2 Insights into extremes	257
9.3 Final notes	259

Chapter 1

Introduction

Scientists have calculated that the chances of something so patently absurd actually existing are millions to one.

But magicians have calculated that million-to-one chances crop up nine times out of ten.

–Terry Pratchett, *Mort*

People love a sensational headline. Beyond dramatisation, to what extent are economic outcomes actually dictated by large events? Qualitatively distinct events, or extreme values, in economic data pose statistical challenges that accompany today's growing piles of data. One particular type of extremes, namely heavy-tailed phenomena, is actually surprisingly well 'behaved', in the sense that their frequency is in line with what statistical models predict. In practice, these extremes increasingly highlight an awkward gap in economics between theoretical models and empirical fact: the largest observations are often not outliers in the meaningful sense, because they are in line with how we expect heavy-tailed phenomena to appear. Insightful examples range from income distributions, stock market volatility, industrial organisation, to climate change.

The relationship between the frequency and the size of a variable constitutes its *distribution*. Economic research typically relegates the study of distribution to a handful of summary statistics, notably the variance of a quantity. These are not always necessary or sufficient to describe a given distribution. For example, the mean and variance of a variable uniquely reconstruct a Gaussian distribution, which fails to represent other patterns prominent in economic quantities. This leads to a major blind spot on heavy-tailed fluctuations in economics, both in terms of statistical predictions as well as modelling.

This shortcoming is increasingly highlighted with access to larger datasets that highlight the stark disparities in economic quantities among agents. The first part of this thesis highlights two such examples, namely stock price fluctuations and the concentrations of labour productivity. The second part of this thesis builds a framework with which to model distributions for economic quantities.

Many areas of economic research increasingly take such features more seriously, for instance in the adoption of Heterogeneous Agent New Keynesian (HANK) models in macroeconomics. This thesis takes issue with the ad hoc justifications in these explanations for

what are, ultimately, common patterns in economics. Common patterns can emerge from different processes by definition. The alternative view pursued in this thesis is motivated by a perspective borrowed from statistics; the fashion by which the frequency of a variable scales with its quantity depends on features of the aggregate. For example, the equilibrium price of a valuable stock depends largely on the availability of alternative valuable stocks.

1.1 Granular patterns in stock price and firm data

This thesis is motivated by the role of extremes and heterogeneity in economic quantities. It offers two empirical studies that demonstrate the importance of dispersion in economic aggregates.

Role of aggregation Activity on stock markets relies on a diversity of information held by participants ([Grossman & Stiglitz 1980](#)). The emergence of online investment forums that promote the coordination of strategies among retail investors presumably reduces this diversity, and potentially leads to unstable prices.

Chapter 2 constructs a novel dataset from one such forum, namely WallStreetBets. Each day, thousands of users publish submissions in which they can detail their sentiments about the expected performance of a stock. Others are able to reply and comment on these submissions, to develop larger discussions. We track user activity over time, as well as their interactions with other users, to empirically test the extent to which sentiments are formed as a function of their peers' expressed sentiments, despite their own personal experience.

This is an example where the aggregation of information is more than a simple sum of users' beliefs. The transmission of beliefs by a few prominent users via discussions on social media introduces an extra multiplicative factor that can impact overall sentiments beyond those of users who would have acted in a social void.

Of particular interest are 'viral' conversations, which receive orders of magnitude more attention from users relative to typical submissions. Using the method of [Gabaix & Koijen \(2023\)](#), these 'granular' sentiments have an outsized impact on stock returns. We find more intense discourse on WallStreetBets during bubble-like stock price run-ups, which we distinguish from non-bubble-like stocks using the approach of [Greenwood et al. \(2019\)](#).

Aggregation plays a large role in macroeconomics, both empirically and theoretically ([Olley & Pakes 1996](#), [Baqaee & Farhi 2020](#)). Important benchmarks, such as labour productivity, are reported across a variety of industries, firms and even products. Endogenous growth theories have largely embraced the importance of variety underlying aggregate economic outcomes since [Schumpeter \(1943\)](#). This idea highlights the tension between efficiency – the productivity of a frontier firm – versus heterogeneity – a body of firms constituting 'laggards'. They must co-exist: innovations diffuse to laggards from the frontier, while the frontier is compelled to innovate further and retain its lead.

Chapter 5, an excerpt from [Goldin et al. \(2024\)](#), investigates the extent to which the general slowdown in labour productivity growth can be attributed to an even slowdown across all industries, versus a re-allocation of economic activity out of more productive industries,

into less productive ones. Re-allocation does not appear to be a primary explanation for the slowdown in most cases. This is largely because significant structural changes develop over decades, a point developed further in Chapter 6.

Consensus formation The role for social contagion in asset markets dates back to [Shiller \(1984\)](#). Due to the popularity of WallStreetBets, as well as its demonstrated impact on stock prices, our data is uniquely well-suited for an empirical assessment of patterns in price data that are subject to social contagion. Between Chapters 2 and 3, we use a behavioural finance model to predict the temporal behaviour of investor sentiment and stock prices. The version in Chapter 2 considers an equilibrium in which prices are set to meet asset demand in each period. In Chapter 3, we investigate an interesting case in which investors base their decision on the previous price return, as they seek to extrapolate a trend.

All reach the same conclusion: prices are expected to cycle or oscillate at the time scale with which investors decide to buy or sell an asset. The dynamic system in Chapter 3 offers more insight in this regard, as the model predicts a bifurcation once social contagion dominates the decision-making process of investors. These investors alternate between two fixed points, either buy or sell, when i) an exogenous shock pushes them off a current steady state, or ii) an endogenous overshooting of the steady state due to their impact on price bounces the system between the fixed points.

Heavy tails in firm productivity Chapter 4 is an excerpt from [Yang et al. \(2024\)](#), in which we document the remarkable ubiquity of the shape with which labour productivity is distributed at the firm level. In particular, this distribution is closely fitted by a Lévy alpha-stable distribution across many countries, years, and industries. The standout features are the heavy left and right tails of the distributions, which imply that firms regularly boast value added per employee in the millions of euros. This holds for the ‘frontier’ at the top of the distribution, but also for those that reported losses, in the negative part of the distribution.

This paper spurs a discussion on the appropriate metrics for dispersion. This is because the dispersion of a Lévy alpha-stable distribution is regulated by two parameters; a scale parameter, analogous to the standard deviation in a Gaussian distribution, and a tail parameter, which separately measures the frequency of extreme outcomes. Research on misallocation in economics has not emphasised this important distinction. Common measures, such as the standard deviation, are outright inappropriate because their sample estimates vary with the sample size, and fail to converge.

1.2 Distributions in economics

The second part of this thesis offers a theoretical framework to make sense of these patterns. It focuses on the tension between aggregate quantities acting as forces on economic decisions, against the inertia imposed by the vast variety of heterogeneous types. The idea

fundamental to this thesis is that the manner with which this tension resolves – how quantities *scale* across varieties – dictates the common shapes and patterns in economic data.

Role of heterogeneity Many economic variables demonstrate remarkable variety, for example in goods, assets, skills etc. This manifests in the vast arrays of *types* that constitute product, financial, and labour markets, among many others.

In the face of this variety, economic research often focuses on key aggregates. For example, traded products are aggregated into quantity indices, which help with the important problem of evaluating welfare or policy outcomes. This aggregation is usually interpreted to occur from the perspective of a *representative agent*, who is tasked with a choice between alternative types: products to consume, people to employ, and stocks to invest in.

Chapter 6 introduces entropy as a measure for such variety in types. It focuses on the problem of price indexation, which is used to track the evolution of living standards over time. A useful measure for living standards is the monetary amount of goods traded in the economy, which has been recorded for roughly a century in most advanced economies. However, these monetary transactions not only increase when more goods are exchanged – which is a meaningful quantity for assessing the welfare of citizens – but also when these goods command higher prices – which has no implications for welfare. Crucially, a change in entropy reflects shifts in the make-up of economic activity, and is shown to introduce a substantial bias in quantity indices.

Entropy in economics Chapters 7 and 8 use entropy as a measure to study the shape of distributions. There are different justifications for the use of entropy from an economic perspective, which vary with the application. In the case of indexing output in Chapter 6, I rely on a decomposition and the statistical properties of entropy.

Chapters 7 and 8 use entropy as a measure that should be maximised. This approach is common in other sciences. Famously, the Maximum Entropy Principle (MEP), discussed by Jaynes (1978), asserts that models in statistics and physics should seek to maximise entropy. This principle found important uses in many contexts in which the hierarchy of types in a system matters for its function (Frank 2018).

Unlike physics and biology, microeconomic and macroeconomic behaviours are more difficult to disentangle. For instance, economic agents possess the ability to foresee changes in their environment, and may act on it, in a fashion that gasses or animal species are unable to. In any case, Foley (1994) first demonstrates the relevance of the MEP in economics, and Scharfenaker & Foley (2017) use it to study firm profit distributions. Arguably, the Olley & Pakes (1996) decomposition of labour productivity is similar in spirit to the MEP; the authors rely on covariance, instead of entropy, to measure the tendency with which output fails to align with productivity at the firm level.

In this thesis, I demonstrate that one important purpose for entropy in economics is to introduce a coherent framework with which to study out-of-equilibrium dynamics without corner solutions. I elaborate on this point briefly in the context of the canonical *Capital Asset Pricing Model* (CAPM) in finance below. Another example is the Olley-Pakes decomposition,

which would suggest that only the most productive firm should dominate the economy. Providing an extra consideration for variety, using entropy as a measure, induces diminishing returns to concentrating economic quantities in one, or few types. In this manner, a distribution of firms with varying levels of productivity co-exists in equilibrium.

Why should entropy increase? I use the measure of entropy to study distributions in two, well-established contexts. In Chapter 7, the distribution of interest income over maturities – which constitutes the *Forward Rate Curve* – is reached by a representative bank that seeks to maximise its return on assets. The ‘entropy’ of the curve is pushed up, because any bond that pays in excess of the exogenous short term rate, scaled by maturity, is an opportunity to increase the long term return on the bank’s assets.

In Chapter 8, entropy stems from a utility function that compels investors to diversify their portfolio of stocks. In this scenario, entropy emulates diminishing marginal returns to investing in any one stock. This prevents corner solutions when investors desire holding stocks with, say, high book-to-market values. Investing more in stocks with a high perceived value can be less useful than investing in stocks with lower values if their small shares present an opportunity to better diversify the portfolio.

The role of constraints Pursuing variety does not come without a price. A competing incentive might assert a constraint on the degree to which entropy can be maximised. In Chapter 7, banks are liable to a constraint on the maturity of their interest income. Hypothetically, this is because of the uncertain timing with which deposits might be withdrawn. In order to avoid a run on deposits, banks thus prefer investments with shorter maturities compared to those with longer maturities. The relative importance of meeting a constraint on maturity relative to diversifying interest income results in a closed form solution for the Forward Rate Curve, whereby the interest rate scales with maturity.

Chapter 8 follows the traditional approach in portfolio theory, by emphasising the role of price signals ahead of entropy. For a typical portfolio selection model, including the CAPM and modern factor models, the stated purpose is to maximise the exposure to some measure of expected returns. These measures are constructed from various characteristics of stocks, such as the straightforward statistical mean and variances in Markowitz (1952), the covariance with market returns from Sharpe (1964) and Lintner (1965), or balance sheet items as in Fama & French (1992, 2015).

These characteristics have always been central to portfolio theory, and so they are in Chapter 8. A hypothetical investor seeks to maximise their exposure to a measured characteristic by allocating wealth shares that align most closely with the desired characteristic. By itself, a portfolio grows to be progressively more concentrated in a few stocks with the highest measured characteristic. In the absence of any changes, one stock dominates the entire portfolio – and possibly the market.

In this context, entropy reigns in concentration as the investor is compelled to diversify their exposure to a certain characteristic. In particular, the returns to holding more of the desired characteristic are decreasing, a feature I accomplish by simply assuming exponen-

tial utility for the payoff to the investor. Thus, entropy manifests as a constraint on the concentration of the portfolio, since entropy is only maximal when all wealth shares are equal. In portfolio theory, the relationship between maximising value versus diversification is referred to as a trade-off between a portfolio's *depth* versus *breadth* (Grinold 1989, Ding & Martin 2017).

Novel insights Chapter 7 explains patterns in forward rate fluctuations from changes in duration, which governs the equilibrium slope of the Forward Rate Curve. An additional adjustment for curvature using an extra constraint on the geometric average duration produces a Gamma pattern for forward rates, which holds up well empirically. This pattern suggests that extreme forward rate fluctuations occur at a maturity of around two to three years, where rates are susceptible to changes in maturity on both the linear as well as the geometric scales.

Prominent models of the Forward Rate Curve interpret future prices as expectations for future spot interest rates, and the role of duration is thus relatively unexplored. The model in Chapter 7 is fit to daily Eurodollar futures data, and could readily extend to micro data on mortgages and bonds, promising to help explain why long-dated interest rates appear to fluctuate in excess of short term rates.

Entropy in Chapter 8 produces a natural and intuitive interpretation of factor models. At the market level, a factor model tracks the momentum of returns to stock characteristics – investment, systemic risk, book value – because investors seek exposure to a portfolio of stocks that yield these characteristics, instead of concentrating in a few stocks that feature the greatest measurements for them. In addition, this constraint on wealth shares via entropy manifests as a size factor, which has dominated factor models since Banz (1981).

The existence of that factor has not been well explained thus far. Furthermore, its role in tracking the concentration of the stock market serves to reconcile the recent literature on asset demand with more traditional factor models. Specifically, Koijen & Yogo (2019) use a logit representation for mutual fund portfolios, which Chapter 8 replicates as the equilibrium allocations for the hypothetical investor maximising the entropy of their wealth, subject to a desired characteristic. Moreover, the factor model representation lends itself to a straightforward model for the *changing* preferences of that investor, thus moving the equilibrium allocations over time. Empirically, this type of momentum – changes in allocations due to preferences – explains considerably more variation in stock returns than changes in characteristics. This is an important observation in the pursuit to explain why stock prices fluctuate in excess of fundamentals.

1.3 Overview

This thesis is organised in two parts. The first part ranges from Chapter 2 to Chapter 5. It presents a series of empirical findings on large fluctuations in economic quantities, notably stock prices and labour productivity, by focusing on disaggregated data. In doing so, each chapter contextualises its findings and methods in the current literature. Chapter 2 studies

the role of social contagion in retail investor sentiments, and the subsequent impact on stock prices. Chapter 3 further investigates the temporal dynamics in stock return fluctuations in the presence of consensus formation. Chapter 4 documents the presence of heavy tails in labour productivity at the firm level. Chapter 5 investigates the sources of the slowdown in aggregate labour productivity growth through a survey of the literature, and by revisiting models that decompose aggregate productivity by sources of growth with recent data.

The second part synthesises these patterns by introducing the notion of entropy in Chapter 6. Chapter 7 studies fluctuations in the Forward Rate Curve based on distributional methods in which entropy manifests as returns derived by banks. In Chapter 8, entropy serves as a measure for the impact of the evolving demand for various types of assets by mutual funds on stock return variation. Chapter 9 concludes.

Bibliography

- Banz, R. W. (1981), ‘The relationship between return and market value of common stocks’, *Journal of Financial Economics* **9**(1), 3–18.
- Baqaei, D. R. & Farhi, E. (2020), ‘Productivity and misallocation in general equilibrium’, *Quarterly Journal of Economics* **135**(1), 105–163.
- Ding, Z. & Martin, R. D. (2017), ‘The fundamental law of active management: Redux’, *Journal of Empirical Finance* **43**, 91–114.
- Fama, E. F. & French, K. R. (1992), ‘The cross-section of expected stock returns’, *The Journal of Finance* **47**(2), 427–465.
- Fama, E. F. & French, K. R. (2015), ‘A five-factor asset pricing model’, *Journal of Financial Economics* **116**(1), 1 – 22.
- Foley, D. K. (1994), ‘A statistical equilibrium theory of markets’, *Journal of Economic Theory* **62**(2), 321 – 345.
- Frank, S. A. (2018), ‘The Price equation program: Simple invariances unify population dynamics, thermodynamics, probability, information and inference’, *Entropy* **20**(12).
- Gabaix, X. & Koijen, R. S. (2023), ‘Granular instrumental variables’, *Journal of Political Economy* (forthcoming) .
- Goldin, I., Koutroumpis, P., Lafond, F. & Winkler, J. (2024), ‘Why is productivity slowing down?’, *Journal of Economic Literature* **62**(1), 196–268.
- Greenwood, R., Shleifer, A. & You, Y. (2019), ‘Bubbles for Fama’, *Journal of Financial Economics* **131**(1), 20–43.
- Grinold, R. C. (1989), ‘The fundamental law of active management’, *Journal of Portfolio Management* **15**(3), 30–37.

- Grossman, S. J. & Stiglitz, J. E. (1980), 'On the impossibility of informationally efficient markets', *The American Economic Review* **70**(3), 393–408.
- Jaynes, E. T. (1978), Where do we stand on maximum entropy?, in R. D. Levine & M. Tribus, eds, 'The Maximum Entropy Formalism', MIT Press, Cambridge, Mass., pp. 15–118.
- Koijen, R. S. J. & Yogo, M. (2019), 'A demand system approach to asset pricing', *Journal of Political Economy* **127**(4), 1475–1515.
- Lintner, J. (1965), 'The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets', *The Review of Economics and Statistics* **47**(1), 13–37.
- Markowitz, H. (1952), 'Portfolio selection', *The Journal of Finance* **7**(1), 77–91.
- Olley, G. S. & Pakes, A. (1996), 'The dynamics of productivity in the telecommunications equipment industry', *Econometrica* **64**(6), 1263–1297.
- Scharfenaker, E. & Foley, D. K. (2017), 'Quantal response statistical equilibrium in economic interactions: Theory and estimation', *Entropy* **19**(9).
- Schumpeter, J. A. (1943), *Capitalism, socialism, and democracy*, G. Allen & Unwin Ltd, London.
- Sharpe, W. F. (1964), 'Capital asset prices: A theory of market equilibrium under conditions of risk', *The Journal of Finance* **19**(3), 425–442.
- Shiller, R. J. (1984), Stock prices and social dynamics, Technical Report 2, The Brookings Institution.
- Yang, J., Heinrich, T., Winkler, J., Lafond, F., Koutroumpis, P. & Farmer, J. D. (2024), 'Measuring productivity dispersion: a parametric approach using the Lévy alpha-stable distribution', *Industrial and Corporate Change* p. dtae021.

Chapter 2

Social Contagion and asset prices: Reddit's self-organised bull runs

(With Valentina Semenova)

Abstract

In recent years, online discussion platforms have amplified the impact of social forces on financial markets. This paper uses text data from Reddit's WallStreetBets (WSB) forum to document that user sentiments about assets' future performances are motivated, in part, by peers. The paper then estimates the effect of WSB activity on asset prices in three ways. We document: higher retail trader demand following WSB discussions, the price impact of idiosyncratic investor sentiment in viral content, and the significant activity of WSB users around bubble-like stock price run-ups. Our findings suggest that social media can destabilise financial markets.

2.1 Introduction

In investigating the stock market crash of May 28, 1962, the Securities and Exchange Commission (SEC) found that: ‘investor “psychology” being what it is, the increasing decline in one or several issues can easily spread to others. Once the process becomes generally operative, the stage is set for a serious market break’ (1963). The SEC concluded that large institutions acted as a balancing force during the collapse. The report pointed at retail traders as the *key players* behind the panic. Over half a century later, we are again confronted with the consequences of investors’ social behaviours. As online discussants on Reddit’s ‘WallStreet-Bets’ (WSB) forum drove up the price of GameStop shares in January, 2021, retail investors regained a spotlight on the (virtual) trading floor. A key difference between 1962 and today is the internet, which offers both a coordination platform on an unprecedented scale, and a new datasource on investor narratives, interactions and psychology.

In this paper, we study how behaviours on social media have impacted financial markets using data from the WSB forum. Our analysis proceeds in two steps. First, we leverage the quasi-random variation in how users are exposed to information on Reddit to study peer effects. WSB users build expectations based on the perspectives shared by peers *in addition to* their own experiences. Second, we adopt several recent developments in finance to estimate how social dynamics online have impacted trade volumes, asset prices and volatility. We use the method of [Boehmer et al. \(2021\)](#) to demonstrate that retail trading volumes follow online discussions. The paper leverages the techniques developed by [Gabaix & Koijen \(2023, 2021\)](#) to show how viral content can impact asset prices and volatility. Finally, we adapt the framework of [Greenwood et al. \(2019\)](#) to investigate the interest of investors on social media in assets exhibiting bubble-like dynamics.

Social contagion This paper complements the existing literature on social investing and narrative finance by studying peer effects in sentiments on a relatively unexplored text-based dataset. Compared to typical investor surveys, online platforms produce real-time observations for expressed beliefs, allowing us to examine reactions to information at a granular timescale.

A particular challenge to the study of peer effects on social media platforms is the presence of homophily – the tendency for users to select content as a function of their own beliefs ([Cookson et al. 2022](#), [Chen & Hwang 2022](#)). Unlike many other investor discussion platforms, WSB does not present content through self-selected followership ties. Instead, it presents users with posts sorted on their recency and popularity. The distinction is critical – other platforms offer limited scope for studying peer effects, since users self-select to follow other investors with similar views. In our data, on the other hand, individuals are exposed to the current and popular content shared on the platform, regardless of whether this matches their own perspective. Put differently, if two users were to log onto WSB simultaneously within our sample period, they would be exposed to identical content (regardless of their own prior activity).

We leverage this variation in sentiments expressed regarding the future performance of

assets to propose two Instrumental Variable (IV) strategies that quantify the effect of peer sentiment on users. First, we select individuals who express sentiments about an asset multiple times and observe peers discussing the same asset in-between. We use historic peer sentiment as an IV, which mitigates the common shock problem. This approach is inspired by the peer effects in the classroom literature, which gauges future student performance based on entry exams (Duflo et al. 2011). Indeed, more recent work (posted a year after this paper) leverages a similar estimation strategy to ours for studying peer effects in cryptocurrencies (Han et al. 2023) further underscoring the usefulness of our proposed methodology. Second, we leverage the WSB network of interactions to identify the information to which an individual investor has been exposed. The network links an older submission about an asset to a new submission if the author of the new submission comments on the older submission. To estimate a parameter for social contagion, we regress the sentiment expressed by the new submission on the average sentiment of older, linked submissions. We use the timing of our IVs to control for common shocks and instrument the sentiment of linked submissions to control for an author’s endogenous choice to comment (Zenou 2016).

In both approaches, exogenous variation in average peer sentiment is a statistically significant predictor for the change in author sentiment. This finding suggests that retail investors experience investment complementarities and adapt their strategies based on those of their peers. The IV results reveal that when the odds of peers expressing bullish over bearish sentiments double, the odds of a given user expressing bullish over bearish sentiment increase by an average of 15%.

Our empirical exercise sheds light on several other phenomena. First, we confirm that retail investors on WSB tend to be trend-followers. Specifically, a log-return of 0.1 on a given day increases the probability of a user posting a bullish over bearish submission – regarding the *future* performance of the stock – by twenty percent. Overall, peer effects appear to play *a greater role* in sentiments expressed by WSB users than extrapolation. Second, we explore how sentiments on WSB respond to market surprises (defined as log-returns that are two standard deviations above or below a stock’s monthly average), as well as reinforcement between peer sentiments and market performance. Interestingly, the effect is negative and significant for negative market surprises, however, positive market surprises appear to have no impact. This suggests a mechanism through which downside panic can spread quickly among investors.

Asset prices The paper subsequently explores the impact of WSB activity on asset prices in three ways. First, we consider whether changes in retail investor trading patterns can be explained by variation in WSB discussions. We extract intra-day trade data in the most popular stocks on WSB from the Trades and Quotes (TAQ) dataset and identify retail trades as those that are traded at the sub-penny increment (Boehmer et al. 2021). Our approach demonstrates that changes in interest in certain assets over others on WSB can explain changes in the fraction of retail trades observed in the market. Contrary to the narrative that WSB’s sole purpose was to fuel pump-and-dump schemes, the effect appears more pronounced for stocks with large market caps, which are less prone to market manipulation (such as AAPL,

AMZN and MSFT).

Second, we use a Granular Instrumental Variable (GIV) framework to demonstrate how online virality drives asset prices and volatility. To construct our GIV, we compare the average sentiment of all submissions to the popularity-weighted sentiment of submissions in a given week. Our analysis reveals a statistically significant relationship between our GIV and future returns. Specifically, a doubling in the odds of a very popular submission expressing bullish over bearish sentiments results in an average increase in log-returns of approximately 0.4pp the following week.

Finally, we demonstrate a high propensity for WSB investors to take interest in assets experiencing bubble-like dynamics by extending the method for classifying index bubbles in [Greenwood et al. \(2019\)](#) to individual stocks. There is a statistically significant difference in activity levels on WSB for stocks that experience run-ups in price with a subsequent downturn, compared to price run-ups without a downturn. The findings underscore the importance of understanding the timing of retail investor trading patterns, and the redistributional consequences of social investing ([Pearson et al. 2021](#), [An et al. 2022](#)).

Economists have long deliberated to what extent social dynamics and human psychology play a role in economic decision-making ([Shiller 1984](#), [Black 1986](#)). We show how psychological mechanisms, amplified by social media, have impacted the investment landscape and demonstrate concurrently the opportunities for using unstructured text data and Machine Learning (ML) tools for finance ([Kuchler & Stroebel 2021](#)).

Related literature This paper contributes to several rapidly-evolving domains within the finance literature, including those: documenting investor narratives and dynamics, studying peer effects in finance, and bubble dynamics. The paper quantifies peer effects and other psychological biases which are prevalent among WSB users, and reconciles our estimates with a proposed modelling framework. We, subsequently, directly evaluate market impact in three different ways: using TAQ data to track shifting retail investor preferences, using a GIV to quantify the impact of WSB discussions on returns, and by distinguishing different stock price run-ups into bubbly versus stable.

A prominent strand of literature highlights the importance of peers and narratives in shaping investor perspectives. In his seminal work, [Shiller \(1984\)](#) discusses the excess volatility in stock prices relative to dividends. Since then, ‘narrative economics’ has played an increasingly important role in our understanding of investor decision-making and market moves ([Shiller 2017](#), [Hirshleifer 2020](#)). Social media forums have been used to uncover various aspects of investor behaviour, such as: the myForexBook platform to study the disposition effect ([Heimer 2016](#)), StockTwits to study echo chambers ([Cookson et al. 2022](#)), and SeekingAlpha to study impression management and the propagation of noise ([Chen & Hwang 2022](#)), to name a few (see [Kuchler & Stroebel \(2021\)](#), [Cookson et al. \(2024\)](#) for a comprehensive overview). The WallStreetBets forum has been used to study: the interactions of different types of investors ([Hu et al. 2021](#)), the informational content of posts for prediction ([Bradley et al. 2021](#)), and the dynamics behind the GameStop short squeeze ([Mancini et al. 2022](#)). We study a different question to those listed above and propose an empiri-

cal methodology to more precisely identify how information shared by peers contributes to sentiment formation, as well as the impact of market surprise and reinforcement. Beyond studying a distinct question in behavioural finance, we explore three different avenues for market impact.

The impact of peers has been studied in several different financial contexts such as: the diffusion of micro-finance decisions in a social network ([Banerjee et al. 2013](#)), the effect of peers on risk taking ([Lahno & Serra-Garcia 2015](#)), the effect of social networks on saving ([Breza & Chandrasekhar 2019](#)), and the role of ‘social learning’ versus ‘social utility’ in financial decision-making ([Bursztyn et al. 2014](#)). By studying a broader set of investors in a natural experiment, our research question is similar to [Pool et al. \(2015\)](#), who demonstrate that socially connected fund managers hold similar stocks. Social media data, however, presents a novel identification challenge. Fortunately, other papers in economics offer promising methods to leverage naturally occurring variation in peers for identification. An area which pioneered many of these techniques investigates peer effects in the classroom (see [Epple & Romano \(2011\)](#), [Sacerdote \(2011\)](#) for a general overview, and [Duflo et al. \(2011\)](#) for a prominent example). Social networks are also an active area of study (see [Bramoullé et al. \(2020\)](#) for a recent review). The present paper highlights how to transfer well-established techniques from the empirical peer effects in the classroom and networks literatures to social media data.

Investor discussion forums have been documented to have predictive power for market returns ([Antweiler & Frank 2004](#), [Sabherwal et al. 2011](#), [Chen et al. 2014](#), [Azar & Lo 2016](#), [Agrawal et al. 2018](#)). The present paper quantifies several additional channels through which social media has had an impact. We leverage the methodology of [Boehmer et al. \(2021\)](#) to show how retail trading volumes follow the attention paid to specific assets on the forum, complementing other works tracking WSB positions, such as [Welch \(2022\)](#). We present a new channel for price impact and use a GIV approach to demonstrate how virality online may be driving volatility in the financial markets ([Gabaix & Koijen 2023](#)). Since its development, the GIV methodology has been used to study several important questions, including credit risk ([Galaasen et al. 2020](#)) and capital flows ([Gabaix & Koijen 2021](#)). This is the first work, to our knowledge, to study the impact of social media and viral content. The GIV approach allows us to study fluctuations in returns at the weekly time-scale, however, we also consider that social movements can have more long-term consequences, namely through driving asset price bubbles. Several papers empirically study the formation bubbles in the markets. [Greenwood et al. \(2019\)](#) propose a methodology to study price run-ups in sectors. We adapt their framework to our setting and demonstrate the increased activity of hype investors in assets that exhibit bubble-like dynamics. Our brief exploration prompts an important question for further research. Recent studies use investor-level holdings data to track positions and profits / losses of individuals ([Pearson et al. 2021](#), [Balasubramaniam et al. 2023](#)). Given the quickly growing hype investor population in the market (a statement justified in our conclusion and data appendix), there is a pressing need to study these individual investors’ profit and loss profiles as the re-distributional consequences of social influence among investors ([An et al. 2022](#)).

Road map We present our results in four sections. The following section comprehensively describes the data source and relevant variables. Section 2.3 presents empirical evidence for peer effects on WSB. Section 2.4 empirically evaluates the effect of WSB activity on financial markets. Section 2.5 concludes.

2.2 What is *WallStreetBets*?

Reddit, launched in 2005, is a social news aggregation, web content rating, and discussion website. It was ranked as the 8th most visited site globally in March 2024,¹ with over 268 million anonymous, active, weekly users.² The website’s contents are self-organised by subject into smaller sub-forums, ‘subreddits’, which discuss a unique, central topic.

Structure of WSB Within subreddits, users publish titled posts (called ‘submissions’), typically accompanied with a body of text or a link to an external website. These submissions can be commented and ‘upvoted’ or ‘downvoted’ by other users. A ranking algorithm raises the visibility of a submission with the amount of upvotes it receives, but lowers it with age. Therefore, the first submissions that visitors see are i) highly upvoted, and ii) recent, with the precise algorithm considered private intellectual property and discussed further in Appendix 2.A.1.³ Comments on a submission, visible to anyone, are subject to a similar scoring system, and can, themselves, be commented on.

Features The WSB subreddit was created on January 31, 2012, and reached one million followers in March 2020.⁴ As per a Google survey from 2016 (later reinforced by more recent surveys⁵), the majority of WSB users are ‘young, male, students that are inexperienced investors utilising real money (not paper trading); most users have four figures in their trading account’.⁶ Individuals on the forum discuss and express their sentiments about stock-related news. In addition to market discussions, there is ample evidence of users pursuing the investment strategies encouraged in WSB conversations. Users post screenshots of their investment gains and losses, which subreddit moderators are encouraged to verify – a dynamic reminiscent of Shiller’s (2005) description of an asset bubble. The discussions are whimsical, but mostly investment-focused.

Available data We downloaded WSB data using the PushShift API.⁷ PushShift records all comment and submission data at the time of creation. The full dataset consists of two parts. The first is a total of 1.4 million submissions, with their authors, titles, text and timestamps. The second is comprised of 16.5 million comments, with their authors, text, timestamp, and

¹<https://www.similarweb.com/top-websites/>

²<https://backlinko.com/reddit-users>

³https://www.reddit.com/r/help/comments/717686/order_of_posts/

⁴<https://subredditstats.com/r/wallstreetbets>

⁵https://www.reddit.com/r/wallstreetbets/comments/a2qvj3/wsb_demographics_survey_result_s_a_portrait_of/

⁶<https://andriymulyar.com/blog/how-a-subreddit-made-millions-from-covid19>

⁷<https://pushshift.io/>

the identifier of the parent comment or submission. Submission and comment numbers have grown exponentially since 2015 – Figure 2.A.3 in Appendix 2.A.1 displays the forum’s exponential growth.

Our dataset spans January, 2012 to July, 2020. Importantly, our results do not hinge on the 2021 GameStop (GME) short squeeze. The decision to focus on this timeframe is intentional: before the GME short squeeze, WSB received less attention from institutional investors, as well as less bot-activity. Indeed, Bradley et al. (2021), for example, find that the quality of information content and discussion changed on the forum around the time of the short squeeze. Our sample tracks retail investor discussions more precisely, without systematic external influence. However, in our supplementary material we also present results using the data of Semenova et al. (2024), which extend through mid-2021.

Identifying assets The following sections predominantly rely on submissions for text data, since they are substantially richer than individual comments. Comments are used to trace interactions between discussants. In order to understand how users discuss specific assets, we extract mentions of *tickers* from the WSB submissions’ text data. A ticker is a short combination of letters, used to identify an asset on trading platforms. For example, ‘AAPL’ refers to shares in Apple, Inc. Appendix 2.A.2 documents how tickers are extracted from submissions, as well as the heavy-tailed nature of ticker discussions. In total, we are left with 111,765 submissions that mention one, unique ticker. These submissions have 1.9 million comments in total.

Sentiment model In addition to extracting tickers, we gauge whether submissions express an expectation for an asset’s future price to rise, the *bullish* case, to fall, the *bearish* case, or to remain unpredictable/stable, the *neutral* case. We identify sentiment using a supervised-learning approach, with a hand-labelled dataset of almost five thousand submissions for training, validation and testing (Araci 2019). The sentiment model outputs a probability for each sentiment category, achieving 70% accuracy in categorising the manually labeled test set. Appendix 2.A.3 discusses details of this Natural Language Processing (NLP) model and the distribution of labels. We test several other more recent large language models, for example Mistral 7B (Jiang et al. 2023), and do not observe in a significant improvement in performance (though context-enhanced LLMs, such as Zhu et al. (2021), have potential to improve the performance - we leave this for further research). We also test less computationally complex models, such as a bi-term regex, and observe a decline in performance.

Key sentiment variable The sentiment classifier assigns three probability scores to each submission about a ticker: the probability of a submission being bullish, $P(\phi = +1)$, bearish, $P(\phi = -1)$, neutral, $P(\phi = 0)$. The probabilities sum to one. At the time t when an author i posts about asset j , we use the probability scores above to calculate a continuous sentiment

score between $(-\infty, \infty)$:

$$\Phi_{i,j,t} = \frac{1}{2} \log \left(\frac{P(\phi_{i,j,t} = +1)}{P(\phi_{i,j,t} = -1)} \right). \quad (2.1)$$

Submissions labelled as bullish ($P(\phi = +1) = 1$) (or bearish ($P(\phi = -1) = 1$)) are set to $P(\phi = +1) = 0.98$ (or $P(\phi = -1) = 0.98$) to retrieve a finite value for the log-odds. For certain exercises, we assign three categorical variables (bullish, bearish, neutral) to a submission's sentiment, encoded with a one if the corresponding label received the highest probability from our classifier. As such, this categorical variable $\phi_{i,j,t}^{+1}$ will be equal to one if author i 's post about asset j at time t is categorised as bullish; $\phi_{i,j,t}^0$ and $\phi_{i,j,t}^{-1}$ will be zero. We leverage these variables to investigate investor sentiment throughout the paper. Appendix 2.A.3 shows the distribution of our key sentiment variable $\Phi_{i,j,t}$.

We argue that the error rate and distribution of errors from our sentiment classifier are appropriate for the task we study. In most of our analysis, we use a continuous transformation of the sentiment – the log-odds of bullish over bearish sentiment. In our confusion matrix, we observe that the model mistaking bullish for bearish sentiment and vice-versa is rare: 6% in the former case and 7% in the latter case. The model has a greater difficulty distinguishing positive/neutral and negative/neutral, however, this is expected, given the continuous nature of language – fortunately, these errors have a lesser impact on our analytical result. We observe human annotator-to-annotator disagreement of 10%.

Additional Information We provide a detailed data appendix – Appendix 2.A – which discusses: i) the growth of the forum over time, ii) the way content is presented to users, followership ties on Reddit and why our identification assumptions hold, iii) the details of the ticker and sentiment extraction methodologies, and iv) the prevalence of hype traders. In a preliminary analysis, we manually extract positions from screenshots publicised on WSB and sentiments (as described above). Expressing a positive sentiment about an asset on WSB raises the probability of a long investment in the same asset in the future by over 6 times (compared to the probability of a short investment) – Appendix 2.A.6 presents this analysis, while additional evidence of trading is offered through our use of TAQ data in Section 2.4. We also perform a preliminary analysis of how beliefs relate to asset prices and document a positive relationship between returns preceding sentiments (evidence of extrapolation) and a subsequent reversal. The details of this exercise are presented in Appendix 2.A.6.

2.3 Social dynamics on WSB

What impacts the sentiments of investors active on the WSB forum? This section provides empirical evidence for the existence of two mechanisms – namely peer effects and extrapolation – among investors on WSB. Appendix 2.B explores how these effects can impact asset prices using a stylised model inspired by the literature on strategic information complementarities (Hellwig & Veldkamp 2009, Zenou 2016) and ‘cognitive’ finance (Bordalo et al. 2021).

Estimating equation The target independent variable of interest for studying hype investor sentiment is the log-odds of bullish over bearish sentiment, modelled as

$$\Phi_{i,j,t} = g(b_{i,j,t}) + f(\bar{\phi}_{-i,j,(t-1,t)}) + \epsilon_{i,j,t}, \quad (2.2)$$

using a utility framework detailed in Appendix 2.C.1. An author i chooses a bullish over bearish strategy in asset j depending on: i) a signal $b_{i,j,t}$, and ii) the observed sentiments of peers, $\bar{\phi}_{-i,j,(t-1,t)}$. The goal of this paper is to better understand $g(\cdot)$ and $f(\cdot)$. We set out to empirically whether: whether $f(\cdot)$ is increasing (underscoring the existence of strategic information complementarities among these investors, as explained in the model appendix), the importance of extrapolation in $g(\cdot)$, as well as other potential impacts documented in cognitive finance (such as reinforcement and surprise).

Empirical strategy: consensus formation among investors A key challenge for estimating Eq. 2.2 is the existence of confounding variables, especially when estimating $f(\cdot)$ – peer influence. We, therefore, use two approaches to estimate Eq. 2.2: i) the *Frequent Posters* approach, and ii) the *Commenter Network* approach. Both leverage different features of our data. For the *Frequent Posters* approach, we leverage the fact that certain users post multiple submissions about the same asset (hence, *frequent*) and the quasi-random variation in exposure to content on Reddit. For the *Commenter Network* approach, we use instances in which users comment on others’ submissions in combination with an IV proposed in the networks literature (Zenou 2016, Patacchini & Zenou 2016, Bifulco et al. 2011) to more precisely gauge the transmission of sentiments about the same asset.

Consuming content on WSB Understanding the way in which content is consumed by users on WSB underpins our IV strategies. The options for sorting and filtering content on WSB are *identical* across users – put differently, two users who log on at the same time and choose the same sorting option will be exposed to identical content, regardless of their own posting or commenting history. Reddit does not have friendship/follower ties within specific forums.⁸ Individuals cannot filter exposure to certain sentiments over others and do not receive specific, tailored notifications to view content related to their past commenting or posting history. To illustrate content consumption, consider two users **A** and **B**: **A** expressed bullish sentiments about AAPL, while **B** created several bearish posts (or commented on several bearish posts) about the same asset. If both users were to subsequently view WSB simultaneously (within our sample period), they would be exposed to the same exact content (verified through Reddit documentation and website AB testing). The posts that an individual is exposed to on WSB depend on what other anonymous, disconnected users have posted on the forum shortly before the author logs on, and what topic has recently gained popularity (see Section 2.2 and Appendix 2.A.1 for a detailed description). This exposure of users to different opinions (similar in spirit to Weidmann & Deming (2021)) allows

⁸Followership ties on Reddit more broadly and the fact that they do not impact our approach are discussed in Appendix 2.A.1.

us to estimate direct peer effects.

For readers less familiar with Reddit, we present screenshots of submissions, comments and content sorting options in Appendix 2.A.

2.3.1 Identifying peer influence: Frequent Posters

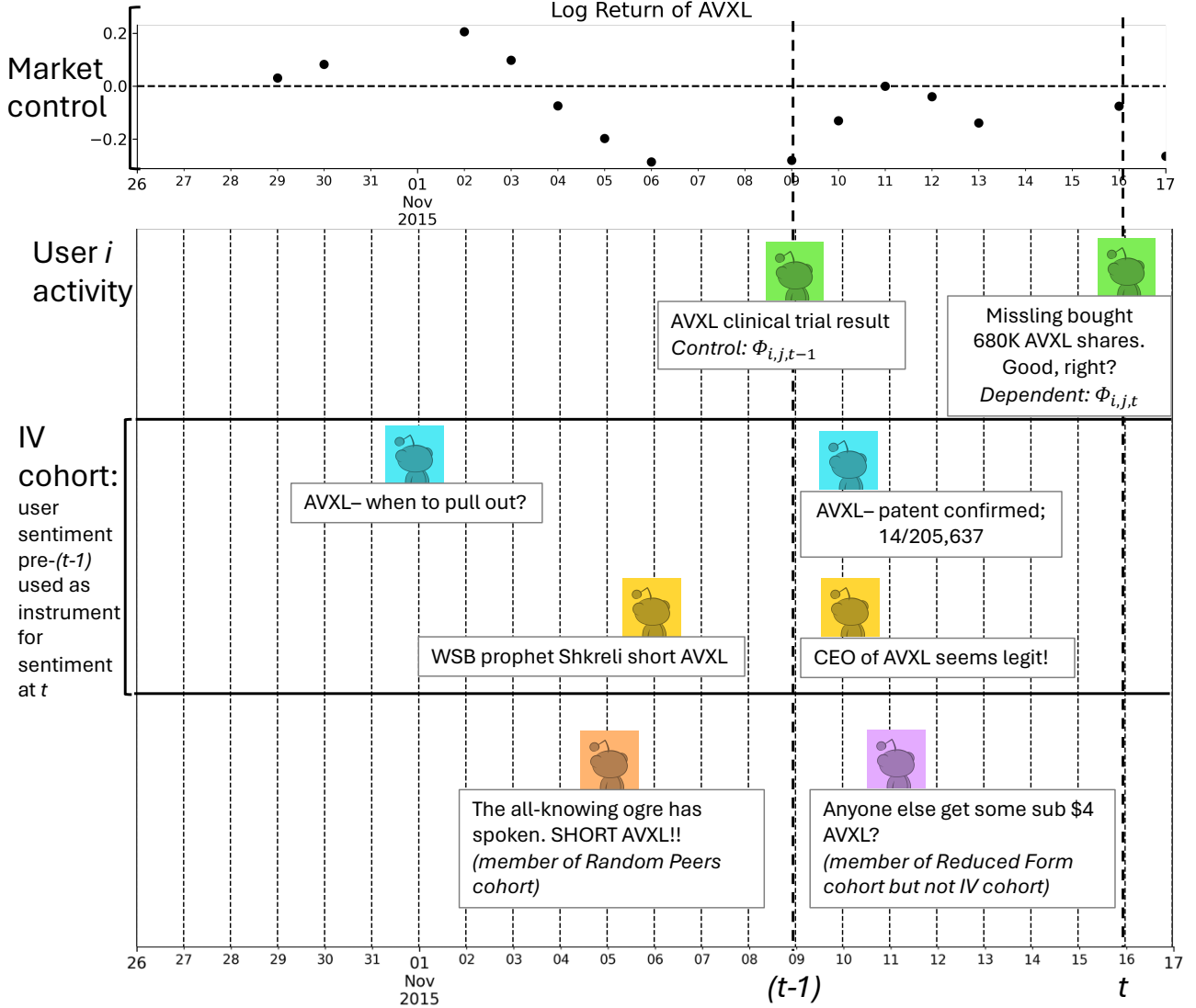


Figure 2.1: Illustration of Frequent Posters Approach using Observations from WSB Data: We observe user i create two posts about asset j (AVXL). She observes posts of peers between her two posts at time $(t-1)$ and t . In our IV approach, we use previous sentiments of peers to predict their sentiments between times $(t-1)$ and t – leveraging opinions that are formed exogenously to market conditions between $(t-1)$ and t . Market returns, as well as ticker fixed effects, are included as a control.

For the *Frequent Posters* approach, we observe that a subset of authors create at least two submissions about the same ticker. We quantify peer influence by identifying the impact of other authors who write submissions about the same asset *between* an individual's two submissions. We use an IV of previous, expressed peer sentiments to control for exogenous shocks (see Figure 2.1 for an illustration and Appendix 2.C.4 for a Directed Acyclic Graph

(DAG) illustration). Our approach allows us to control for the author’s sentiment prior to exposure to his peers, in addition to market moves. We complement this approach with the *Commenter Network* approach, described in the next section.

Within WSB, we observe author i initially express a sentiment about an asset j , $\Phi_{i,j,(t-1)}$ (the continuous log-odds of a post expressing bullish over bearish sentiment, as per Eqs. 2.1&2.2), and, subsequently, write a new submission about the same asset at a later time, with an updated sentiment $\Phi_{i,j,t}$ (where time t is in event time). In the time between these posts, the author may observe submissions by others on the same asset expressing average sentiment $\bar{\Phi}_{-i,j,(t-1,t)}$, in addition to outside information related to the asset. Our goal is to identify the effect that expressed peer sentiments have on changing author i ’s sentiment.

Reduced form We first estimate the effect of average peer sentiment between an author’s two submissions with the following linear model:

$$\Phi_{i,j,t} = \kappa_1 \bar{\Phi}_{-i,j,(t-1,t)} + \kappa_2 \Phi_{i,j,t-1} + \kappa_3 r_{j,t} + \kappa_4 \bar{r}_{j,t} + \kappa_5 \sigma_{j,t}^2 + \kappa_6 X_j + \epsilon_{i,j,t}, \quad (2.3)$$

where the controls include: stock-specific fixed effects X_j , author i ’s past sentiment $\Phi_{i,j,t-1}$, stock log returns on the day of the author’s post $r_{j,t}$ and the five days prior $\bar{r}_{j,t}$, the variance in log returns on the five days prior to the author’s post $\sigma_{j,t}^2$. Even though an individual cannot filter exposure to certain sentiments over others, an exogenous shock in the period $(t-1, t)$ may affect the views of both peers and the author in question simultaneously. For this reason, the OLS estimates do not enable us to precisely estimate peer influence.

Instrumenting peer sentiment To tackle this issue, we use the historical views of peers as an IV for their views expressed within $(t-1, t)$. Our choice of IV is founded in psychology: [Ross et al. \(1975\)](#) find that ‘once formed, impressions are remarkably persevering and unresponsive to new input’, with later studies, such as [Anderson et al. \(1980\)](#), supporting these findings.

We reason about our choice of IV through the example in Figure 2.1, which presents a true example from the WSB discourse. The author of interest i creates two posts about AVXL and observes the posts of peers in-between. We use the historical sentiments of peers (created before $t-1$) as our IV. In this way, we isolate variation in peer sentiment *exogenous* to market conditions and news about asset j between time $(t-1, t)$.

We perform several robustness checks to ensure the validity of our instrument (discussed in greater detail after our results): we show that *predicted* peer sentiment is not correlated with asset returns between time $(t-1, t)$ demonstrating that there is no selection bias for a subset of peers whose past sentiment correlates with future returns or news. This is further reinforced by us including several controls for returns. We check that our estimation is not a lagged news effect by replacing the peers with a random peer cohort which posts before time $t-1$, and find that the random peer cohort has no effect. We test that there is no selection bias for peers to post a second time only when their sentiment agrees with average forum sentiment about the stock and show that probability of posting is uniform with respect to

the level of agreement with peers. We also elaborate our reasoning about our choice of IV using a DAG in Appendix 2.C.4.

We estimate user k 's sentiment – a peer of investor i – about asset j , $\Phi_{k,j,t}$, based on the sentiment they expressed previously, $\Phi_{k,j,t-1}$, and control for asset returns at the time of their original post, $r_{j,t-1}$:

$$\Phi_{k,j,t} = \kappa_1^0 \Phi_{k,j,t-1} + \kappa_2^0 r_{j,t-1} + \epsilon_{k,j,t}^0, \quad (2.4)$$

where $\epsilon_{k,j,t}^0$ is an idiosyncratic error. The coefficient κ_1^0 estimates the true effect of an individual's historical sentiment. Controlling for $r_{j,t-1}$ allows us to accurately estimate κ_1^0 , while controlling for confounders. Eq. 2.4 is estimated using a sample containing submissions by all authors who post multiple times. The F-statistic for this first stage estimate, presented in Table 2.2, suggests that this is a strong instrument. Our choice of IV gives a good approximation for author sentiment, while allowing us to control for common shocks affecting the sentiments of peers and investor i in the period $(t-1, t)$. We use the predicted outlook of peers between an author's posts, $\hat{\Phi}_{-i,j,(t-1,t)}$, to estimate peer effects as our Second Stage regression, while keeping all other controls the same – historic peer sentiment is used for prediction. Appendix 2.C provides further details on our variable construction and method. Appendix 2.A.4 describes the construction of market variables, and their matching to WSB data.

Results

Table 2.1 presents the results of our estimation with column (1) presenting OLS estimates for κ (from Eq. 2.3) using observed variation in peer sentiments, and column (2) using predicted variation in peer sentiments – independent variables are normalised with respect to their mean and standard deviation (explained further in Appendix 2.C). The *Frequent Posters* approach indicates that peer effects are approximately 1.5 times *more important* in individual sentiment formation, as compared to extrapolation. Our non-normalised coefficient estimates in Table 2.C.2 of 0.19 on predicted peer sentiments means that doubling in the odds of peers expressing bullish over bearish sentiments increases the odds of a given submission to be bullish, over bearish, by 14.1%. In all cases, the robust standard errors, clustered at the ticker level, produce estimates statistically significant at the 1% level. First Stage estimates are presented in Table 2.2.

Support for identification – Frequent posters

One potential concern is that individuals who post multiple times about the same asset may differ from the rest of the population on the forum. If this were the case, our findings would not allow us to draw valid conclusions about the overall population of investors. We provide evidence that sentiments expressed in our samples are similarly distributed to those of the overall user population in Appendix 2.C.

Table 2.1: Peer influence: *Frequent Posters* – full regression estimates

		Dependent Variable: $\Phi_{i,j,t}$		
		Reduced Form (1)	Full Second Stage (2)	Random Peers (3)
Independent Variables	$\Phi_{i,j,t-1}$	0.154 (0.010) ***	0.129 (0.011) ***	0.158 (0.010) ***
	$\bar{\Phi}_{-i,j,(t-1,t)}$	0.055 (0.011) ***	0.036 (0.010) ***	0.005 (0.009)
	$r_{j,t}$	0.022 (0.004) ***	0.025 (0.005) ***	0.023 (0.004) ***
	$\bar{r}_{j,t}$	0.007 (0.004)	0.007 (0.005)	0.006 (0.004)
	$\sigma_{j,t}^2$	-0.003 (0.004)	0.003 (0.008)	-0.003 (0.004)
	Ticker Fixed Effects	Yes	Yes	Yes
No. Observations:		14,396	11,122	14,371
R^2 :		0.12	0.08	0.11
R^2_{adj} :		0.08	0.06	0.08

Notes: The dependent variable is individual investor sentiment about an asset, scaled continuously between $(-\infty, \infty)$, and is estimated by the individual's previously expressed sentiment about the same asset ($\Phi_{i,j,t-1}$) and a set of market control variables ($r_{j,t}, \bar{r}_{j,t}, \sigma_{j,t}^2$), using OLS. The sentiment of peers ($\bar{\Phi}_{-i,j,(t-1,t)}$) is estimated in several ways. In Column (1), we use observed, average sentiment of peers between an author's two posts. In Column (2), we estimate the sentiment of peers using an IV. In Column (3), we select a random cohort to estimate peer sentiment. Robust standard errors, clustered at the ticker level, are presented in parentheses. Observations with incomplete market data are dropped.

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Table 2.2: Peer influence in WSB sentiments – first stage

			Dependent Variable: Sentiment of Peers	
			Frequent Posters (1)	Network (2)
Historical Sentiment of Peers			0.31 (0.01) ***	
Sentiment of Neighbours' Neighbours				0.14 (0.01) ***
Author controls ($X_{i,j,t}^0$)			No	Yes
Controls for returns ($r_{j,t-1}$)			Yes	No
Number of obs.			19,370	24,013
F-statistic			1,105	118

Notes: This table presents the First Stage estimates for peer influence on WSB. In column (1), the First Stage is estimated using the initial sentiment expressed by an author about an asset to estimate his sentiment in the following post. In column (2), the First Stage is estimated using the sentiment of previous submissions that an author commented on, regarding the same asset – the second IV of the neighbour's own historical sentiment is presented in Appendix 2.C.

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Endogenous choice to post multiple times An additional concern may arise around the endogenous choice for certain individuals to post multiple times. Specifically, individuals may choose to post a second time: i) only if their peers agree with them, or ii) when their predictions about the market are correct (in which case the IV would be documenting author i 's response to general market moves, rather than peer influence).

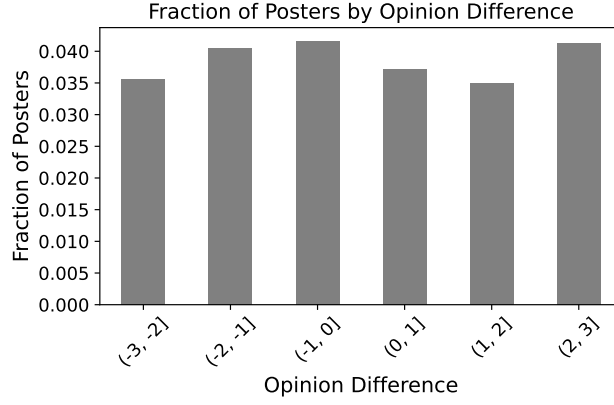


Figure 2.2: **Probability of posting on 5th day bucketed by disagreement level with peers;** the fraction of individuals that post five days after their original post bucketed by the difference between their sentiment and that of peers (the difference is taken between the mean peer sentiment in the intermediate five days and the sentiment expressed by the author in their follow-up post).

We test for (i) by looking at the probability that an individual posts X days after their original post depending on the difference between their sentiment and that of peers (the difference is taken between the mean peer sentiment in the intervening X days and the sentiment expressed by the author in their follow-up post). Figure 2.2 presents our results at the five day time horizon – we observe that the distribution for the probability of posting with sentiment difference is uniform, invalidating the idea that individuals who post are only those who peers agree with. Appendix 2.C presents the same results at different time horizons.

We test for (ii) by including the average return of stock j in the period between $(t - 1)$ and t as an additional control. If individual i was swayed by market moves that coincide with the sentiment of peers, rather than peer effects themselves, this control should erode the significance of peer effects. However, adding this control does not change our results and is not significant - this implies that it is not peers echoing market returns that sways sentiments, but rather the peer effects themselves that impact sentiment. Additionally, we test to see if our predicted peer sentiment, $\hat{\Phi}_{-i,j,(t-1,t)}$, is correlated with the returns on day t when user i posts – the two do not have a statistically significant relationship. Appendix 2.C further discusses both robustness checks.

Distinguishing information from peer effects Another concern is whether our proposed independent variables – asset price movements, ticker fixed effects and author historical sentiments – are effective controls for unobserved ticker characteristics. Along the same lines, one might worry about our ability to distinguish between peer effects and information being absorbed slowly from peers from their initial post before time $(t - 1)$. If our controls in the *Frequent Posters* formulation and identification technique are valid, then a randomly selected cohort of individuals who post on the same ticker *before* the author’s first post, should have no effect on the sentiments expressed in dependent submissions. This would demonstrate that what we observe is not information being observed slowly by peers or lingering

forum sentiment, but rather peer effects in the absence of any additional information. The results are presented in Column (3) in Table 2.1. We find no statistically significant correlation from the randomly selected cohort, which rules out longer lags in market conditions manifesting as peer effects in user sentiments. We also run several robustness checks validating that peer effects persist if we select a cohort of peers that posts initially over three days before the author’s post ($t - 1$). We observe that peer effects persist in this truncated sample, further invalidating the slow spread of information argument in favour of our peer effects estimation.

Longer time horizon – GME short squeeze and beyond A final question is whether the effects we study persist outside of our selected timeframe. We test our approach on the data from Semenova et al. (2024), which offers a time horizon spanning through mid-2021.⁹ We observe that through the end of the GameStop short squeeze, our estimate for peer effects is close to that of this paper. However, after late February 2021, peer effects and extrapolation no longer appear significant. The individual’s own perspective about the stock, $\Phi_{i,j,t-1}$, not only remains significant but its coefficient also increases in magnitude. Full results are presented in Appendix 2.C. This is consistent with the findings of Bradley et al. (2021) documenting that the informational content and quality of the forum steeply declined after GME. Our analysis points to the fact that after GME, individuals ceased to update their perspectives based on that of their peers or the market, but became firmly entrenched in their own beliefs. This upholds the narrative that WSB transformed from an investor discussion platform into a social movement.

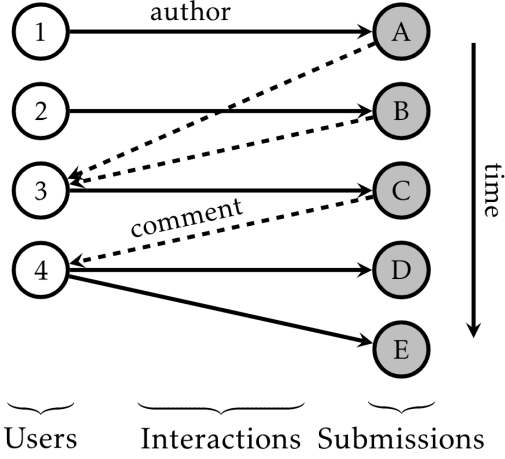
2.3.2 Identifying peer influence – Commenter Network

The *Commenter Network* approach considers a submission-to-submission network, with an earlier submission exerting peer influence on a future submission if the author of the later submission commented on the earlier one. The submission-to-submission network helps identify peers an author interacts with more precisely. Here, we also control for market moves, and employ a set of IVs to address endogeneity concerns. As our IVs, we measure: i) sentiments of submissions to which the influencing submission is connected (the ‘friends of friends’ – detailed in Figure 2.3b), and ii) the historic sentiment of neighbours. The underlying argument rests on the premise that neighbours of network distance two exert an influence on user sentiments through peer effects (consistently with Bond et al. (2012)). A user’s endogenous choice to comment on certain posts over others would therefore not account for users one step removed.

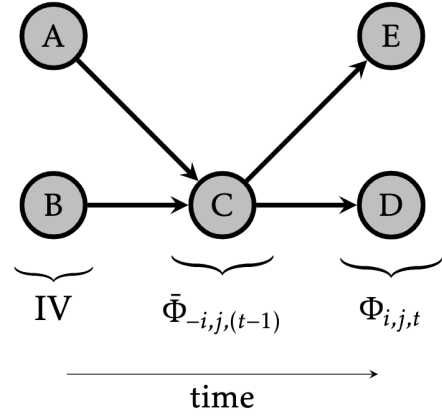
WSB allows us to trace the interactions of users through a commenting network, even though the forum does not support friendship ties (see Appendix 2.A.1). We exploit a submission-to-submission interaction network for each asset, tracking which submissions in the past influence future submissions based on authors’ commenting histories. This method

⁹It is worthwhile to note that our data through 2020 is not identical to theirs due to their selection of a different subset of tickers from AlphaVantage and due to different data processing methodologies.

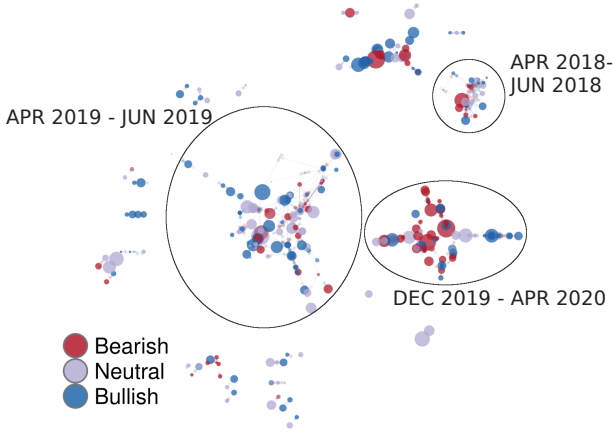
offers a more precise way to identify a user's peers by observing which individuals, and submissions, an author explicitly interacts with. Figures 2.3a and 2.3b illustrate the approach.



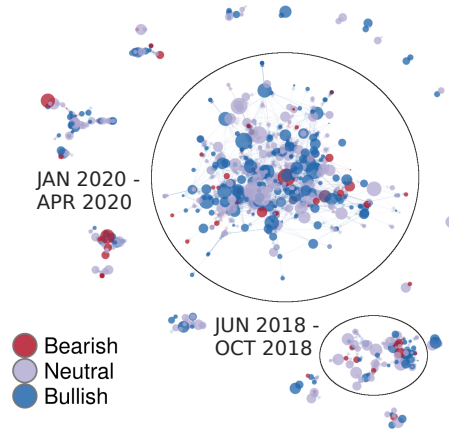
(a) Bipartite network between authors and submissions



(b) Submission-to-submission projection of network in Figure 2.3a



(c) Submission-to-submission network: DIS



(d) Submission-to-submission network: MSFT

Figure 2.3: User networks in WSB conversations: WSB data is summarised as a bipartite graph, illustrated in Figure 2.3a, where users (left) are linked to submissions (right) when they author the submission (solid edge) or comment on the submission (dashed edge). The resulting projection of submissions, in Figure 2.3b, tracks the propagation of sentiments Φ . The submission-to-submission networks for two stocks in Figures 2.3c and 2.3d reveal that individuals post more submissions that are bullish(bearish) at times when the price of an asset increases(decreases) dramatically, with some visual evidence that similar sentiments tend to cluster.

Two examples of submission-to-submission networks in our data are displayed in Figures 2.3c and 2.3d. Distinct temporal clusters emerge, as a certain asset gains and loses prominence on WSB. Some discussions appear fragmented: the *DIS* discussion in Figure 2.3c, for example, contains several smaller clusters, with perceptible differences in overall sentiments. Others, such as the *MSFT* discussion in Figure 2.3d, contain a giant component where investors with different sentiments interact.

Our network approach uses a similar Reduced Form and Second Stage to the *Frequent*

Posters approach in Eq. 2.3. We modify our control for an author’s past sentiment about the stock to account for authors who post for the first time: a dummy variable encodes whether the author’s most recent previous post is bearish, neutral, bullish or missing.

Instrumenting peer sentiment We use an IV approach to estimate peer influence. As the First Stage, we estimate the sentiments of neighbours to estimate an author’s view. As indicated in Figure 2.3b, the sentiments in submissions **A**, **B** can be used to predict that of submission **C**. The *predicted* sentiment of **C** can then, in turn, be used to predict the sentiments of **D** and **E**. This choice of IV is well-established in the networks literature (Zenou 2016, Patacchini & Zenou 2016, Bifulco et al. 2011), and helps control for the exogenous choice to comment on certain submissions and not others – the method further benefits from the fact that users are not alerted to the activity of specific other individuals. We also include the neighbour’s own historical sentiment, as a set of categorical variables, as the second IV (similarly to the *Frequent Posters* approach).

Timing of observations We use the timings of events to mitigate the common shock problem for both our IVs: the neighbour’s historical sentiment and the ‘friends or friends’ submissions. For the latter, we calculate the time period of influence for a given post, which ends when the last comment is made on a submission. This effectively marks the point when a particular submission ceases to be of interest to the WSB community. We filter for instances where the period of influence for a submission used as an IV for another submission ends before the new submission we are modelling is created. In practice, if submission **C** in Figure 2.3b occurs on July 1st at 2:31PM, the final comments on posts **A** and **B** must occur before, in order to ensure that our IV is not affected by a common shock. We also include an author’s own, historical sentiment as an IV only if his previous submission occurs at least two business days before the current one.

The *Commenter Network* offers certain upsides, but also certain shortcomings, as compared to the *Frequent Posters* approach. The network method more precisely identifies the channels of influence between authors. However, the allocation of peers is no longer random, since the network structure is governed by a *choice* to comment.

Results

The results of the *Commenter Network* approach are consistent with those of the *Frequent Posters* approach, presented in Table 2.3 – column (1) presents OLS estimates for using average, observed peer sentiment, while column (2) uses predicted sentiment from our IV approach. In both cases, the robust standard errors, clustered at the ticker level, produce estimates statistically significant at the 1% level. First Stage estimates for estimating peer sentiment using their neighbours are presented in Table 2.2, while the estimates for using the neighbour’s own historical sentiment are presented in Appendix 2.C.

The estimated coefficients in both approaches suggest that an exogenous increase in average peer outlook appears to increase an investor’s own future view about an asset. These findings suggest that the data supports a model where strategic complementarities govern

Table 2.3: Peer influence: *Commenter Network* – full regression estimates

		Dependent Variable – $\Phi_{i,j,t}$		
		Reduced Form (1)	Full Second Stage (2)	Random Network (3)
Independent Variables	$\phi_{i,j,t-1}^{-1}$	-0.226 (0.026) ***	-0.215 (0.021) ***	-0.342 (0.040) ***
	$\phi_{i,j,t-1}^0$	0.047 (0.023) **	0.034 (0.022)	0.073 (0.036) **
	$\phi_{i,j,t-1}^{+1}$	0.160 (0.028) ***	0.141 (0.031) ***	0.244 (0.042) ***
	$\bar{\Phi}_{-i,j,(t-1,t)}$	0.041 (0.009) ***	0.022 (0.009) **	0.009 (0.009)
	$r_{j,t}$	0.020 (0.003) ***	0.023 (0.006) ***	0.031 (0.005) ***
	$\bar{r}_{j,t}$	0.005 (0.003)	0.006 (0.006)	0.008 (0.005)
	$\sigma_{j,t}^2$	0.044 (0.289)	0.512 (0.508)	0.078 (0.449)
	Ticker Fixed Effects	Yes	Yes	Yes
No. Observations:		24,902	16,514	25,220
R^2 :		0.09	0.07	0.09
R_{adj}^2 :		0.06	0.06	0.06

Notes: The dependent variable is individual investor sentiment about an asset expressed in a single submission, scaled continuously between $(-\infty, \infty)$. We estimate it using the individual's previously expressed sentiment about the same asset ($\phi_{i,j,t-1}$) as a categorical variable, with the author not having posted previously ($\phi_{i,j,t-1}^{NA}$) as the baseline. We control for a set of market control variables ($r_{j,t}, \bar{r}_{j,t}, \sigma_{j,t}^2$). The sentiment of posts that the author commented on previously ($\bar{\Phi}_{-i,j,t-1}$) is estimated several ways. In column (1), we present the estimate using the sentiment of posts the author previously commented on. In column (2), we use an IV to predict the sentiment of posts the author comments on. In column (3), we randomly rewire the network, connecting the author to a random set of posts about the same ticker. Robust standard errors, clustered at the ticker level, are presented in parentheses. Observations with incomplete market data are dropped.

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

the investment decisions of retail traders sampled on WSB and where extrapolation plays an important role.

Support for identification – Commenter network

One potential concern is that individuals who comment on other's posts may differ from the rest of the population on the forum. If this were the case, our findings would not allow us to draw valid conclusions about the overall population of investors. We provide evidence that sentiments expressed by our sample are similarly distributed to those of the overall user population in Appendix 2.C.

This approach is more comprehensively discussed with the networks literature, so we rely on the established robustness checks. If our controls are useful in the *Commenter Network* formulation, a random rewiring of the network should yield no effect. The results are presented in column (3) of Table 2.3: no statistically significant correlation emerges from the randomly selected cohort. This provides further evidence that unobserved factors influencing within-ticker variation in both peer composition and author sentiment are not confounding.

An additional concern with our *Commenter Network* approach is overidentifying restrictions. A J-statistic of 0.43, and a corresponding p-value of 51%, leads us to believe that our

additional instruments are exogenous (see Appendix 2.C for further details).

2.3.3 Further insights

WSB data provide additional opportunities to test investor responses to a market surprise, and the reinforcement mechanism between peers and asset prices. We consider the sentiments expressed by investors i about asset j at time t , $\Phi_{i,j,t}$, as our dependent variable and use the controls from Eq. 2.3 to test for two additional effects: market surprise and reinforcement.

We define two types of surprises: i) a positive surprise if asset j experiences a return which is two standard deviations higher than the 30-day historical average for the stock on day t or on the day before, and ii) a negative surprise if asset j experiences a return which is two standard deviations lower than the average for the stock on day t or on the day before. We compute the average and standard deviation for stock j using data of the thirty trading days before t . We also interact returns and predicted peer sentiment to see the extent to which peer effects are reinforced by returns. We use the *predicted* peer sentiment from our *Frequent Posters* approach in our regressions to control for sentiments that respond to current price changes.

Table 2.4: Additional effects: surprise and reinforcement

		Dependent Variable: $\Phi_{i,j,t}$	
		(1)	(2)
Independent Variables	$\hat{\Phi}_{-i,j,(t-1,t)}$	0.037 (0.009) ***	0.036 (0.012) ***
	$r_{j,t}$	0.019 (0.006) ***	0.017 (0.005) ***
	Positive Surprise	-0.018 (0.031)	
	Negative Surprise	-0.114 (0.038) ***	
	$(r_{j,t} \times \hat{\Phi}_{-i,j,(t-1,t)})^+$		0.074 (0.025) ***
	$(r_{j,t} \times \hat{\Phi}_{-i,j,(t-1,t)})^-$		0.070 (0.108)
	Author & asset controls ($X_{i,j,t}$)	Yes	Yes
	No. Observations:	11,073	11,116
R^2 :		0.08	0.08
R^2_{adj} :		0.06	0.06

Notes: The dependent variable – individual investor sentiment about an asset, scaled continuously between $(-\infty, \infty)$ – is estimated using the variables in Eq. 2.3 and additional variables, using OLS. The additional variables in column (1) are categorical variables for positive and negative market surprises at time t in asset j ; in column (2) the additional variables are a cross term between asset j 's returns and the estimated sentiments of peers: $(r_{j,t} \times \hat{\Phi}_{-i,j,(t-1,t)})^+$ is the product if the predicted sentiment of peers is positive and returns are also positive, and zero otherwise; $(r_{j,t} \times \hat{\Phi}_{-i,j,(t-1,t)})^-$ is product if the predicted sentiment of peers is negative and returns are also negative, and zero otherwise. $(r_{j,t} \times \hat{\Phi}_{-i,j,(t-1,t)})^+$ captures the extent to which positive peer predictions correspond to observed market moves; the reverse is true for $(r_{j,t} \times \hat{\Phi}_{-i,j,(t-1,t)})^-$. Peer sentiment $\hat{\Phi}_{-i,j,(t-1,t)}$ is estimated using the *Frequent Posters* approach to control for confounders. Robust standard errors, clustered at the ticker level, are presented in parentheses. Observations with incomplete data are dropped.

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Surprise Table 2.4 presents the results from our exploration of surprise and reinforcement. Column (1) contains the OLS estimates when including positive and negative categorical variables for market surprise. A negative market surprise appears to significantly affect investor sentiments. The result is not symmetric – a positive surprise does not appear to convince investors of the upside potential of a stock. This observation suggests that downside panic spreads quickly within the investor population. This effect is in addition to the significant impact returns have on sentiment.

Reinforcement Column (2) considers the effect from market reinforcement of peer sentiments by including the cross term between returns and the predicted sentiments of peers. The cross terms are separated depending on whether the predicted peer sentiment $\hat{\Phi}_{-i,j,(t-1,t)}$ is positive or negative: $(r_{j,t} \times \hat{\Phi}_{-i,j,(t-1,t)})^+$ is the *bullish* interaction when $\hat{\Phi}_{-i,j,(t-1,t)}$ and returns are both positive and zero otherwise, whereas the bearish interaction $(r_{j,t} \times \hat{\Phi}_{-i,j,(t-1,t)})^-$ is positive if predicted sentiment and returns take negative values, zero otherwise. Therefore, a large value for the bullish interaction corresponds to peers forecasting positive returns in asset j and the asset j simultaneously experiencing positive returns on the day of author i 's submission.

In Table 2.4, the bullish interaction is highly significant. WSB users are spurred by peers predicting positive returns and subsequently observing the asset outperform in the market, possibly suggesting ‘irrational exuberance’ (Shiller 2005). The reverse is not true for bearish reinforcement.

2.4 Has WSB destabilised markets?

In Section 2.3, we focus on identifying the extent to which investors consider information from peers, as well as recent market moves, while updating their own sentiments about future asset returns. The evidence for peer effects in asset demand is robust in two separate estimation strategies. In this section, we analyse asset demand and returns in order to validate the link between these behaviours and stock moves. Strategic complementarities can impact asset returns in different ways, as illustrated by our stylised model in Appendix 2.B.

We first investigate changes in retail trader asset demand by tracking retail investor trades, using the methodology from Boehmer et al. (2021). We observe that changes in asset interest on WSB explain a significant fraction of changes to retail trading behaviour (Section 2.4.1).

A key challenge to analysing returns is that sentiments and returns are co-determined in equilibrium. If returns are high, sentiments are also high. Conversely, if sentiments are high, buying pressure will also increase returns contemporaneously. Our strategy exploits the heavy tails in the popularity of discussions on WSB for identification and demonstrates that sentiments expressed in submissions move asset prices as a function of their virality – as proxied by their comment count (Section 2.4.2).

In a final empirical test, we identify bubble-like dynamics in assets and verify that WSB users have a significantly greater interest in assets whose price experiences a sharp rise,

but subsequently implodes, as compared to those that have a sustained increase in price followed by price stability (Section 2.4.3).

2.4.1 Evidence of trading

We test for a link between discussions on WSB and changes to retail investor demand, approximated by the fraction of retail investor trades executed in the market.

Variable Definition Retail trader activity is identified from Trade and Quote (TAQ) data – a dataset containing all transactions for listed stocks in the United States. We leverage the fact that retail transactions are offered price improvements and, therefore, may execute at a fraction of a penny. To identify retail traders, we first filter trades to those with *exchange code* = ‘D’ in TAQ. In the remaining trades, we identify those that execute at a fraction of a penny as retail transactions. Specifically, let $Z_{j,t} = \text{mod}(100 * P_{j,t}, 1)$, the fraction of a penny associated with the transaction price in stock j at time t . If $Z_{j,t}$ is in the interval $(0, 0.4)$ or $(0.6, 1)$, the transaction is coded as a retail transaction. We define a metric for retail trade fraction in asset j in week t_w as:

$$F_{j,t_w}^R = V_{j,t_w}^{\text{retail}} / V_{j,t_w}^{\text{total}},$$

where $V_{j,t_w}^{\text{retail}}$ is the sum of the sizes of all trades labelled as retail transactions using the method above in asset j in week t_w , and V_{j,t_w}^{total} is the sum of the sizes of all trades in asset j in week t_w from TAQ. Our variable of interest – the change in retail investor trading fraction – is defined as:

$$\Delta F_{j,t_w}^R = F_{j,t_w}^R - F_{j,t_w-1}^R.$$

We design a similar monthly metric, $\Delta F_{j,t_m}^R$, to track changes in retail trade volumes on a monthly scale. Importantly, a change in news or asset-level characteristics cannot justify a change to $\Delta F_{j,t_w}^R$, since it would affect retail traders and institutional investors. Our metric, therefore, allows us to distinguish between changes to *retail investor preferences* versus overall market shifts.

Our goal is to consider whether discussions on WSB explain variation in the fraction of retail investor transactions in the market. We define a metric tracking the prevalence of discussions about asset j in week t_w on the forum versus the overall number of posts about assets on the forum, F_{j,t_w}^W , defined as the number of posts mentioning asset j in week t_w over the total number of posts mentioning assets in week t_w . Our predictor of interest is defined as:

$$\Delta F_{j,t_w}^W = F_{j,t_w}^W - F_{j,t_w-1}^W,$$

the change in the fraction of posts discussing asset j on the forum.

Analytic Approach Our goal is to demonstrate that changes in discussions among retail investors are accompanied by changes in retail trading volumes. We regress change in weekly (monthly) retail trade fractions on the changes in ticker importance on the forum:

$$\Delta F_{j,t_w}^R = \beta_v \Delta F_{j,t_w}^W + \eta_j + \epsilon_{j,t_w}, \quad (2.5)$$

where η_j are ticker fixed effects, and ϵ_{j,t_w} is an error term. We repeat the same exercise, except with variables computed at the monthly time scale, t_m .

Table 2.5: Retail trade volume versus the WSB discussion

<i>Dependent variable: $\Delta F_{j,t}^R$</i>						
	Weekly Changes			Monthly Changes		
	(1) Pooled	(2) Pooled	(3) Top 25% (market cap)	(4) Pooled	(5) Pooled	(6) Top 25% (market cap)
$\Delta F_{j,t}^W$	0.098*** (0.019)	0.098*** (0.019)	0.167*** (0.030)	0.261*** (0.049)	0.253*** (0.049)	0.505*** (0.033)
Ticker FE	No	Yes	No	No	Yes	No
Observations	3,503	3,503	1,040	776	776	232
R_{adj}^2	0.007	0.000	0.028	0.035	0.011	0.116

Notes: This table presents the OLS estimates for the influence of changes in WSB discussion interests on retail trading patterns. Columns (1), (2), (3) present the weekly estimates, while (4), (5), (6) present the monthly ones. Columns (3) and (6) consider the estimates for the top quartile of stocks, by market cap, within our sample. All standard errors are clustered at the ticker level.

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Results We propose that tickers that are popular on WSB are more likely to be linked to changes in retail trader order flow. For this reason, we look at the trading patterns of the twenty most popular tickers, by year, on WSB between the years 2017 and 2020. Table 2.5 presents our main result. Changes in discussion popularity of tickers are statistically significant for explaining changes in retail trading behaviour at the weekly and monthly level. The monthly estimates appear more significant and help explain a greater variation in the dependent variables, as per the R_{adj}^2 .

In columns (3) and (6), we repeat the exercise but only consider the stocks that are within the highest quartile by average market cap between the years of 2012-2020 within our sample: AAPL, AMZN, BA, BAC, DIS, FB, GE, GILD, MSFT. We observe that the coefficients on $\Delta F_{j,t}^W$ at the monthly and weekly time frame are more significant in this formulation and explain a greater fraction of the variation in changes in retail investor trading activity. At the monthly time horizon, 10% increase in the prevalence of AAPL discussions on WSB is associated with a 5% increase in retail trading activity in the market. Interestingly, the stocks within the top quartile do not include ‘meme’ stocks - indicating that WSB discus-

sions track retail preferences across a broad spectrum of assets. Including ticker fixed effects in our model specifications for columns (3), (6) decreases the R_{adj}^2 , indicating that the effect cannot be explained by stock-level differences.

Our experiments demonstrate that WSB discussions track changes in retail trading behaviour at the weekly and monthly timescales. These results further justify our modelling approach, highlighting how WSB activity is linked to retail trader demand for assets. We focus on the directional impact of hype investor discussions on assets in Appendix 2.D.2.

2.4.2 Evidence of granularity

Thus far, our paper demonstrates that strategic information complementarities can drive persistence in sentiments and oscillations in returns. Our empirical exercise in Section 2.3 also shows that individual sentiments are not *fully* explained by information from peers or recent returns, as demonstrated by the R^2 in our regressions. There is, therefore, unexplained heterogeneity in individual investor sentiments. An outstanding question is whether these heterogeneous opinions can survive aggregation across peers and impact asset returns. We offer a some theoretical context in Proposition 3 of Appendix 2.B. This section demonstrates that viral content can influence future asset returns; Appendix 2.D.1 offers additional insights into our identification and presents the diversity in influential, viral content identified through our methodology.

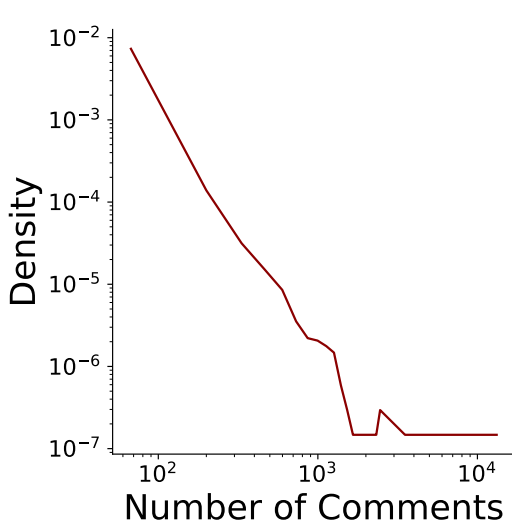
Model framework We remind ourselves of the proposed model framework, which captures the impact of individual, idiosyncratic demand shocks for an asset by investor i , e_i , on price:

$$p = \sum_{i=1}^N s_i \mathbb{E}_i(v) - \gamma \sigma^2 S + \sum_{i=1}^N s_i e_i.$$

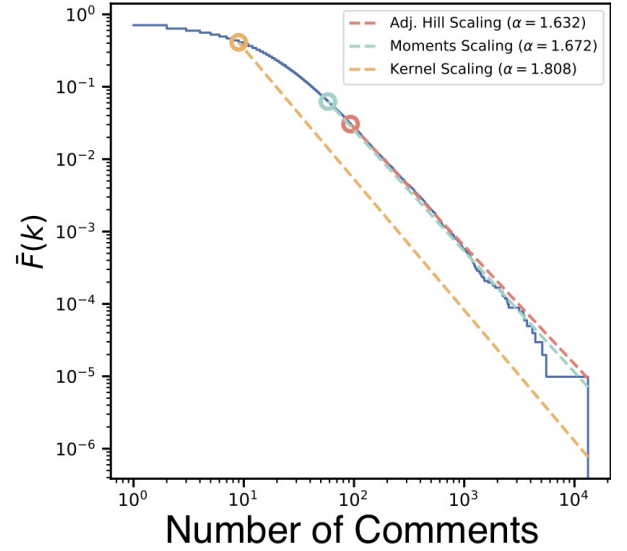
In the absence of ‘granularity’ among investors, the shocks average out to zero. However, when certain investors are weighted differently to others due to differences in capital or popularity, the shocks can have an impact on price, as highlighted above. If there is a granular shock at time t , we would therefore expect a change in the log price, which would manifest in a correlation between the granular social shock and returns, as well as increased volatility.

Granularity of social attention We leverage the heavy-tailed structure of WSB discussions for our empirical strategy. The intuition is that certain submissions gather many more followers than others, which we measure using the number of comments they receive. Appendix 2.A.2 gives a sense for the heavy-tail in discussions *between* assets – a handful of tickers are mentioned in thousands of posts, while most assets receive just a small number of mentions. Attention *within* stocks is also dominated by a few, heavily-commented submissions. In our modification to our model, the heavy-tail of attention can result in unexplained variation in sentiments surviving aggregation and impacting returns. We use a

Granular Instrumental Variable (GIV) approach to investigate the impact on returns (Gabaix & Koijen 2023, 2021, Galaasen et al. 2020).



(a) Distribution of comments across all submissions



(b) Tail estimate for the distribution of comments

Figure 2.4: **Attention is heavy-tailed:** These graphs demonstrate the distribution of comments on submissions that mention a ticker on WSB. The left figure plots the distribution on a log-log scale. The right plots the tail exponent, estimated using the three different methods outlined in Voitalov et al. (2019). The tail exponent is estimated to be less than two across all methods – this implies that the tail distribution obeys a power law, and is heavy-tailed.

We begin by establishing that the *distribution of attention* that information shared by investors online receives is heavy-tailed. We proxy attention by the number of comments that a particular submission receives. Figure 2.4 shows the extreme tail in the distribution of attention – some submissions appear to receive a large following, while the majority are of little interest. We measure the size of the tail in the comment distribution using the approach in Voitalov et al. (2019), who propose several methods for estimating the power-exponent of a distribution’s tail – all methods estimate the tail exponent to be less than two, implying that the tail is power-law and heavy-tailed.

Heavy-tailed attention online implies that idiosyncratic information contained within the most popular submissions potentially persist after pooling across all investor’s opinions, and may have a disproportionate effect on returns. For our identification strategy, we exploit within-ticker-week variation in attention. Therefore, we filter our sample to weeks and tickers where a sufficient number of submissions are made to distinguish between highly popular and less popular sentiments – we choose five submissions for our cutoff. We subsequently test for whether the week in question exhibits granular social attention by fitting a pareto distribution to the popularity of posts within a week and selecting ticker-week combinations where the exponent is less than two.

The approach has certain shortcomings as there are challenges to finding the exponent with a small sample of data, however, allows us to estimate whether the activity in a ticker in

a given week is a good candidate for granular social dynamics. We ensure our result is robust by imposing several different thresholds and varying our approach between applying and not applying the heavy-tail filter; we observe that all results are consistent with the findings presented below.

Preliminary analysis As a preliminary test, we measure the link between granular social activity and volatility. Table 2.6 shows that there is a statistically significant link between the standard deviation of asset j 's returns in week t_w and the existence of granularity in social attention in that week (as defined above). The effect indicates that granularity in social discussions are linked to a 0.8pp increase in asset volatility in a given week, on average.

Table 2.6: Relationship between social granularity and volatility

	<i>Dependent variable:</i> σ_{j,t_w}^2
Granularity Indicator	0.0076 (0.000) ***
σ_{j,t_w-1}^2	0.126 (0.001) ***
r_{j,t_w}	0.000 (0.000)
Week FE	Yes
Number of obs.	2,479,664
Adjusted R ²	0.029

Notes: The dependent variable is the variance in the log returns of asset j in week t_w , σ_{j,t_w}^2 . Explanatory variables include: the variance in log returns in the previous week σ_{j,t_w-1}^2 , the average log returns in the present week, r_{j,t_w} , as well as an indicator for whether we observe social granularity in the week (the indicator equals one if the ticker receives five posts in a given week and a Pareto-fit indicates that the distribution of popularity is heavy-tailed). The sample includes all tickers discussed on WSB since 2016; the sample period begins in 2016 and ends at our WSB cutoff time. Accompanying standard errors, displayed in brackets, are clustered at the stock level.

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

The analysis is only a preliminary indication and does not imply causality. Furthermore, we cannot perform the same test for returns, since the direction of idiosyncratic demand (long or short) is important to quantify. Identifying the link between granular social shocks and returns is tricky due to confounding variables – returns and popularity of online content about an asset may both be driven by news or other market factors. The next step of our empirical approach therefore consists of extracting idiosyncratic social shocks, measured as unexplained idiosyncratic variation in the sentiments of submissions.

Estimates of idiosyncratic social shocks To extract unexplained variation in sentiments, we regress the sentiment expressed in a given post on the return on day t , $r_{j,t}$, the cumulative returns on the week containing day t , r_{j,t_w} , and average sentiments expressed by peers in the prior week $\bar{\Phi}_{j,t_w-1}$. Since our analysis includes weekly and daily variables to better extract idiosyncratic social shocks, we distinguish between t time in days, and t_w time in weeks. For a post about asset j made by author i at time t , we estimate the idiosyncratic information content of the post as $e_{i,j,t}$ in the following regression:

$$\Phi_{i,j,t} = \beta_1^A r_{j,t} + \beta_2^A r_{j,t_w} + \beta_3^A X_j + \beta_4^A \bar{\Phi}_{j,t_w-1} + e_{i,j,t}, \quad (2.6)$$

where X_j are asset fixed effects. These control for the persistence in sentiments for certain assets – for example, the forum’s enthusiasm about TSLA. We note that the results remain similar with and without ticker fixed effects.

The strategy follows the reasoning outlined in Section 2.3.2, Figure 2.C.1. We posit that any news that emerges at time t about a company is assimilated by the market and manifests in returns. Any variation in sentiment that is unexplained by market performance at time t and by past discussions is submission-specific and is idiosyncratic to news and information more broadly available about that stock at that time.

The object of interest, residual $e_{i,j,t}$, is information shared in the submissions that is orthogonal to asset j ’s returns at time t or within week t_w . We, therefore, would expect $e_{i,j,t}$ to have an impact on the market through social forces, rather than through purely informational content. Figure 2.5 plots the distribution of idiosyncratic social shocks. The distribution is somewhat asymmetric and the modal, unexplained sentiment is bullish, but the left tail of discussions demonstrates the presence of intense bearish discourse.

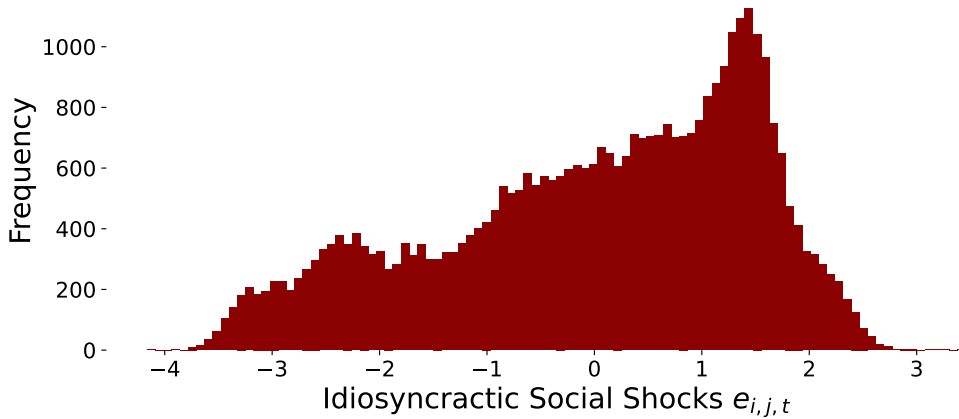


Figure 2.5: **Idiosyncratic social shocks:** This graph plots the distribution of idiosyncratic sentiment heterogeneity, $e_{i,j,t}$, from Eq. 2.6.

Estimating the effect of granular attention In order to assess the impact on asset returns of granular social attention, we proceed by analysing the following relationship:

$$r_{j,t_w+1} = \beta_1^B \bar{e}_{j,t_w} + \beta_2^B r_{j,t_w} + \beta_3^B X_{t_w} + v_{j,t_w}, \quad (2.7)$$

where X_{t_w} are week fixed effects, r_{j,t_w+1} is the cumulative return of stock j in week $t_w + 1$, r_{j,t_w} is the cumulative return in asset j in week t_w , $\bar{e}_{j,t_w} = \sum_i s_{i,j,t_w}^c e_{i,j,t_w}$ the popularity-weighted average idiosyncratic sentiment shared about stock j in week t_w ($\sum_i s_{i,j,t_w}^c = 1$), and v_{j,t_w} is a stock-week specific error. s_{i,j,t_w}^c is calculated by summing the total number of

comments received across all posts in stock j in week t_w – the comment count for a specific submission is then normalised by the total comment count within week t_w . Comment count is re-indexed so that the minimum number of comments a post receives is one. We choose to model future returns in order to mitigate any confounding variables.

A key identification challenge stems from our goal to identify the impact of social forces and popularity of some content over other content, versus general idiosyncratic sentiment on the forum. We rely on a GIV for our identification strategy. The GIV is defined as the difference between popularity-weighted and equally-weighted social shocks, each aggregated for the stock j at period t_w :

$$GIV_{j,t_w} = \sum_i s_{i,j,t_w}^c e_{i,j,t_w} - \sum_i \frac{1}{N_{j,t_w}} e_{i,j,t_w}, \quad (2.8)$$

where N_{j,t_w} is the total number of posts about stock j in week t_w , and i at the authors who are active in the discussion about asset j in week t_w . We subsequently replace $\bar{e}_{j,t_w}$ in Eq. 2.7 with $\hat{u}_{j,t_w}$, where $\hat{u}_{j,t_w}$ is the predicted values from the regression of the GIV on the social shocks $\bar{e}_{j,t_w}$. The outcome variable $\hat{u}_{j,t_w}$ is driven by the popularity of certain posts over others, rather than by general idiosyncratic sentiment.

GIV requirements and threats to identification In addition to capturing the impact of social forces, our specification allows us to disentangle asset properties which affect both sentiment and returns simultaneously. The use of the GIV allows us to mitigate the common shocks problem, where certain stock-specific shocks could affect all idiosyncratic sentiments and future returns – this is similar to [Galaasen et al. \(2020\)](#) in reasoning. Specifically, stock j 's returns in week $t_w + 1$ may be driven in-part by social forces, but also by asset properties which affect both sentiment and returns, which we are unable to control for while estimating idiosyncratic sentiments in Eq. 2.6. A correlation of these shocks with $\bar{e}_{j,t_w}$ may result for a biased estimator for β_1^B . More formally, outcome variable, after imposing controls from Eq. 2.7, y_{j,t_w+1}^r (return variable in week t_w+1 renamed to demonstrate the incorporation of controls), may be of the form:

$$y_{j,t_w+1}^r = \beta_1^B \bar{e}_{j,t_w} + \eta_{j,t_w},$$

where η_{j,t_w} is the ‘common shock’ to asset j in period t_w .

We assume that the idiosyncratic social shock from a post can be expressed as having a stock level component, common to all posts about the stock within that time period, and a post-level component:

$$e_{i,j,t_w} = \beta_j^{cs} \eta_{j,t_w} + u_{i,j,t_w},$$

where β_j^{cs} is the sensitivity of posts within week t_w to the common shock to stock j .

The identification strategy rests on the assumption that the popularity of the idiosyncratic content of posts $s_{i,j,t_w}^c e_{i,j,t_w}$ is not correlated with stock-week shocks η_{j,t_w} . More for-

mally, we require $E[u_{i,j,t}\eta_{j,t_w}] = 0$: the idea is that there are social shocks which make certain content on WSB popular over other content, but that is orthogonal to shocks affecting asset j in week t_w . This is not a problem in this setting for several reasons. Firstly, the popularity of a post could potentially be linked to a stock through its informativeness about the asset's price. However, in the creation of our social idiosyncratic shocks, we extract shocks while controlling for returns on the day and the week of the post. Our social shock time series is, therefore, orthogonal to asset returns at time t and in week t_w and is, therefore, orthogonal to new information available to investors at the time. Furthermore, we find post popularity to be uncorrelated with asset returns on day t and week t_w on which the post is made. As a final precaution, we look at the time period during which a post is active on WSB, where the final time that a post is active is the final comment activity on the post (if the post receives no comments, it is simply the time of the post). We remove posts from our sample that receive commenting activity into the week following the post. On WSB we observe data about relatively unsophisticated retail investors where the sentiments of posts at time t about an asset are systematically linked to *negative* future returns (this holds both when we take a raw average and popularity-weighted average sentiment). This is additional proof that investors we observe do not have access to information on stock-level shocks. Finally, both our shock and popularity time series are constructed at time t_w , while the dependent variable is observed at time $t_w + 1$, avoiding contemporaneity issues.

Table 2.7: Stock returns versus granular social shocks

	Dependent variable: r_{j,t_w+1}				
	Average	Popularity-weighted	Instrumented by GIV: $\hat{u}_{j,t_w}$		
	$\bar{e}_{j,t_w}$ (1)	$\bar{e}_{j,t_w}$ (2)	(3)	(4)	(5)
Granular Social Shock	0.007 (0.005)	0.010*** (0.004)	0.012*** (0.005)	0.013*** (0.005)	0.011** (0.005)
r_{j,t_w}	-0.099*** (0.037)	-0.100*** (0.037)	-0.059*** (0.012)	-0.102*** (0.037)	-0.052* (0.028)
Ticker FE	Yes	Yes	No	Yes	No
Week FE	No	No	No	No	Yes
Observations	2,201	2,201	2,201	2,201	2,201
R^2_{adj}	0.160	0.164	0.013	0.163	0.124

Notes: This table presents the OLS estimates for the influence of idiosyncratic social shocks on WSB. Columns (1) and (2) present the effect of average idiosyncratic shocks and popularity-weighted idiosyncratic shocks, respectively. Columns 3-5 present various specifications of our instrumented idiosyncratic social shocks $\hat{u}_{j,t_w}$. Robust standard errors, clustered at the ticker level, are shown in parentheses. The F-statistic for the first stage regression is 2,297. Observations with incomplete data are dropped.

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Results In order to study the impact of granular social attention, we run the following regression on weekly returns and posts:

$$r_{j,t_w+1} = \tilde{\beta}_1^B \hat{u}_{j,t_w} + \beta_2^B r_{j,t_w} + \beta_3^B X_j + v_{j,t_w},$$

where $\hat{u}_{j,t_w}$ are the *fitted values* of idiosyncratic social shocks on our GIV (we rename our coefficient to $\tilde{\beta}_1^B$ to highlight this), r_{j,t_w+1} is the cumulative weekly return of stock j in week $t_w + 1$, X_j are ticker fixed effects. The formulation closely follows that of our extended model with granular shocks, from Eq. 2.20.

Several patterns emerge from our empirical exercise, presented in Table 2.7. Firstly, we observe that the data appears to follow the structure proposed in our model – idiosyncratic social shocks are positively linked to future returns. In column (4), the estimated effect can be summarised as follows: an estimate for $\tilde{\beta}_1^B$ at 0.013, means that the idiosyncratic doubling in the odds of a very popular post expressing bullish over bearish sentiments (while less popular posts do not express an idiosyncratic sentiment) increases returns in the following week by approximately 0.4 percentage points ($0.013 \times \frac{1}{2} \log 2$), on average. The effect is small, but persists across specifications. Consistently with our model prediction, the average idiosyncratic noise, in column (1), has no effect.

We explore the data in greater detail in Appendix 2.D.1. We consider instances where we observe a large (in magnitude) $\hat{u}_{j,t_w}$ and the types of posts that triggered the viral discussions. We observe that the viral, influential content has very different style. Some posts present a thorough, substantiated discussion backing the author’s claim. Others attract attention through the large risk the author has taken, potentially inspiring others to follow in their footsteps.

2.4.3 Evidence of bubbles

Can we identify Reddit’s bull runs? Greenwood et al. (2019) propose a transparent classification scheme to identify potential bubbles. They determine time windows in which the price indices for various industry market capitalisations grew at a ‘rapid’ rate, constituting a sample of ‘run-ups’. These run-ups are then separated into those whose price levels remained constant, versus those whose price levels crashed (where the rate of increase and price level of the crash are pre-selected). A run-up followed by a crash constitutes an instance in which the index experienced a ‘bubble’.

We adapt the method of Greenwood et al. (2019) to identify large run-ups in stock prices, and test whether activity on Reddit during the price run-up is related to an eventual downturn. The two main variables we use to measure social activity are the number of submissions that mention a given stock in the month preceding a price run-up, as well as the average sentiment expressed in that time window.

The challenge to finding price run-ups for *individual stocks* is that large price swings are more erratic compared to the returns on broad industries considered by Greenwood et al. (2019). Importantly, run-ups in stock prices can be immediate under small market capitalisation, but slower with large market capitalisation. An arbitrary condition on time

and return magnitude thus introduces considerable selection concerns into a sample of run-ups in stock prices. This calls for a more flexible method.

We tweak the methodology from [Greenwood et al. \(2019\)](#), but still rely on their theoretical framework: the identification of price run-ups, in relation to corresponding price crashes. The goal of our method is specifically to find instances when a price run-up either precedes or follows a subsequent price decline.

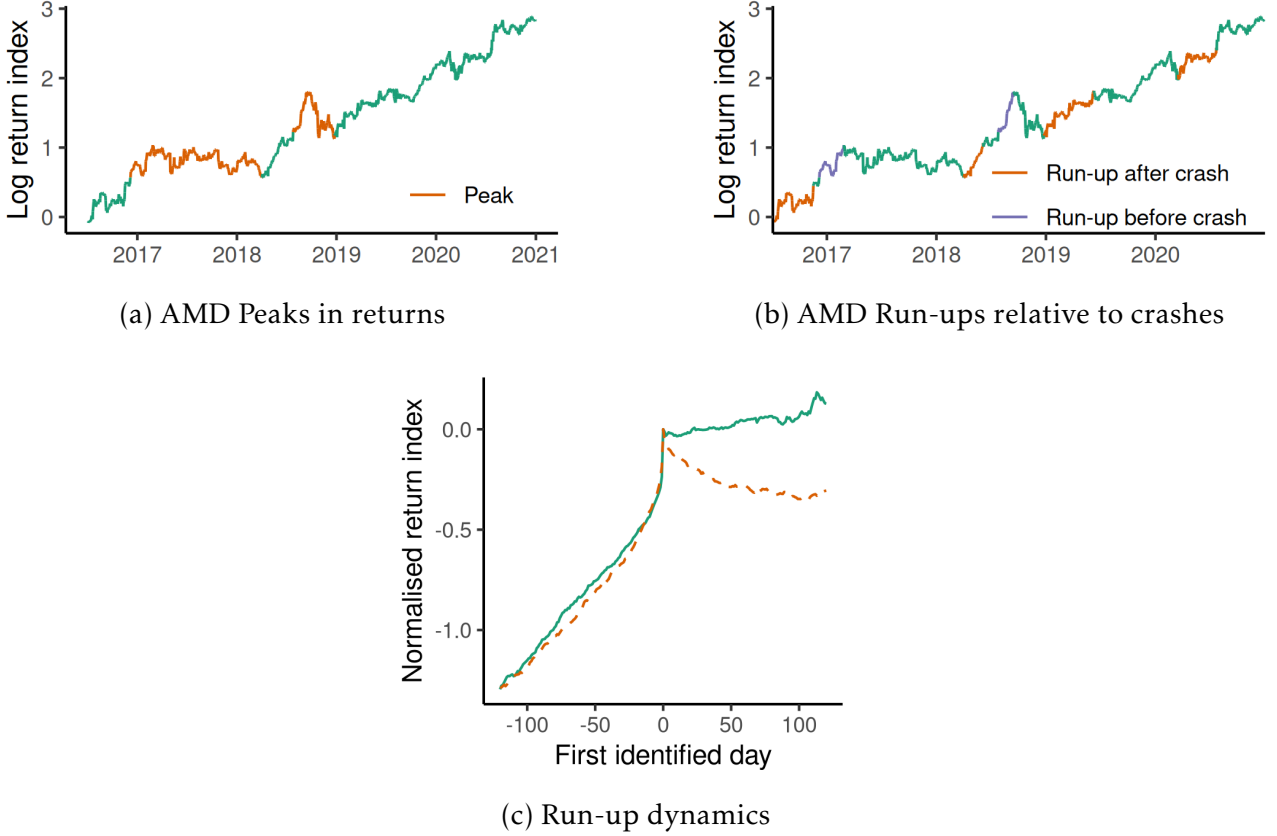


Figure 2.6: **Price run-ups identification:** Figures 2.6a, 2.6b display identified peaks and run-ups in AMD. Figure 2.6c displays run-up dynamics averaged across our sample: returns following a ‘recovery’ run-up, in solid green, remain stable, while they substantially decline after a run-up followed by a crash.

Identifying price run-ups In order to remain as systematic as possible, we identify ‘excursions’ from local minima in the time series of each stock’s log cumulative return index. We define excursions as observations that precede a future minimum, starting from a previous point that matches the cumulative return at that minimum. Specifically, this future minimum is a price level to which the stock does not return at any point following that date. Intuitively, the price of those stocks reverts to a future minimum at the start of the excursion, and thus constitutes a ‘peak’. Here, we filter excursions with a maximum return from the minimum of over 50% as peaks. Figure 2.6a includes two examples of peaks for AMD, one slow run-up starting in late 2016, and an abrupt one in mid 2018.

In addition to peaks, we can also define troughs as excursions from local maxima. An excursion from a local maximum, as opposed to a minimum, is an observation with a cumulative return lower than the highest value in the stock’s history. As such, these function

as ‘inverted’ peaks: a series of negative returns following a high price, which ends when the local maximum is recovered. As with peaks, we isolate troughs with a maximum cumulative return over 50% from the lowest return to the following local maximum.

In summary, both peaks and troughs constitute price swings of over 50% in either direction, one with a period of negative returns (troughs) and one of positive returns (peaks). Peaks and troughs differ in the ordering of price run-ups versus crashes: for peaks, the run-up is followed by a crash, whereas the crash is followed by a run-up recovery period for troughs. In this fashion, we can leverage Greenwood et al.’s (2019) classification of bubbles by studying run-ups that precede crashes (during a peak) and run-ups that follow a crash (during a trough). We illustrate this distinction for the AMD price series in Figure 2.6b, where segments in purple correspond to the run-up periods during bubbles (where peaks are highlighted in orange Figure 2.6a), whereas segments in orange in Figure 2.6b represent run-ups that follow crashes.

We implement this procedure for all stocks traded on the NYSE, NYSE Mkt and NASDAQ, with share codes 10 and 11. In what follows, we restrict the sample of peaks to those after January 1, 2016, and for stocks with an average market capitalisation over one billion USD during the peak. Finally, to harmonise all instances of run-ups with varying lengths of time, we filter out run-ups with an average daily annualised return less than 200% (corresponding to 0.29% daily) to match the criterion used by Greenwood et al. (2019). This leaves us with a sample of 329 run-ups, of which 31 occur during a ‘peak’ period and precede a crash. We visualise the average cumulative return index by type of run-up in Figure 2.6c. Importantly, the trends prior to the turning point are similar for run-ups followed by downturns, and run-ups following downturns.

Table 2.8: WSB features for price runups and crashes

	Run-up after downturn (1)	Run-up before crash (2)	Difference (3)	T test (4)
Log Mentions	0.02 (0.19)	0.19 (0.63)	0.17	15.27
Sentiment	0.02 (0.31)	0.13 (0.52)	0.10	11.41
No. Observations	39,392	3,410		

Notes: This table presents the prevalence of WSB discussions in assets that experience bubble-like dynamics - the mean for each stock sample is shown, with its standard deviation in parentheses. The first column displays that discussions are limited in stocks that are in recovery and do not experience a downturn; however, both mentions and sentiment are higher for stocks that experience a run-up and a subsequent crash, as shown in column (2). A single observation is a day during a run-up period.

Results For each type of run-up – before or after a crash – we count the number of times a submission is made on WSB in the corresponding time window, mentioning the stock experiencing the run-up. Similarly, we compute the average sentiment expressed in these submissions. Table 2.8 summarises our findings, along with the corresponding standard deviation. The striking result is that WSB activity, in both the log of mentions as well as

the average sentiment, features significantly more prominently in run-ups before crashes, rather than after crashes – the difference in means is highly significant.

Are users similarly aware of both types of stocks? One might argue that stocks that experience a crash may have some underlying characteristics which make them of greater interest to the WSB crowd. The share of run-ups preceding crashes during which the corresponding stock is mentioned on WSB, 64.52%, is close to the same share for run-ups following crashes, at 54.36%. The samples are thus similarly represented in terms of discussions in WSB. Figure 2.6c demonstrates that the average price trends for both types of run-ups are nearly identical before the cutoff date, then subsequently diverge.

Summary Our findings demonstrate a link between bubble-like dynamics in the markets and discussions among investors, and prompt additional important questions for research. Recently, rich datasets have become available detailing individual investor portfolios, allowing the study of investor attributes and portfolio choices (Balasubramaniam et al. 2023), individual extrapolation during asset price bubbles (Pearson et al. 2021), and other important characteristics. Combining discussion data and portfolio data offers a promising venue for further research into the profit and loss profiles of hype investors, as well as the behaviour of other market participants faced with hype investor demand.

2.5 Conclusion

We report economically significant estimates for peer effects in expressed sentiments among retail investors on WSB. User sentiments are, on average, 15% more likely to be bullish rather than bearish, if the odds of peers expressing bullish over bearish sentiments double. These results are consistent with the findings of Pool et al. (2015) and Bursztyn et al. (2014). Our observed group of traders on WSB appear to weight the sentiments of peers more heavily than extrapolation, when forming their expectations for future price movements.

We tie observations from WSB to financial markets in three ways. Shifts in attention to different assets on the forum correspond to changes in retail investor trading patterns. In the presence of ‘granular’ heavy-tailed attention (viral content), *heterogeneous* investment decisions are not averaged out and impact returns. Finally, the paper highlights that WSB users exhibit a preference for discussing assets that exhibit bubble-like dynamics.

Even though investor discussion forums have existed for decades, WSB was arguably the first to reach an unprecedented retail following. It succeeded at attracting followers not only through lucrative trade ideas, but also through the promise that coordinating with others would reinforce winning strategies. Our paper offers some degree of evidence that WSB users induced the purchase of assets through popular discourse.

The GameStop short squeeze drew regulators’ attention to investor communities on social media. The potential for price manipulation is cause for concern as retail traders have developed an ability to coordinate investment strategies online. While we cannot address the challenge of legally pursuing actors active on decentralised and anonymous platforms, patterns in unstructured text data should have alerted researchers to the possibility for

market-moving activity stemming from online discussions. Given the recent trends, it is likely that we have only seen the tip of the iceberg, in terms of the impact that social media can have on the market.

Bibliography

- Agrawal, S., Azar, P. D., Lo, A. W. & Singh, T. (2018), 'Momentum, mean-reversion, and social media: Evidence from StockTwits and Twitter', *The Journal of Portfolio Management* **44**(7), 85–95.
- An, L., Lou, D. & Shi, D. (2022), 'Wealth redistribution in bubbles and crashes', *Journal of Monetary Economics* **126**, 134–153.
- Anderson, C. A., Lepper, M. R. & Ross, L. (1980), 'Perseverance of social theories: The role of explanation in the persistence of discredited information.', *Journal of Personality and Social Psychology* **39**(6), 1037.
- Antweiler, W. & Frank, M. Z. (2004), 'Is all that talk just noise? the information content of internet stock message boards', *The Journal of finance* **59**(3), 1259–1294.
- Araci, D. (2019), 'FinBERT: Financial sentiment analysis with pre-trained language models', *arXiv preprint arXiv:1908.10063*.
- Azar, P. D. & Lo, A. W. (2016), 'The wisdom of Twitter crowds: Predicting stock market reactions to FOMC meetings via Twitter feeds', *The Journal of Portfolio Management* **42**(5), 123–134.
- Balasubramaniam, V., Campbell, J. Y., Ramadorai, T. & Ranish, B. (2023), 'Who owns what? a factor model for direct stockholding', *The Journal of Finance* **78**(3), 1545–1591.
- Banerjee, A., Chandrasekhar, A. G., Duflo, E. & Jackson, M. O. (2013), 'The diffusion of microfinance', *Science* **341**(6144).
- Barberis, N., Greenwood, R., Jin, L. & Shleifer, A. (2018), 'Extrapolation and bubbles', *Journal of Financial Economics* **129**(2), 203–227.
- Bifulco, R., Fletcher, J. M. & Ross, S. L. (2011), 'The effect of classmate characteristics on post-secondary outcomes: Evidence from the Add Health', *American Economic Journal: Economic Policy* **3**(1), 25–53.
- Black, F. (1986), 'Noise', *The Journal of Finance* **41**(3), 528–543.
- Boehmer, E., Jones, C. M., Zhang, X. & Zhang, X. (2021), 'Tracking retail investor activity', *The Journal of Finance* **76**(5), 2249–2305.
- Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D., Marlow, C., Settle, J. E. & Fowler, J. H. (2012), 'A 61-million-person experiment in social influence and political mobilization', *Nature* **489**(7415), 295–298.

- Bordalo, P., Gennaioli, N., Kwon, S. Y. & Shleifer, A. (2021), 'Diagnostic bubbles', *Journal of Financial Economics* **141**(3), 1060–1077.
- Bouchaud, J.-P. & Potters, M. (2003), *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management*, 2 edn, Cambridge University Press.
- Bradley, D., Hanousek Jr., J., Jame, R. & Xiao, Z. (2021), Place your bets? The market consequences of investment research on Reddit's WallStreetBets, Technical report, SSRN.
- Bramoullé, Y., Djebbari, H. & Fortin, B. (2020), 'Peer effects in networks: A survey', *Annual Review of Economics* **12**, 603–629.
- Breza, E. & Chandrasekhar, A. G. (2019), 'Social networks, reputation, and commitment: evidence from a savings monitors experiment', *Econometrica* **87**(1), 175–216.
- Bursztyn, L., Ederer, F., Ferman, B. & Yuchtman, N. (2014), 'Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions.', *Econometrica* **82**(4), 1273.
- Chen, H., De, P., Hu, Y. J. & Hwang, B.-H. (2014), 'Wisdom of crowds: The value of stock opinions transmitted through social media', *The Review of Financial Studies* **27**(5), 1367–1403.
- Chen, H. & Hwang, B.-H. (2022), 'Listening in on investor's thoughts and conversations', *Journal of Financial Economics* **145**(2, Part B), 426–444.
- Cookson, J. A., Engelberg, J. E. & Mullins, W. (2022), 'Echo Chambers', *The Review of Financial Studies* **36**(2), 450–500.
- Cookson, J. A., Mullins, W. & Niessner, M. (2024), 'Social media and finance', *Available at SSRN 4806692*.
- Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. (2018), 'BERT: Pre-training of deep bidirectional transformers for language understanding', *arXiv preprint arXiv:1810.04805*.
- Duflo, E., Dupas, P. & Kremer, M. (2011), 'Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya', *American Economic Review* **101**(5), 1739–74.
- Epple, D. & Romano, R. E. (2011), Peer effects in education: A survey of the theory and evidence, in 'Handbook of social economics', Vol. 1, Elsevier, pp. 1053–1163.
- Gabaix, X. & Koijen, R. S. (2021), In search of the origins of financial fluctuations: The inelastic markets hypothesis, Technical report, National Bureau of Economic Research.
- Gabaix, X. & Koijen, R. S. (2023), 'Granular instrumental variables', *Journal of Political Economy* (forthcoming).

- Galaasen, S., Jamilov, R., Juelsrud, R. & Rey, H. (2020), Granular credit risk, Technical report, National Bureau of Economic Research.
- Galeotti, A., Golub, B. & Goyal, S. (2020), ‘Targeting interventions in networks’, *Econometrica* **88**(6), 2445–2471.
- Greenwood, R., Shleifer, A. & You, Y. (2019), ‘Bubbles for Fama’, *Journal of Financial Economics* **131**(1), 20–43.
- Han, B., Liu, H. & Sui, P. (2023), ‘Social learning and sentiment contagion in the Bitcoin market’, *Available at SSRN* 4543326 .
- Heimer, R. Z. (2016), ‘Peer pressure: Social interaction and the disposition effect’, *The Review of Financial Studies* **29**(11), 3177–3209.
- Hellwig, C. & Veldkamp, L. (2009), ‘Knowing what others know: Coordination motives in information acquisition’, *The Review of Economic Studies* **76**(1), 223–251.
- Hirshleifer, D. (2015), ‘Behavioral finance’, *Annual Review of Financial Economics* **7**, 133–159.
- Hirshleifer, D. (2020), ‘Presidential address: Social transmission bias in economics and finance’, *The Journal of Finance* **75**(4), 1779–1831.
- Hu, D., Jones, C. M., Zhang, V. & Zhang, X. (2021), The rise of Reddit: How social media affects retail investors and short-sellers’ roles in price discovery, Technical report, SSRN.
- Jiang, A. Q., Sablayrolles, A., Mensch, A., Bamford, C., Chaplot, D. S., Casas, D. d. l., Bressand, F., Lengyel, G., Lample, G., Saulnier, L. et al. (2023), ‘Mistral 7B’, *arXiv preprint arXiv:2310.06825* .
- Kuchler, T. & Stroebe, J. (2021), ‘Social finance’, *Annual Review of Financial Economics* **13**, 37–55.
- Lahno, A. M. & Serra-Garcia, M. (2015), ‘Peer effects in risk taking: Envy or conformity?’, *Journal of Risk and Uncertainty* **50**(1), 73–95.
- MacKinnon, J. G. & White, H. (1985), ‘Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties’, *Journal of Econometrics* **29**(3), 305–325.
- Mancini, A., Desiderio, A., Di Clemente, R. & Cimini, G. (2022), ‘Self-induced consensus of Reddit users to characterise the GameStop short squeeze’, *Scientific reports* **12**(1), 13780.
- Patacchini, E. & Zenou, Y. (2016), ‘Social networks and parental behavior in the intergenerational transmission of religion’, *Quantitative Economics* **7**(3), 969–995.
- Pearson, N. D., Yang, Z. & Zhang, Q. (2021), ‘The Chinese warrants bubble: Evidence from brokerage account records’, *The Review of Financial Studies* **34**(1), 264–312.

- Pool, V. K., Stoffman, N. & Yonker, S. E. (2015), 'The people in your neighborhood: Social interactions and mutual fund portfolios', *The Journal of Finance* **70**(6), 2679–2732.
- Ross, L., Lepper, M. R. & Hubbard, M. (1975), 'Perseverance in self-perception and social perception: biased attributional processes in the debriefing paradigm.', *Journal of Personality and Social Psychology* **32**(5), 880.
- Sabherwal, S., Sarkar, S. K. & Zhang, Y. (2011), 'Do internet stock message boards influence trading? Evidence from heavily discussed stocks with no fundamental news', *Journal of Business Finance & Accounting* **38**(9-10), 1209–1237.
- Sacerdote, B. (2011), Peer effects in education: How might they work, how big are they and how much do we know thus far?, in 'Handbook of the Economics of Education', Vol. 3, Elsevier, pp. 249–277.
- Securities & Commission, E. (1963), 'Report of special study of securities markets of the Securities and Exchange Commission', *House Document* **95**(4).
- Semenova, V., Gorduza, D., Wildi, W., Dong, X. & Zohren, S. (2024), 'Wisdom of the crowds or ignorance of the masses? a data-driven guide to WallStreetBets', *Journal of Portfolio Management* **50**(4).
- Shiller, R. J. (1984), Stock prices and social dynamics, Technical Report 2, The Brookings Institution.
- Shiller, R. J. (2005), *Irrational Exuberance: (Second Edition)*, Princeton University Press.
- Shiller, R. J. (2017), 'Narrative economics', *American Economic Review* **107**(4), 967–1004.
- Voitalov, I., van der Hoorn, P., van der Hofstad, R. & Krioukov, D. (2019), 'Scale-free networks well done', *Physical Review Research* **1**(3), 033034.
- Weidmann, B. & Deming, D. J. (2021), 'Team players: How social skills improve team performance', *Econometrica* **89**(6), 2637–2657.
- Welch, I. (2022), 'The wisdom of the Robinhood crowd', *The Journal of Finance* **77**(3), 1489–1527.
- Zenou, Y. (2016), 'Key players', *The Oxford Handbook of the Economics of Networks* **11**.
- Zhu, J., Cui, Y., Liu, Y., Sun, H., Li, X., Pelger, M., Yang, T., Zhang, L., Zhang, R. & Zhao, H. (2021), Textgnn: Improving text encoder via graph neural network in sponsored search, in 'Proceedings of the Web Conference 2021', pp. 2848–2857.

Appendix

2.A Data appendix

2.A.1 Extended description of WSB

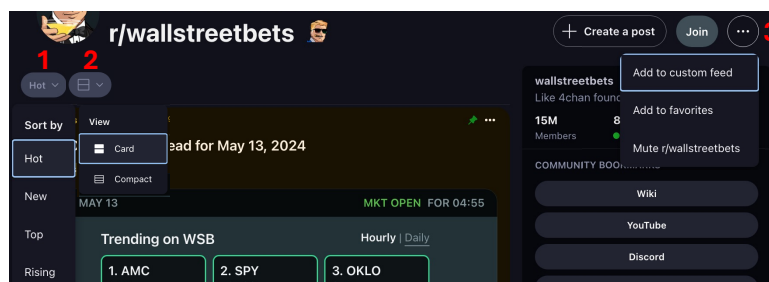
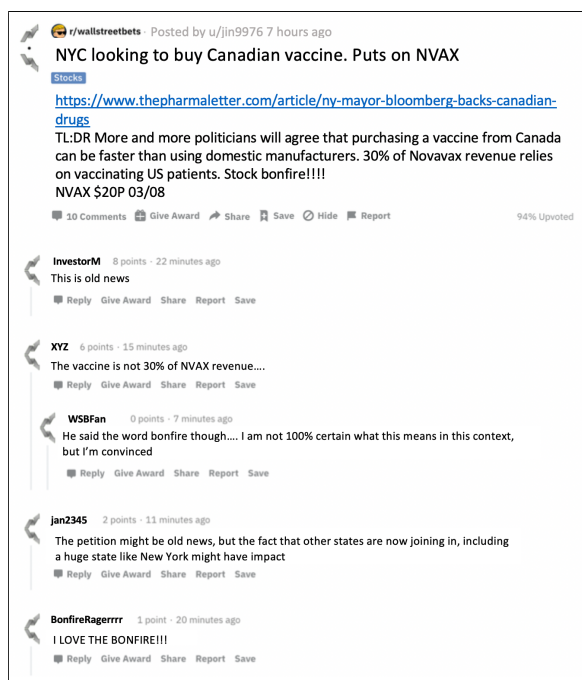
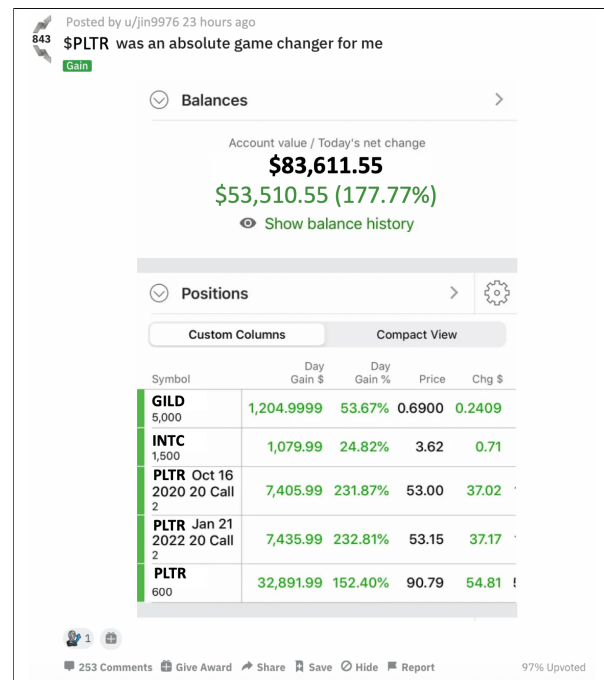


Figure 2.A.1: **WSB Landing Page:** Users have the ability to sort content by 1, the way that posts are displayed by 2 and to add WSB to a feed by 3 (a way of presenting recent content from several forums together).



(a) A typical discussion on WSB



(b) A sample screenshot of user profits

Figure 2.A.2: **What does WSB look like?** These snapshots display typical discussions on WSB. The exact text, usernames, and conversation details have been modified to protect user identities.

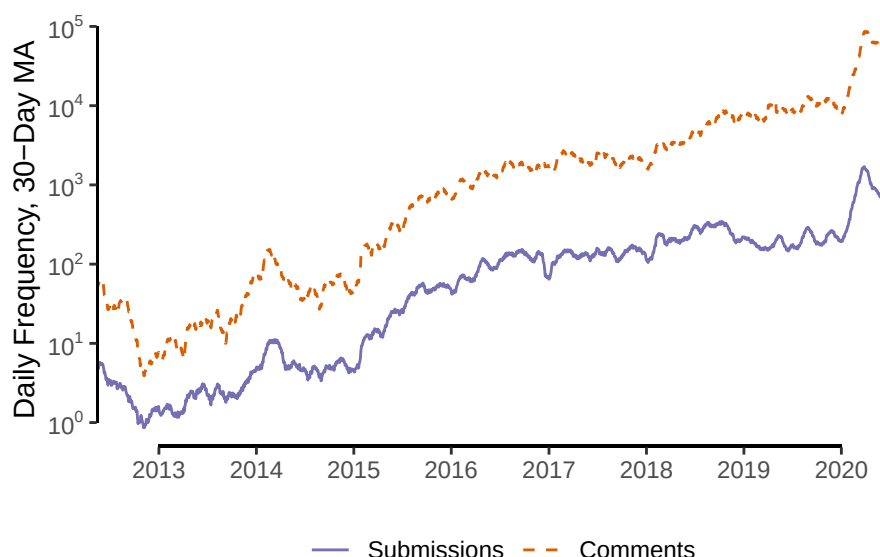


Figure 2.A.3: **Daily activity on WSB plotted on a logarithmic scale:** The daily submission and comment counts, averaged over 30 days, demonstrate a persistent exponential increase in activity on the WSB forum from 2015 to 2020, with a substantial jump in early 2020.

Figure 2.A.2a displays a typical exchange on the WSB forum: individuals discuss stock-related news and their sentiments on whether this will affect stock prices in the future. In addition to market discussions, there is ample evidence of users pursuing the investment strategies encouraged in WSB conversations. Users post screenshots of their investment gains and losses, which subreddit moderators are encouraged to verify, as illustrated in Figure 2.A.2b.

Figure 2.A.1 displays the different options that users have for sorting and viewing content on WSB. It is important to note that the viewing and sorting options are *not* user specific - two users who log on at the same time and choose the same sorting option will be exposed to identical content, regardless of their own posting or commenting history.

Figure 2.A.3 displays the evolution of WSB over time. Two jumps are notable: a smaller, seemingly idiosyncratic rise in early 2018, and a sharp spike during the COVID-19 pandemic.

Reddit user content presentation Our empirical identification strategy in the *Frequent Posters* approach rests on the premises that users are exposed to random variation in peer sentiments. In this section, we discuss the details of how users are presented with content on Reddit. Upon logging into Reddit, users are presented with a ‘home feed’. Historically, the home feed has contained the ‘top posts’ from the subreddits to which a user has subscribed. Subreddit top posts are not individually tailored to the specific user. Users have several sort options based on whether they prefer to see most recent or most highly rated content.¹⁰

More recently, Reddit has implemented an algorithm to try and match users to content

¹⁰https://www.reddit.com/r/help/comments/717686/order_of_posts/

based on a machine learning algorithm, via the home feed.¹¹ However, this change has only taken effect recently – during 2021 and after our sample period.¹² Reddit began to experiment with personalisation seven years ago, however, focused exclusively on gaming subreddits.¹³ However, personalisation is focused on identifying *new* subreddits that a user might be interested in, rather than the way that they are exposed to content from subreddits they already subscribe to. The explicit goal of personalisation is described as an effort to identify new communities users might be interested in:

After 15+ years and millions of feedback comments, survey responses, customer interviews, and Mod Council conversations, we know that whether you’ve been here since the great Digg migration or because you heard about a little community called r/wallstreetbets, we want to help you find communities that you will love on Reddit.

Reddit has, relatively recently, introduced the option to follow individual users, however, following a user means getting exposed to what they post directly to *their own page*, similarly to following an additional subreddit dedicated exclusively to this user. The experience is described as:

Following is just like subscribing to a subreddit, except the subreddit is your profile page. Reddit recently added the ability to post directly to your profile instead of to a specific subreddit. Posts you post to your profile will be seen on a user’s front page feed if they follow you. Outside of that, following does nothing else except your username will be listed in their subscribed subreddits list.¹⁴

The content viewed on WSB would, therefore, remain consistent for all users regardless of their followership, with random temporal variation, across users. Furthermore, given the fact that the anonymity of the forum is of great appeal, followership ties are rare.¹⁵ We consider how pervasive followership relationships are on Reddit by studying which users actually post to their own profiles (the only way to target content directly at followers). We look through all 42,036 authors who create posts about individual tickers within our sample and observe that less than 3% of users create content on their own individual user profile pages prior to our data cutoff time, demonstrating the relative lack of content generated for followers and the insignificance of followership relationships on Reddit. Furthermore, we observe that content posted to WSB user’s own profile (which we retrieve) is generally unrelated to investment – investment advice is typically shared on investment-related forums to reach a targeted audience. We test the sensitivity of our results to the users that post to

¹¹<https://reddithelp.com/hc/en-us/articles/4402284777364-What-are-home-feed-recommendations->

¹²https://www.reddit.com/r/blog/comments/o5tjcn/evolving_the_best_sort_for_reddits_home_feed/

¹³https://www.reddit.com/r/changelog/comments/5zrtmr/testing_community_recommendations/

¹⁴https://www.reddit.com/r/NoStupidQuestions/comments/9dzp9y/what_does_following_someone_on_reddit_do/

¹⁵https://www.reddit.com/r/NoStupidQuestions/comments/71kwqs/do_reddit_users_actually_follow_other_people/

their own profiles remaining in our sample by removing them and rerunning the *Frequent Posters* estimation procedure: the results remain unchanged when these users are removed.

We conclude that individual users were exposed to WSB content based on the content on the forum that was most recent and popular at the time of their logging on, rather than based on personal preference. This, in turn, allows us to assert that users are exposed to random, temporal variation in peer sentiment.

2.A.2 Tickers mentioned on WSB

Table 2.A.1: Most frequent ticker mentions

Ticker	Name	Comments	Submissions	Sum
SPY	S&P 500 Index	291,279	9,408	300,687
AMD	Advanced Micro Devices, Inc.	124,685	5,721	130,406
TSLA	Tesla, Inc.	124,222	5,910	130,132
MU	Micron Technology, Inc.	86,611	3,941	90,552
AAPL	Apple Inc.	48,345	1,880	50,225
AMZN	Amazon.com, Inc.	44,426	1,534	45,960
MSFT	Microsoft Corporation	41,152	1,799	42,951
SNAP	Snap Inc.	40,766	2,043	42,809
NVDA	NVIDIA Corporation	38,012	1,556	39,568
SPCE	Virgin Galactic Holdings, Inc.	30,758	1,640	32,398
FB	Facebook, Inc.	26,143	1,446	27,589
DIS	The Walt Disney Company	25,611	1,088	26,699
BYND	Beyond Meat, Inc.	23,299	906	24,205
NFLX	Netflix, Inc.	20,800	936	21,736
JNUG	Direxion Daily Jr Gld Mnrs Bull 3X ETF	15,761	1,095	16,856
GE	General Electric Company	15,730	929	16,659
RAD	Rite Aid Corporation	14,781	839	15,620
SQ	Square, Inc.	14,003	824	14,827
ATVI	Activision Blizzard, Inc.	13,076	674	13,750
USO	United States Oil	12,949	667	13,616

Notes: This table lists the 20 most mentioned assets on WSB, observed by submissions which uniquely mention the related ticker. ‘Comments’ is the number of comments posted on these submissions, ‘Submissions’ counts submissions, and ‘Total’ is the sum of the two. The name of the asset corresponding to the identified ticker is retrieved from *Yahoo Finance*.

Conventionally, submissions or comments that mention a ticker will spell it using uppercase letters, or following a dollar sign. However, a challenge is that not all uppercase words are valid tickers.

We first match words in WSB submissions to assets by identifying any succession of two to five capital letters. Subsequently, we used a pre-determined list of tickers from CRSP to check whether a match is indeed present in the available financial data. Some abbreviations or capitalised words which are not valid tickers might still show up, such as ‘USD’ (*ProShares Ultra Semiconductors*), ‘CEO’ (*CNOOC Limited*), and ‘ALL’ (*The Allstate Corporation*). Single characters also appear, such as ‘A’ (*Agilent Technologies, Inc.*). We manually created a list of

such tickers, and removed matches featured in WSB submissions, to build a preliminary list of candidate ticker mentions. We refined a second list of candidates by checking whether a collection of one to five letters, lower or uppercase, is preceded by a dollar sign. Any mentions of ‘\$CEO’ or ‘\$a’ count as the tickers ‘CEO’ and ‘A’, respectively. These extracts are, again, checked against the list of available tickers.

A small fraction of the 4,650 tickers we extract dominate the discourse on WSB. 90% of tickers are mentioned fewer than 31 times, and more than 60% are mentioned fewer than five times. The frequency distribution of tail of ticker mentions demonstrates this point. We arbitrarily selected tickers with the number of mentions in the top 10th percentile. Even though threshold of mentions for this top decile is 30 submissions, the most popular, SPY, features in almost 8,000 submissions.

2.A.3 Sentiment modelling in WSB posts

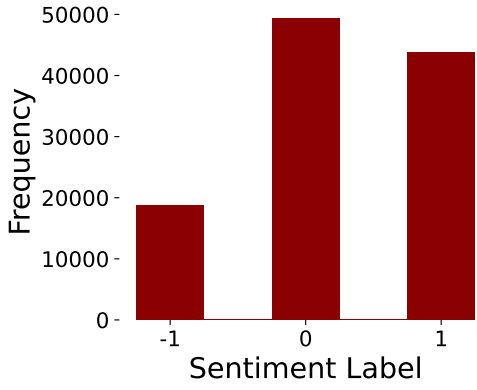
Table 2.A.2: Fine-tuned FinBERT confusion matrix

		Predicted Label		
True Label		-	0	+
	-	64%	28%	7%
	0	6%	77%	17%
	+	6%	27%	67%

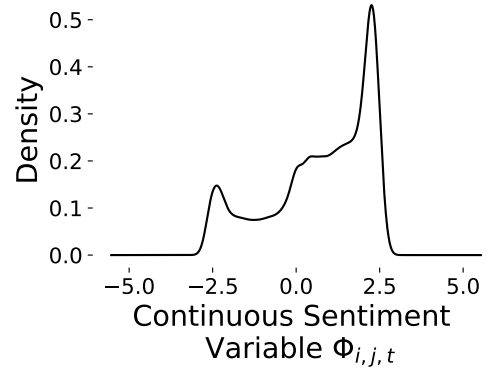
Notes: We use 10% of our hand-labelled data to test the performance of FinBERT on out-of-sample sentiment prediction. The results highlight the model’s ability to predict sentiment with reasonably high accuracy.

Our goal, with regards to the text data in WSB, is to gauge whether discussions on certain assets express an expectation for their future price to rise, the ‘bullish’ case, to fall, the ‘bearish’ case, or to remain unpredictable, the ‘neutral’ case. Among other alternatives, we pursued a supervised-learning approach to identify the sentiment expressed about an asset within a WSB submission. This required a training dataset, for which we manually labelled 4,932 random submissions with unique ticker mentions as either ‘bullish’, ‘bearish’ or ‘neutral’ with respect to the authors’ expressed expectations for the future price. We used the FinBERT algorithm for labelling ([Araci 2019](#)) - a financially oriented modification of Google’s Bidirectional Encoder Representations from Transformers (BERT) algorithm ([Devlin et al. 2018](#)). Work not shown here implements an alternative regression-based approach as a robustness check, but FinBERT performs better out-of-sample.

We trained FinBERT on 75% of the labelled data, and used the remaining 25% for validation and the test set. Table 2.A.2 plots the out-of-sample confusion matrix. For the out-of-sample test, we train FinBERT on 75% of the available data and use 15% for validation; we then compute what the algorithm predicts for the remaining 10% of data. We achieve 70% accuracy on the test set. This is better than a LASSO regression’s accuracy, which was implemented separately and is not covered here.



(a) Distribution of Sentiment Labels - Output from FinBERT Classifier



(b) Continuous Sentiment Variable Distribution

Figure 2.A.4: Distribution of Expressed Sentiments on WSB: We present the density plot of the labelled posts on WSB, and our key continuous variable of log-odds of positive over negative sentiment $\Phi_{i,j,t}$.

Data Description Our final sample contains 111,765 submissions that have a mention of a single, identifiable asset. Figure 2.A.4a shows that the sample is slightly unbalanced - more posts are labelled as neutral than the other two categories, and more posts appear bullish than bearish. The distribution of our continuous variable is shown in Figure 2.A.4b.

2.A.4 Market variables

We include a set of market return and volatility control variables. The data source for these variables are the daily stock files issued by the Center for Research in Security Prices (CRSP), accessed through Wharton Research Data Services.

Market variables in Sections 2.A.6&2.3 The following market variables serve as controls.

$r_{j,t}$: the log return for asset j on trading day t . From CRSP, we calculate it using their ‘RET’ variable: $r_{j,t} = \log(1 + RET_{j,t})$, which automatically corrects the percentage change in closing prices for share splits and dividend distributions.

$\bar{r}_{j,t}$: the average log returns for asset j in the five days prior to t (the log return on day t is not included). A minimum of three daily log-return observations is required, otherwise the observation is set as missing.

$\sigma_{j,t}^2$: the variance of log returns for asset j in the five days prior to t (the log return on day t is not included). A minimum of three daily log-return observations is required, otherwise the observation is set as missing.

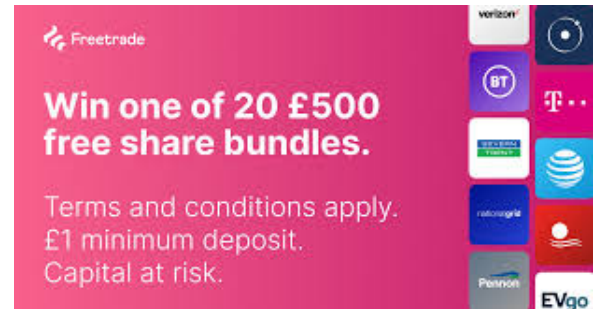
Matching submission timings to trade timings If a post occurs before 16:00:00 EST on day t , we match it with the log-return on the same day t . If a post occurs after 16:00:00 EST on a given day, we match it with market data for the next trading day, $t + 1$. This is done to capture the fact that many news announcements occur after hours and someone posting after the market close may be exposed to these after-hour moves. Instance in which submissions are made on weekends, or holidays, are matched to the next possible trading

day. For example, a submission made at 5pm on Friday is paired to the observed log return for the following Monday.

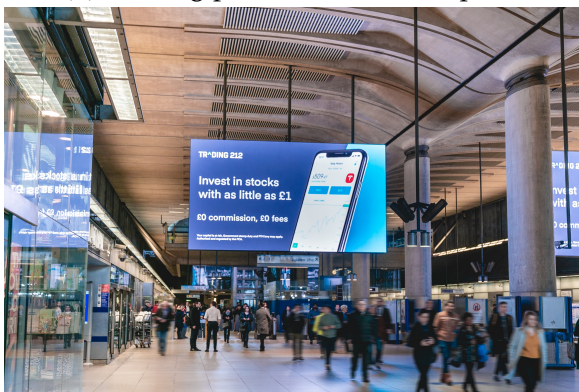
2.A.5 How prevalent are hype traders?



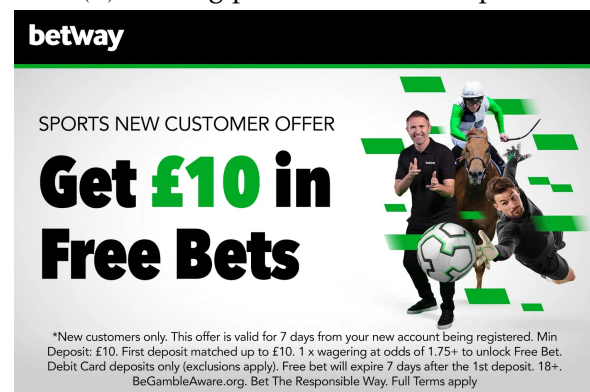
(a) Trading platform ad - example 1



(b) Trading platform ad - example 2



(c) Trading platform ad - example 3



(d) Gambling ad

Figure 2.A.5: **Sample trading and gambling advertisements:** We note the similarity in the style in certain trading advertisement to that of a gambling advertisement - encouraging the retail trader to get ‘hooked’ with an initial free share offering, or to get in with virtually no money in their account.

This study focuses primarily on the WSB discussion forum. However, a potential outstanding question is whether our findings extend to the broader trading population. We present several facts to support the broader relevance of our findings. Firstly, we note that anonymous stock market related forums have skyrocketed in their popularity. WSB, as we had noted previously, has experienced exponential growth and currently boasts fourteen million followers – putting these numbers into perspective, The Times newspaper recently boasted 7.5 million subscriptions.¹⁶ However, retail trader appetite for the hype is not satiated, as new forums are gaining popularity: [r/StockMarketLeakz](#) was one of the top-growing subreddit forums, while [r/personalfinance](#), [r/CryptoCurrency](#), [r/bitcoin](#), [r/stocks](#) are all among the top 100 forums by number of subscribers, as of July 2023.

¹⁶<https://www.nytimes.com/2021/02/04/business/media/new-york-times-earnings.html>

The rise of hype traders has not gone unnoticed as a prospective business opportunity with a tremendous rise in the number of retail trading platforms. Even though many cite long-term investment as a key reason to join, others use an advertisement approach with clear parallels to the gambling industry: offering a free trade or a free stock upon taking up the platform, or to trade with virtually no initial money in their account, as shown in Figure 2.A.5.^{17,18,19,20} Other trading platforms, in turn, provide investors the opportunity to seamlessly execute upon the social and psychological biases studied within this text: ‘Copy-Trader’, for example, on the platform eToro allow investors to automatically copy the trades of others. A recent report by the UK Parliament discusses several important anecdotes pointing to the future relevance of this study: i) UK’s largest investment platform, Hargreaves Lansdown, reported a 40% jump in net new business in the final six months of 2020, ii) the average age of platform users has dropped from 54 in 2012, to 47, reflecting a rise in younger investors (which are more prevalent on discussion forums, such as WSB), iii) Trading 212 (a different platform) announced on 2 February 2021 that it would temporarily pause new account openings due to huge demand.²¹ Even though investor discussion forums have had an influence on markets in the past, evidence points to the fact that the popularity of WSB has permanently altered the composition and behaviour of traders²² and has changed the dynamics of financial markets, as evidenced in this paper.

2.A.6 Preliminary Data Exploration

Isn’t all of this just talk?

Why should we care about the sentiments people express about assets online? Anecdotally, the GameStop short squeeze demonstrated that the online discussions on the WSB forum have an impact on assets. However, this does not constitute evidence of the fact that people follow through on the investment strategies they discuss online *systematically*.

To address the concern that WSB sentiment data has limited impact on investment decisions, we utilise screenshots that users post of their investment positions to test whether they follow through on their expressed sentiments. We extract approximately 9,000 images from WSB – we focus only on image-related URLs (such as ones with the domain name ‘imgur’, an image-hosting site) mentioned in posts of authors who had previously posted about a ticker. We hand-annotate a third of the images. Specifically, we manually annotate the image, if it is a position screenshot, with i) the tickers in the screenshot and ii) the positions (long or short) the author displays. The position taken by author i in asset j , $B_{i,j}$, is annotated as +1 if the author is long in the asset, and -1 if the author is short.

¹⁷<https://www.wallstreetsurvivor.com/robinhood-free-stock/>

¹⁸<https://freetrade.io/legal/ps500-free-share-bundles-april-2022>

¹⁹<https://europe1.discourse-cdn.com/business20/uploads/trading212/original/1X/0d0133c0864441c22fd265d247422b8d5c37e5cb.jpeg>

²⁰<https://talksport.com/sport/betting-tips/1304509/sports-betting-offer-premier-league-racing-free-bets-betway/>

²¹<https://commonslibrary.parliament.uk/the-rise-of-armchair-retail-trading-risks-and-regulation/>

²²<https://markets.businessinsider.com/news/stocks/reddit-retail-investor-trend-survey-social-media-wallstreetbets-wsb-stock-2021-4-1030366600>

Table 2.A.3: Follow-through on WSB Advice

<i>Dependent variable: Position in Asset j of Author i</i>		
	$B_{i,j}$ - categorical	
	(1)	(2)
$\Phi_{i,j}$	0.72 (0.09) ***	
$\phi_{i,j}^{-1}$		-0.97 (0.29) ***
$\phi_{i,j}^0$		0.66 (0.21) ***
$\phi_{i,j}^{+1}$		1.84 (0.27) ***
Observations	278	278
Pseudo- R^2	0.14	0.17

Notes: This table presents estimated log-odds coefficients for two logit models for the relationship between the sentiment expressed by user i about asset j and the subsequent long/short position the user reports (see Eq. 2.9). Sentiment estimates are presented in two ways: (1) the continuous log-odds of the author expressing positive over negative sentiment $\Phi_{i,j}$, and (2) as a categorical variable where $\phi_{i,j}^{-1}$ corresponds to the expression of negative sentiment, $\phi_{i,j}^0$ - neutral, $\phi_{i,j}^{+1}$ - positive.

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

We note that the sample of screenshots is biased. Authors on WSB are socially incentivised to share extreme losses (known as ‘loss porn’ on the forum) or gains. We, therefore, observe relatively few positions, as compared to sentiments. The positions data is also skewed towards long positions, which is consistent with the skew towards bullish sentiments on the forum. However, despite these shortcomings, the positions provide sufficient variety in investment strategies to test whether people trade on their expressed sentiments on WSB.

We match the ticker screenshot to a submission posted before/simultaneously with that screenshot by the same author and about the same ticker. We regress the most recently expressed sentiment by author i about asset j (our key sentiment variable $\Phi_{i,j}$) on the log-odds of the position $B_{i,j}$ extracted from their screenshot being long versus short:

$$\log\left(\frac{P(B_{i,j} = +1)}{P(B_{i,j} = -1)}\right) = \lambda^s \Phi_{i,j} + u_{i,j,t}^p \quad (2.9)$$

where λ^s measures the pass-through rate of sentiment into eventual investment positions. In an alternative formulation, we represent the past sentiment as three categorical variables: $\phi_{i,j}^{-1}$, $\phi_{i,j}^0$, $\phi_{i,j}^{+1}$, which take on a value of one if the author’s sentiment is labelled short, neutral or long, respectively (and a value of zero otherwise).

Results Table 2.A.3 presents the coefficients of past sentiments regressed on future positions, estimated using a logistic regression. We observe that an author’s sentiment is highly correlated to their subsequent holdings of the stock. Let us consider the results in column (2) – an author creating a bullish post about an asset raises the probability of a long versus short investment by over six times. Furthermore, this specification of the logistic regression

predicts an author's position with over 75% accuracy. The positions data gives us confidence that investors do trade based on their discussions and expressed sentiments.

Relationship between stock returns and WSB sentiments

Is WSB activity linked to stock market returns? We run a set of simple exercises to motivate our future analysis. We average the sentiment characteristics in Eq. 2.1 by stock j and trading day t , denoting these mean sentiments by $\bar{\Phi}_{j,t}$. We merge these daily sentiment observations with US common stock returns reported by CRSP (detailed in Appendix 2.A.4), and transform the reported returns into log returns.

We first consider a regression of log returns on mean daily sentiment,

$$r_{j,t} = \lambda_{1,T} \bar{\Phi}_{j,t+T} + \eta_t^r + v_{j,t}^r, \quad (2.10)$$

where $v_{j,t}^r$ is a residual, η_t^r is a daily fixed effect, and T denotes a lag varying from -10 to 15 days. The OLS estimates for coefficients $\lambda_{1,T}$ describe how WSB sentiments are temporally related to stock returns.

Subsequently, we regress daily log returns on current mean sentiments, as well as previous day sentiments:

$$r_{j,t} = \lambda_1 \bar{\Phi}_{j,t} + \lambda_2 \bar{\Phi}_{j,t-1} + \eta_t^r + v_{j,t}^r, \quad (2.11)$$

where η_t^r is a daily fixed effect, $v_{j,t}^r$ an error term, and λ_1, λ_2 our coefficients of interest. This specification gives a sense for the dynamic properties of WSB sentiments – leading up to a trade day, plus their response on that day. In this analysis, we multiply log-returns by 100, for scaling purposes.

Results Figure 2.A.6 plots the OLS estimates for $\lambda_{1,T}$ in Eq. 2.10 as a function of lag T . Generally, past sentiments appear negatively related with current returns, although the effect is small, and not highly significant beyond four lags. Current sentiments are strongly correlated with current and past returns, and this effect is significant for up to five days in the past, before dissipating. This implies that a large return in an asset today will have a persistent impact on investor sentiment for five days into the future. Investor sentiments do not anticipate future returns, but rather follow the trend.

Table 2.A.4 reports OLS estimates for coefficients from Eq. 2.11 in Column 1. Returns again relate positively to contemporaneous sentiment, and negatively with previous day sentiments. Both of these are statistically significant at the 1% level. However, the implied effects are relatively small; on an average day, returns are 0.1 log points lower if sentiments expressed on the previous day are twice more likely to be bullish than bearish.

In Column 2, we estimate Eq. 2.11 with the interaction between lagged, average sentiment and lagged returns, to capture a non-linearity for sentiments in stocks that garner exceedingly high amounts of attention. The slight negative relationship between current and past returns could potentially confound the effect of past sentiment, as seen in the

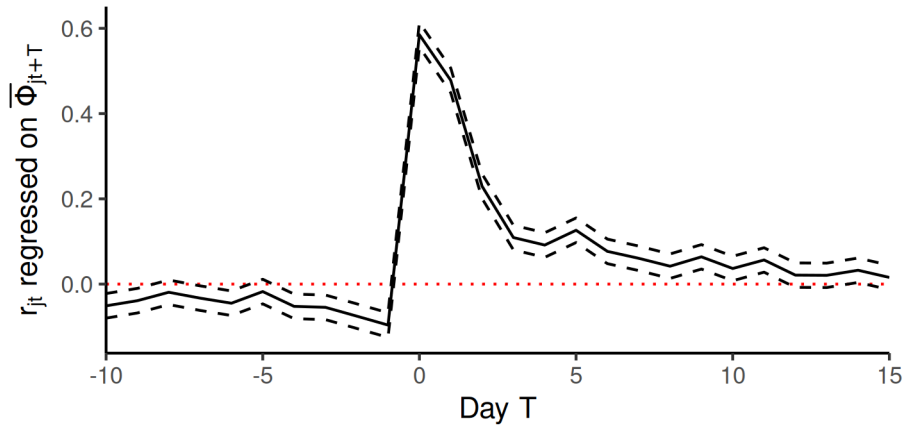


Figure 2.A.6: **Returns correlate negatively with past WSB sentiments, and positively with current and future WSB sentiments:** Daily stock log returns are regressed on daily average sentiments expressed in submissions on that stock on WSB, where sentiments are lagged by -10 to 15 days – the estimated relationship is described in Eq. 2.10. A lag of -5 implies that sentiments precede the returns observation by five days. The point estimate for the correlation is plotted by the solid black line, with the 99% confidence interval in dashed black lines. A correlation of zero is highlighted by the dotted red line. All variables are demeaned by their daily average. Return correlation with past WSB sentiments is negative, but strongly positive with current and future sentiments. Log-returns are multiplied by 100.

Table 2.A.4: Stock returns versus WSB characteristics

	<i>Dependent variable:</i>	
	$r_{j,t}$	
	(1)	(2)
$\bar{\Phi}_{j,t}$	0.60 (0.04) ***	
$\bar{\Phi}_{j,t-1}$	-0.16 (0.02) ***	-0.07 (0.02) ***
$r_{j,t-1}$		-0.06 (0.004) ***
$\bar{\Phi}_{j,t-1} \times r_{j,t-1}$		0.01 (0.01)
Day FE	Yes	Yes
Observations	8,287,639	8,287,639
R ²	0.0004	0.003

Notes: This table presents the OLS estimates for the relationship between stock log returns, $r_{j,t}$ and average expressed sentiment on WSB, $\bar{\Phi}_{j,t}$ and $\bar{\Phi}_{j,t-1}$. t represents time in days. The regression highlights the existence of a positive relationship between current sentiments and current returns, and a negative one between current returns and past sentiments. The negative relationship persists when controlling for previous day returns $r_{j,t-1}$ in Column (2). Accompanying standard errors, displayed in brackets, are clustered at the stock level, and calculated in the manner of [MacKinnon & White \(1985\)](#). Log-returns are multiplied by 100.

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

smaller coefficient for lagged sentiments. However, there is no clear evidence that the interaction between sentiments and outsized returns produce a significant effect on subsequent returns.

2.B Model for social dynamics and asset prices

What motivates investors to share trading advice online, and how does such behaviour affect asset prices? Some seasoned traders might argue that one can only have an investment edge while others are unaware of your strategy. We rationalise observed online information sharing through an asset demand model by incorporating information complementarities in investment decisions. This gives rise to *social contagion* in asset demand; investors buy the asset because others do as well, irrespective of their personal beliefs. The model explores the patterns of returns when this behaviour is present.

Our proposed mechanism is motivated by a growing literature on strategic information complementarities (Hellwig & Veldkamp 2009, Zenou 2016), as well as studies of diagnostic expectations (Bordalo et al. 2021) – we choose to focus on information-sharing, rather than network ties since WSB does not have explicit friendship ties, making the information framework more appropriate for our setting. Investors trade on the momentum of the stock price, against a supply of shares provided by noise traders. Our inclusion of a social component subsequently induces persistence in asset demand over time, which leads to reversal in future returns. The goal of this section is to present what role a social component might play in asset returns and to frame our empirical analysis.

General setup We analyse the price of one asset traded by N investors, indexed by i . Each investor derives CARA utility from consuming c , $U_i(c_i) = -\exp(-\gamma c_i)$, where γ is the constant absolute rate of risk aversion – the model setup is consistent with various behavioural models, models for bubble formation, and is justified by empirical observations on the relationship between investor sentiment and volatility (Bordalo et al. 2021, Barberis et al. 2018). We do not include any discounting in investor decision-making, but assume they evaluate an asset according to a log-normal distributed value $v = \log(V)$ with expectation $\mathbb{E}_i(v)$ and variance $\sigma_i^2(v)$. In the static model, investor i purchases ϕ_i shares at the current market log-price p to optimise the mean-variance objective function

$$\mathcal{L}(\phi_i) = [\mathbb{E}_i(v) - p]\phi_i - \frac{\gamma}{2}\sigma_i^2(v)\phi_i^2, \quad (2.12)$$

$$\Rightarrow \phi_i^* = \frac{\mathbb{E}_i(v) - p}{\gamma\sigma_i^2(v)}, \quad (2.13)$$

where an asterisk denotes the value that maximises objective $\mathcal{L}$. In this way, we distinguish between beliefs about value $\mathbb{E}_i(v)$ from investor i 's decision to buy amount ϕ_i . Eq. 2.13 yields a familiar expression for asset demand in equilibrium: namely a ratio of expected net returns over their variance.

Assuming that asset supply originates from noise traders, S , as in Bordalo et al. (2021), common variance $\sigma^2 = \sigma_i^2(v)$ and averaging expected values $\mathbb{E}(v) = 1/N \sum \mathbb{E}_i(v)$, we can re-arrange Eq. 2.13 to yield the following expression for the market-clearing price p :

$$p = \mathbb{E}(v) - \frac{S}{N}\gamma\sigma^2. \quad (2.14)$$

Eq. 2.14 accounts for the price level by investor's average expected value of the asset, in addition to their ability to absorb the exogenous level of assets supplied. This ability depends on the depth of the investor pool – reflected by the number of investors N – as well as their risk appetite $\gamma\sigma^2$. In this simple market, the price increases with expected value, and decreases with supply. Appendix 2.B.1 further links this model framework to complementarities in asset demand.

2.B.1 Static model with investment complementarities

We study the role of complementary investment decisions. To that end, the model operates in two stages. In the first stage, investors build their expectation for the asset's value, using observed signals from their peers and their expectation of the market-clearing price as a function of the expected, and as of yet hidden, supply shock. In the second stage, the asset supply shock is revealed, and investors execute their trades according to their demand curve.

Investor i who expects value $\mathbb{E}_i(v)$ and understands the price setting mechanism. Investor i 's maximised payoff from Eq. 2.12 is

$$\mathcal{L}(\phi_i^*) = \frac{\mathbb{E}_i^2(v - p)}{2\gamma\mathbb{E}_i(v - p)^2} \quad (2.15)$$

$$= \frac{1}{2}\mathbb{E}_i(v - p)\phi_i^* \quad (2.16)$$

$$= \frac{1}{2}[\mathbb{E}_i(v) - \mathbb{E}(v)]\phi_i^* + \frac{\gamma\sigma^2}{2}\left(\frac{1}{N}\sum_j^N \phi_j^*\right)\phi_i^*, \quad (2.17)$$

where $\mathbb{E}(v) = 1/N \sum \mathbb{E}_i(v)$ as before. Here, investors base their price expectations on the simple equilibrium in Eq. 2.14, but use their personal expectations and constant uncertainty σ^2 to forecast price in the second stage. Eq. 2.17 demonstrates that the investor's payoff depends on their peers in two regards. First, payoffs increase to the degree that the investor in question expects to outperform others, in terms of the value they realise in the asset. This is seen in the first component, by which buying(selling) the asset increases the payoff to the extent that i 's expected value $\mathbb{E}_i(v)$ is higher(lower) than that of their peers. Second, the payoff increase by the average optimal asset demand of all investors in the economy.

The asset demand model predicts that social interactions – knowledge of other's asset purchases – is a significant component of investors' welfare *in expectation*. Eq. 2.17 is a well-known formulation for strategic interactions between agents acting under quadratic loss (Zenou 2016). Deriving Eq. 2.17 with respect to two investors' demands reveals their strategic complementarity:

$$\frac{d^2\mathcal{L}(\phi_i^*)}{d\phi_j^*d\phi_i^*} = \frac{\gamma\sigma^2}{2N} > 0. \quad (2.18)$$

The emergence of strategic complementarities is due to a crowding effect that investors have on price. The higher asset demand by other investors, the higher the realised price will turn

out to be. Before the value of the asset is revealed, investors are motivated to gauge demand by others to better estimate what the price will be, in excess of their personal valuation.

The unweighted average of peer sentiment in Eq. 2.17 emerges because we did not provide a specific mechanism by which information about asset demand is transmitted. The acquisition of information under some cost to the investor is an interesting extension, although already studied by [Hellwig & Veldkamp \(2009\)](#). Their study offers more rigorous insight into the manifestation of strategic complementarities, as well as the emergence of multiple equilibria, when investors seek to learn about the underlying price from a set of possible signals.

Granular shocks An investor's demand may have some idiosyncratic preference which is not captured by common factors. For example, an investor may place particular confidence in products he enjoys using, or admire the corporate strategy of certain company leaders. Such sentiments would manifest as idiosyncratic, heterogeneous investor demand, where e_i is the idiosyncratic component of i 's asset demand. Under the assumption that these shocks have a finite variance and mean zero, they should average out to zero by the Central Limit Theorem. However, consider a scenario where some investors have different levels of importance s_i for aggregate demand:

$$S = \sum_{i=1}^N s_i \phi_i^*, \quad (2.19)$$

$$\Rightarrow p = \sum_{i=1}^N s_i \mathbb{E}_i(v) - \gamma \sigma^2 S + \sum_{i=1}^N s_i e_i, \quad (2.20)$$

where weights $\sum_i s_i = 1$ for demand to equal supply. This could be the case for several reasons: investors could have different amount of capital or, of greater interest to this paper, some investors may have *more sway* in forming public opinion than others. The importance of key players in a social context has been explored in several economic settings – see [Zenou \(2016\)](#) for a thorough overview. Typically, the most central nodes in a social network have the ability to quickly diffuse information and, therefore, have a high influence on others.

We justify this weighting scheme by the fact that certain users on WSB have a disproportionate effect in shaping the broader discourse. Indeed, WSB is structured to promote viral content, and we would expect a consensus to be formed by key players – or, rather, around key submissions. If the distribution of importance does not have a finite variance – i.e. it is ‘heavy-tailed’ – then the idiosyncratic shocks would not average out to zero. We use this framework to evaluate the impact of viral content from WSB on market returns in section 2.4.2.

2.B.2 Equilibrium price dynamics with peer effects

To study the joint dynamics of an asset's price and demand by social investors, we treat aggregate asset demand $\phi = \sum_i \phi_i$ and log price p as state variables for a dynamic system,

indexed by time t . In doing so, we assume that cumulative demand ϕ_t reflects a difference between individual valuations of the asset and the price. We distinguish between two independent components of individual valuations: the private signal of individuals $g(b_{i,t})$ and the signal individuals draw from observations of peers $f(\phi_{i,t})$. Aggregate asset demand and price are

$$\phi_t = \frac{\mathbb{E}_t[g(b_{i,t})] + \mathbb{E}_t[f(\phi_{i,t})] - p_t}{\gamma\sigma^2}, \quad (2.21)$$

$$p_t = \mathbb{E}_t[g(b_{i,t})] + \mathbb{E}_t[f(\phi_{i,t})] - \frac{S_t}{N}\gamma\sigma^2. \quad (2.22)$$

Cumulative demand is, therefore, the difference between individual valuations and the market-clearing price, normalised by risk-aversion. The market-clearing price, on the other hand, is the difference between individual valuations and the rate of asset supply, normalised by the number of investors N and their risk appetite $\gamma\sigma^2$, similar to Eq. 2.14.

A focus of this paper is the relationship between valuations $g(b_{i,t})$ and the social component $f(\phi_{i,t})$. Studies in behavioural finance suggest different expectation formation mechanisms that ultimately deviate from rational expectations (Barberis et al. 2018, Bordalo et al. 2021). We combine two such features in Assumptions 2.1 and 2.2 to propose a testable structure for $\mathbb{E}_t[f(\phi_{i,t})]$ and $\mathbb{E}_t[g(b_{i,t})]$.

Persistent demand The mechanism by which past demand enters current asset demand is by the complementarity in investor payoffs. Investor i 's payoff to holding the asset is assumed to increase linearly in average asset demand by others.

Assumption 2.1 (Persistent Demand). *Social investor i 's expectation of future returns is linearly increasing in average asset demand by others: $f(\phi_{i,t}) = \alpha\phi_{t-1}$, where $\phi_{t-1} = 1/N \sum_i \phi_{i,t-1}$ is average asset demand.*

Mechanical Extrapolation We assume that investors partially trade on the momentum of the asset's price, which Bordalo et al. (2021) term 'mechanical extrapolation'. The functional form of $\mathbb{E}_t[g(b_i)]$ is specified in Assumption 2.2.

Assumption 2.2 (Mechanical Extrapolation). *The average investor projects past price increases into the future using the updating rule:*

$$\mathbb{E}_t[g(b_{i,t})] = p_t + \beta(p_t - p_{t-1}), \quad (2.23)$$

where β captures a fixed degree of price extrapolation.

The validity of these assumptions is discussed in Appendix 2.B.

System for price and demand Combining Assumptions 2.1 and 2.2 into Eqs. 2.21-2.22 yields demands and returns $r_t = p_t - p_{t-1}$:

$$\phi_t = \frac{\alpha\phi_{t-1} + \beta r_t}{\gamma\sigma^2}, \quad (2.24)$$

$$r_t = -\frac{\alpha}{\beta}\phi_{t-1} + \frac{S_t\gamma\sigma^2}{\beta N}. \quad (2.25)$$

In this scenario, asset demand and returns are determined simultaneously. The first mechanism is through market clearing, where demand has to adjust to supply. The second is the adjustment of the expected value for the asset to the realised return, through β , and the social signal, through α .

As a result, returns are accounted for by current and past asset demand:

$$r_t = \frac{\gamma\sigma^2}{\beta}\phi_t - \frac{\alpha}{\beta}\phi_{t-1}, \quad (2.26)$$

since supply must equal demand at time t . This equation uncovers several important mechanisms at play. Returns are related positively to current demand ϕ_t through market clearing – supply S is exogenous and must meet current demand. Higher demand drives up returns. Returns are, however, negatively related to past demand ϕ_{t-1} through the expected value of an asset – a *valuation mechanism*. If there is a positive social signal, investors still value an asset highly, even in the presence of low returns.

To explain the basic intuition, we consider the following scenarios: i) one where the asset has a positive return r_t and no social signal ϕ_{t-1} , and ii) one where investors observe a positive social signal ϕ_{t-1} . In scenario (i), demand is driven by the extrapolation component alone – investors believe that returns will continue to increase based on the current trend. In scenario (ii) on the other hand, investors do not require a large return to demand the asset – positive past sentiment drives current demand ϕ_t . Under exogenous supply, a strong positive signal from peers means that the extrapolated return is *less* important in justifying a higher price. The underlying reasoning relies on the fact that the system is in equilibrium. Therefore, both returns and sentiments have adjusted to reflect a new steady-state, where sentiments are at a certain level ϕ_t .

Finally, we observe that the ratios of the coefficients, γ/β and α/β , play an important role. β effectively anchors the demand of investors in reality – a greater value of β implies that social signals carry less weight, and investors focus on price trends to forecast and expect asset values to grow at some constant rate. As β decreases, returns are determined more by social forces – hype from peers, rather than past performance, now justifies returns and demand. In our data, we observe that α is roughly 1.5 times β – individuals weight the sentiments of peers *more heavily* than recent returns. σ^2 serves to taper the impact of sentiment, since investors are less certain in their signal and demand less of the asset. The model also explains persistent fluctuations in asset demand, as discussed below.

2.B.3 Persistent fluctuations

The reversal in returns is an important feature that emerges from social contagion in investors' price expectations. If large enough, these can produce bubbles in asset prices: initial momentum from positive news creates a price run-up, before an absence of news creates a drought of new asset demand. The subsequent price crash carries on its own momentum. We can treat demand as a latent variable to see these oscillations manifest in return data. Substituting lagged demand into the equation for returns, and iterating infinitely yields

$$r_t = - \sum_{T=1}^{\infty} \left(\frac{\alpha}{\gamma \sigma^2} \right)^T r_{t-T} + \frac{S_t \gamma \sigma^2}{\beta N} \quad (2.27)$$

as long as $\alpha/\gamma\sigma^2 < 1$, so that the contribution of demand fluctuations to returns converges to zero over time. This is an autoregressive model with infinite lags, where the coefficients decrease exponentially with lag size T . Without any knowledge of asset demand, the second term encapsulates an unobservable error term, which the model links to exogenous changes in the asset's supply. Eq. 2.27 demonstrates that an exogenous increase in returns at time t is followed by a smaller decrease in $t + 1$. This oscillation persists indefinitely, and would converge to zero rapidly if the social signal $\alpha/\gamma\sigma^2$ is sufficiently small.

2.B.4 Modeling predictions

Prediction 1: A mechanism for peer effects in asset demand *Given that asset demands by social investors are complementary, a marginal increase(decrease) in peer outlook about an asset will raise(lower) the future outlook of an investor about the asset.*

We dedicate Section 2.3 to investigating strategic complementarities among investors on WSB. Besides testing for the direct effect of peers on investor sentiment (Prediction 1), we also use the opportunity to test our assumption for mechanical extrapolation: A uniform, marginal increase(decrease) in an asset's returns will raise (lower) the future outlook of an investor about the asset. It will also indirectly increase(decrease) the outlook of an investor through increasing(decreasing) the outlook of peers.

Prediction 2: Return predictability *An increase in asset demand explained by past investor sentiment increases the asset's price.*

According to Eq. 2.26, there is a positive contemporaneous correlation between returns and WSB sentiments, but a negative correlation with lagged sentiments. However, one issue with the positive correlation is the challenge of identification. The data typically reflects an equilibrium outcome where sentiments and returns are positive, and vice versa. The main challenge is to identify exogenous variations in current sentiment relative to current returns. If current sentiments do not depend on future returns, we can estimate the impact of WSB sentiments on returns using sentiment scores predicted from preceding discussions. Stock-specific characteristics will also drive persistent heterogeneity in the expressed sentiments of WSB users. We tackle the question of whether social dynamics can account for some

return predictability in Section 2.4.

Prediction 3: Peer-driven bubble dynamics *Peer effects are a mechanism behind asset bubbles.*

We use parameter estimates from Section 2.3 to demonstrate how peer effects can amplify bubbles in the presence of a change to the fundamental value of an asset, and also drive bubble-like dynamics in the presence of a social shock. Section 2.4 demonstrates the presence of bubble-like dynamics in assets discussed on WSB.

Prediction 4: Granular social forces *Idiosyncratic demand shocks do not average out in the presence of heavy-tailed attention and impact asset prices.*

Eq. 2.20 predicts that, in the presence of heavy-tailed attention to certain investors, heterogeneity in investor sentiments will not average out, and will instead have an impact on asset prices. An emphasis on viral content compels WSB discussants to follow specific, popular strategies, which are predicted to have an outsized impact on asset price returns. We investigate the role of these granular idiosyncratic sentiments in Section 2.4.

2.B.5 Discussion on Assumptions 2.1, 2.2

Assumption 2.1 Assumption 2.1 is in line with the finding that our asset demand model produces strategic complementarities in investor asset demands in Eq. 2.18. Several extensions of the simple formulation are possible to account for greater complexities in social interactions. It can, for example, be extended to take into account *key players* (Zenou 2016) through changing the way that people weight the demand of others $\phi_{i,t-1}$ in the sum $1/N \sum_i \phi_{i,t-1}$ to $\sum_i s_i \phi_{i,t-1}$, where s_i captures the influence of player i and $\sum_i s_i = 1$.

In more complex settings, we can consider the unique complementarities between connected individuals as described in Zenou (2016), clusters of investors (Bouchaud & Potters 2003), or alternative information spreading / individual targeting models, which have received attention in the recent literature (Galeotti et al. 2020). The added complexity would affect aggregate demand through the expectations of other's demand ϕ_{t-1} .

Assumption 2.2 Mechanical extrapolation is our preferred way to introduce a relationship between prices and demand (Barberis et al. 2018). A model with mechanical extrapolation has several shortcomings, one of which is the inability to relate expectation updates to psychological underpinnings. However, our assumption is justified by our empirical work in Section 2.3.2 which demonstrates that individuals update their outlook based on recent asset returns.

2.B.6 Additional modelling frameworks: bubbles with peer effects

As noted in our introduction and in the excellent overview of Hirshleifer (2015), 'social investing' may play a role in a variety of price dynamics, beyond the mechanisms highlighted above. Our model considers a specific equilibrium scenario where information sharing by

peers and subsequent price dynamics are explained by strategic information complementarities. Quantifying peer dynamics is, however, equally relevant for other models.

Consider, for example, the model for bubbles with extrapolation by [Barberis et al. \(2018\)](#). We argue that in this scenario, sentiments expressed by peers may also play a role. [Barberis et al. \(2018\)](#) propose a model where ‘extrapolators’ determine their demand based on some weight, w_i , placed on the ‘fundamental trader’ valuation of an asset, while weight $(1 - w_i)$ is placed on an extrapolation component. Consistently with our prior assumptions, traders maximise a CARA utility function defined over their next period’s wealth. The extrapolation signal is determined by discounting past asset returns. We propose that this formulation can be modified to incorporate the effect of information shared by peers. We present the model extension below, and simulate price dynamics using our parameter estimates within the next section. The extension allows us to justify bubble dynamics in the absence of a change in fundamental news about the asset. In our setting, ‘animal spirits’ alone among investors can cause a run-up and subsequent drop in price. Additionally, we observe that information assimilation from peers creates a ‘longer memory’ among hype investors – asset bubbles take longer to materialise, and longer to dissipate.

Bubble dynamics In addition to the equilibrium setting, we demonstrate how peer effects are relevant in modelling bubble dynamics through incorporating them in an extension of [Barberis et al. \(2018\)](#). We highlight the relevant model details below, however, direct the reader to the original paper for the full model setup. In the original model, extrapolators determine their demand from a ‘fundamental signal’ with weight w_i , as well as an extrapolation signal with weight $(1 - w_i)$, and trade with fundamental traders in the market. The demand function for extrapolators with non-varying temporal weights is:

$$w_i \frac{F_t}{\gamma \sigma^2} + (1 - w_i) \frac{M_t}{\gamma \sigma^2}. \quad (2.28)$$

[Barberis et al. \(2018\)](#) define M_t (X_t in the original text) as:

$$\begin{aligned} M_t &= (1 - \theta_E) \sum_k^{t-1} \theta^{k-1} (P_{t-k} - P_{t-k-1}) + \theta^{t-1} X_1, \\ &= (1 - \theta_E)(P_{t-1} - P_{t-2}) + \theta_E M_{t-1}. \end{aligned}$$

where θ_E is the weight placed on recent versus older price changes, and is between zero and one.

We propose to modify extrapolator signal to incorporate a social component, M_t^s :

$$M_t^s = (1 - \theta_E) \phi_t + \theta_E M_{t-1}^s, \quad (2.29)$$

where ϕ_t is the average sentiment determined from past price returns (extrapolation), as

well as a past expressed sentiments (persistent demand driven by peer effects),

$$\phi_t = \beta^*(P_{t-1} - P_{t-2}) + \alpha^*\phi_{t-1}, \quad (2.30)$$

where β^* and α^* sum to one.

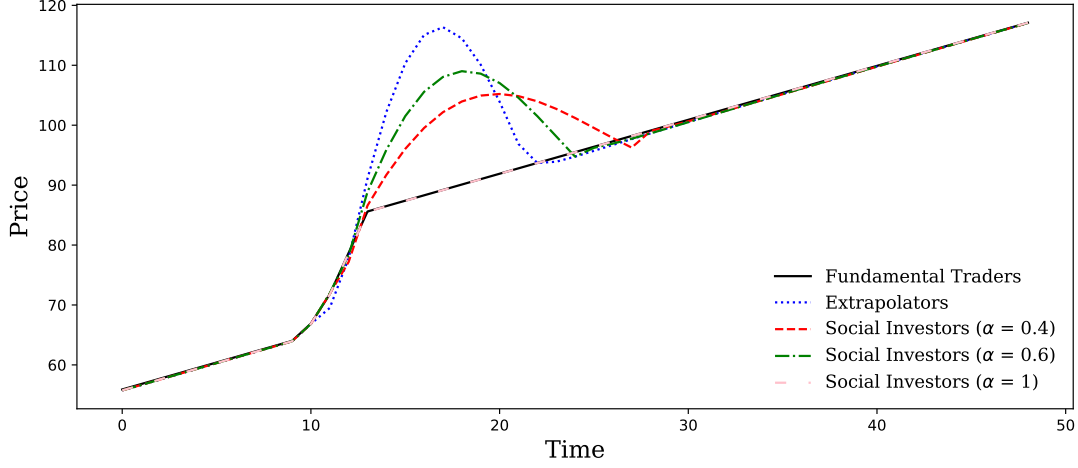


Figure 2.B.1: **Bubbles with Different Parameter Values for α^* , β^* :** We choose initial parameters similar to those from Figure (1) in [Barberis et al. \(2018\)](#): $w_i = 0.1$, $\sigma^2 = 3$, $\gamma = 0.1$, $\theta_E = 0.1$; extrapolators / social investors make up 70% of investors, the remainder are fundamental traders; the quantity of the asset is set to one.

We use our estimates for β and α to compare resulting bubble dynamics in the presence of peer effects to the original findings in [Barberis et al. \(2018\)](#) – our estimates demonstrate that social investors on WSB press a relative weight α^* of 0.6 on the sentiments of peers and a relative weight β^* of 0.4 on recent returns. Figure 2.B.1 demonstrates that the values for α^* , β^* control how long of a memory investors have. As α^* increases, we observe that the bubble takes a longer period of time to form and dissipate.

2.C Extra results for peer effects

2.C.1 Target independent variable

We build a discrete-choice empirical strategy to suit our model. Under the assumption that $u_{i,t}$ is drawn from a standard type-I Extreme Value Distribution, we model the log-odds of expressed investor sentiments ϕ_i by a standard multivariate logistic function,

$$\log\left(\frac{P(\phi_{i,t} = +1)}{P(\phi_{i,t} = 0)}\right) = g(b_{i,t}) + f(\bar{\phi}_{-i,(t-1,t)}) - \theta\sigma_{i,t}^2 + u_{i,t}^+, \quad (2.31)$$

$$\log\left(\frac{P(\phi_{i,t} = -1)}{P(\phi_{i,t} = 0)}\right) = -g(b_{i,t}) - f(\bar{\phi}_{-i,(t-1,t)}) - \theta\sigma_{i,t}^2 + u_{i,t}^-, \quad (2.32)$$

where t denotes time, and $(t-1, t)$ an interval preceding t . We aggregate bullish and bearish sentiments into one continuous variable, $\Phi_{i,t}$:

$$\Phi_{i,t} = \frac{1}{2} \log \left(\frac{P(\phi_{i,t} = +1)}{P(\phi_{i,t} = -1)} \right) = g(b_{i,t}) + f(\bar{\phi}_{-i,(t-1,t)}) + \frac{u_{i,t}^+ - u_{i,t}^-}{2}. \quad (2.33)$$

In the main body, the error term is expressed as $\epsilon_{i,t}$. Under the assumption that $u_{i,t}^+$ and $u_{i,t}^-$ are independent and identically distributed, $u_{i,t}^+ - u_{i,t}^-$ will follow a logistic distribution with finite variance.

2.C.2 First stage regression estimates

Tables 2.1 and 2.3 present our full regression estimates. Table 2.C.1 presents our First Stage estimates for our *Commenter Network* approach, which has multiple IVs.

Table 2.C.1: First Stage estimates for *Commenter Network* approach

	$\phi_{i,j,t-1}^{-1}$	$\phi_{i,j,t-1}^0$	$\phi_{i,j,t-1}^{+1}$	$\bar{\Phi}_{-i,j,t-1}$
<i>Dependent variable:</i>				
Sentiment of Peers	-0.30 (0.04) ***	0.12 (0.03) ***	0.25 (0.03) ***	0.14 (0.01) ***

Notes: The dependent variable is individual investor sentiment about an asset expressed in a single submission, scaled continuously between $(-\infty, \infty)$, modelled using IVs. We estimate it using the individual's previously expressed sentiment about the same asset ($\phi_{i,j,t-1}$) as a categorical variable, with the author not having posted previously ($\phi_{i,j,t-1}^{NA}$) as the baseline, as well as the average sentiment of posts that the author commented on previously ($\bar{\Phi}_{-i,j,t-1}$). We use the timing of IVs to control for common shocks, as discussed in the main text. Our regression has 24,013 observations and an F-statistic of 118.

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

2.C.3 Normalisation procedure and non-normalised coefficient estimates

In order to compare the relative impacts across variables, we perform mean / standard deviation normalisation on the non-categorical variables within our regression. We normalise market variables with respect to the log-returns of all assets discussed on WSB since the forum's creation in 2012. We normalise sentiment variables with respect to the observations within our regressions. The normalisation is performed in order to be able to compare the impact of peer effects and returns on sentiment formation.

Table 2.C.2 presents the non-normalised coefficient estimates for our second stage. We observe that the coefficient on returns and predicted peer sentiment are both higher, and the coefficient on returns is larger than that of predicted peer sentiment: these changes relate to the standard deviations of the two variables which are 0.037 for daily returns and 0.271 for predicted peer sentiment, *Frequent Posters*, and 0.168 for predicted peer sentiment, *Commenter Network*. The first phenomenon is explained by the fact that the standard deviation for both variables is less than one; the second is explained by the fact that daily returns have a substantially smaller standard deviation than that of peer sentiment.

Table 2.C.2: Peer influence in WSB sentiments

	Frequent Posters (1)	Network (2)
Second Stage – peer influence estimated using <i>predicted</i> average sentiment of peers (<i>non-normalised</i>)		
<i>Dependent Variable: Investor Sentiment ($\Phi_{i,j,t}$)</i>		
Average peer sentiment, $\hat{\Phi}_{-i,j,(t-1)}$ (<i>predicted</i>)	0.198 (0.053) ***	0.197 (0.083) **
$r_{j,t}$	0.993 (0.185) ***	0.921 (0.251) ***
Author & asset controls ($X_{i,j,t}$)	Yes	Yes

Notes: This table presents the non-normalised coefficient estimates for the Second Stage of our *Frequent Posters* and *Commenter Network* regressions.

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

2.C.4 Evidence of identification strategy – Frequent Posters

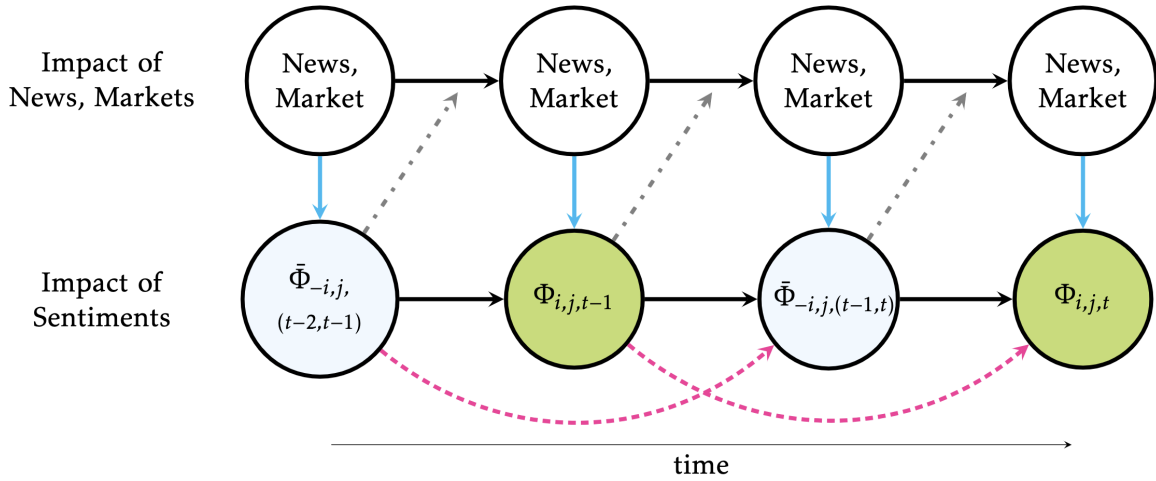
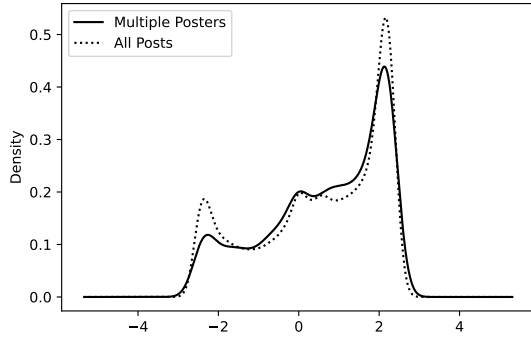


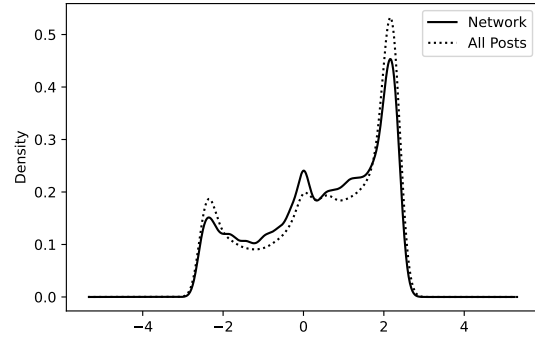
Figure 2.C.1: **Frequent Posters Directed Acyclic Graph (DAG)**: We trace the flow of information within our system. Arrows represent the impact that information from one source has on the next source. Light blue nodes $\bar{\Phi}_{-i,j,(t-2,t-1)}$ and $\bar{\Phi}_{-i,j,(t-1,t)}$ represent peer sentiment; green nodes $\Phi_{i,j,t-1}$ and $\Phi_{i,j,t}$ represent the sentiment of investor i - our target variable. Time t is expressed in event time. The magenta, dashed line represents the impact that historical sentiment expressed about an asset has on the author's own future opinion due to people being reticent to change their minds (Ross et al. 1975). In our first stage, we estimate $\bar{\Phi}_{-i,j,(t-1,t)}$ by $\bar{\Phi}_{-i,j,(t-2,t-1)}$, controlling for the market move at the time of the peer's initial post while estimating the coefficients. In this way, we are able to isolate the impact that peer sentiment $\bar{\Phi}_{-i,j,(t-1,t)}$ has on individual i at time t , $\Phi_{i,j,t}$.

Channels of Influence - Directed Acyclic Graph We present the channels of influence used within the *Frequent Posters* framework in Figure 2.C.1. We reason about our choice of IV through the Directed Acyclic Graph (DAG) shown in Figure 2.C.1. We consider that historic news and market moves are fully reflected in the news and market information available within the following timestep. Information shared by peers is also fully incorporated from one timestep to the next; however, dotted pink lines indicate the persistence of individual author sentiments (the persistence of individual formed impressions). Leverag-

ing the structure of our DAG, we estimate investor k 's sentiment (a peer of investor i) about asset j , $\Phi_{k,j,t}$, based on the sentiment they expressed previously, $\Phi_{k,j,t-1}$, and control for asset returns at the time of their original post,



(a) Frequent Posters Sentiment PDF



(b) Commenters Sentiment PDF

Figure 2.C.2: Density Plot of Sentiments Expressed on WSB: We present the density plot of the sentiments expressed by users on WSB who post multiple times, labelled as *Multiple Posters*, those who comment on others' posts, labelled as *Network*, and that of all submissions, labelled as *All Posts*.

Distribution of sentiments A potential concern with our approach is whether the sentiments expressed by individuals who post multiple times follow the same distribution as all submissions on the forum. Figure 2.C.2a presents the distribution of sentiments for the second or later post of an author about a ticker, which provides evidence that the sentiment distributions are similar to that of other posters on WSB. This supports the hypothesis that our analysis offers insight into how all individuals on WSB form opinions.

Random peers A second concern is whether we effectively control for unobserved ticker characteristics. We replace the composition of an author's peers with a random cohort of people who post on WSB about the same ticker. The random cohort is chosen as follows. We observe how many peers an individual author has. We then select a random sample of the same number of individuals, without replacement, who do not post between an author's two posts but post about the ticker at a different time before $(t - 1)$ for the *Frequent Posters* approach (if fewer individuals post before, we select all of those individuals). The results are presented in Tables 2.1, column (3). We observe that all the coefficients remain close to their original values, except for the peer effect, which becomes insignificant. This lends credibility to our peer identification strategy and shows that unobserved factors that influence within ticker variation are not confounding our estimates. This also supports our argument that we are not observing a longer time period of information absorption from peers – if this were the case, one would expect the random cohort of peers that post before time $(t - 1)$ to remain significant.

Endogenous choice to post multiple times One potential concern is that individuals who post multiple times are those that likely agree with their peers, whereas the sentiments of those that disagree will drop out of the forum and not create a second post. We, therefore, look at the probability of individuals creating a follow up post at various time horizons, depending on the distance between their own sentiments about an asset and that of peers in the interim period. Figure 2.C.3 shows the probability of a user creating a follow up post, plotted in different buckets depending on the magnitude of disagreement between them and their peers – we select several time horizons between four and seven days when an author is most likely to create a follow-up post. We observe that individuals are *not* more likely to create posts if they agree with peers – if this were the case, we would expect higher probabilities of creating a follow-up post in the $(-1, 0]$ and $(0, 1]$ buckets (where the sentiment difference between the individual and the average sentiment of peers is lower). We, therefore, observe that differences in opinion persist among individuals who post multiple times about the same asset on WSB.

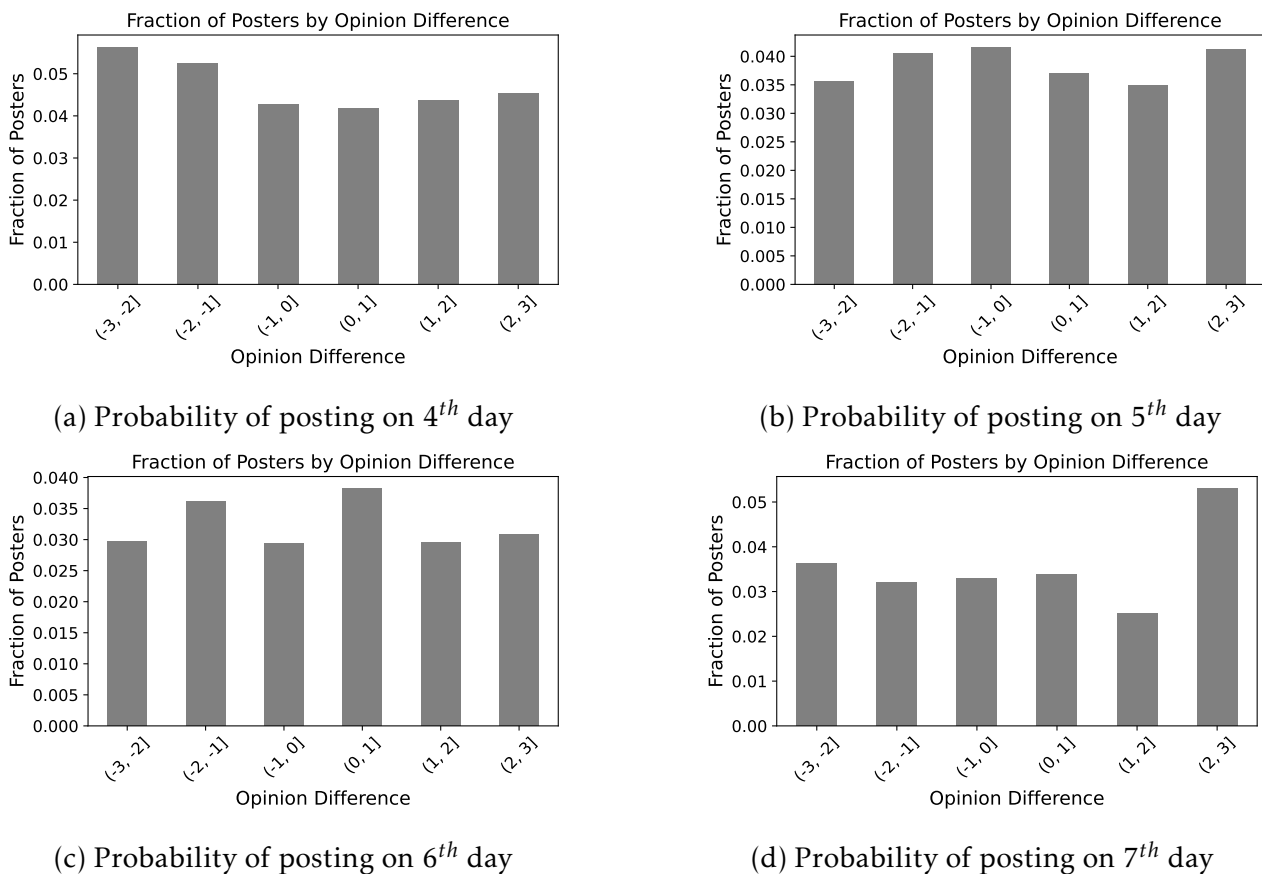


Figure 2.C.3: **Probability of posting:** We look at the probability of a user creating a subsequent post about an asset four, five, six and seven days after their original post in Figures 2.C.3a, 2.C.3b, 2.C.3c, 2.C.3d, respectively. The probability is shown in buckets versus the difference between the author's sentiment in the follow-up post and the average sentiment of peers expressed in the three, four, five, and six days prior to that post, respectively.

We also consider the possibility that there may be another important source of endogeneity driving the choice of certain authors to create a second post rather than others: the fact

that the market moves in line with the author's predictions. If this were the case, instead of peer effects, we may be observing author i 's reaction to general market developments, rather than true peer effects. To ensure that this is not the case, we include the average return in asset j between author i 's two posts ($t - 1, t$). We present the results in Table 2.C.3 – we observe that our original estimates remain consistent with this control. This implies that peers that post multiple times do not systematically reflect market moves, but indeed present a distinct perspective. The potential choice of post in-line with market moves does not confound our estimates.

Table 2.C.3: Robustness check including average inter-post returns

	Frequent Posters (1)
<i>Dependent Variable: Investor Sentiment ($\Phi_{i,j,t}$)</i>	
Average peer sentiment, $\hat{\Phi}_{-i,j,(t-1)}$ (<i>predicted</i>)	0.188 (0.050) ***
$r_{j,t}$	0.990 (0.170) ***
$r_{j,(t-1,t)}$	-0.032 (0.022)
Author & asset controls ($X_{i,j,t}$)	Yes

Notes: This table presents the non-normalised coefficient estimates for the Second Stage of our *Frequent Posters* regression when also including the average return in asset j in the time period between author i 's posts, ($t - 1, t$).
 *** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Temporal robustness check We run several additional robustness checks in order to ensure the validity of our identification strategy. One concern may arise from the fact that individuals may be slow to update their opinions, in relation to historical peer posts. In this way, individual i 's sentiment about asset j would be affected by peer sentiment before time ($t - 1$) at both times ($t - 1$) and t . In order to tackle this concern we run two different experiments and set a minimum time between ($t - 1$) and t to be at least: i) one day and ii) three days. An additional concern may be whether author i is actually exposed to the sentiments of the peers in question – if the time between ($t - 1$) and t is large, it may be more likely that the author is influenced by outside factors, rather than is constantly exposed to peers during such a prolonged time period. In order to tackle this we run an additional experiment capping the time between posts to be fourteen days. The results for these experiments are presented in Table 2.C.4.

Longer time horizon We use the dataset of [Semenova et al. \(2024\)](#) to verify whether our findings hold at a longer time horizon.²³ Specifically, we look at two time periods using the alternative data: i) data through February 1, 2021, and ii) data after February 1, 2021 (the dataset spans through June 27, 2021). GME reached its peak on January 28, 2021. The

²³It is worthwhile to note that our data through 2020 is not identical to theirs due to their selection of a different subset of tickers from AlphaVantage and due to different data processing methodologies.

Table 2.C.4: Controlling for the time between author i 's submissions for Frequent Posters

	Greater than One Day (1)	Greater than Three Days (2)	Less than Twelve Days (3)
<i>Dependent Variable: Investor Sentiment ($\Phi_{i,j,t}$)</i>			
Average peer sentiment, $\bar{\Phi}_{-i,j,(t-1,t)}$ (<i>predicted</i>)	0.241 (0.061) ***	0.195 (0.071) ***	0.223 (0.066) ***
$r_{j,t}$	1.040 (0.195) ***	1.069 (0.223) ***	0.930 (0.190) ***
Author & asset controls ($X_{i,j,t}$)	Yes	Yes	Yes
Number of obs.	9,712	8,406	5,545
F-statistic	61	53	20

Notes: This table presents the Second Stage OLS estimates for peer influence on WSB limiting the time between user i 's posts to be at least one day in Column (1), at least three days in Column (2), and at most fourteen days in Column (3). Variables are not normalised and therefore coefficients are comparable to those in Table 2.C.2. *** Significant at 1% level ** Significant at 5% level * Significant at 10% level

results are presented in Table 2.C.5. We note that for the time period ending shortly after the GME short-squeeze, presented in Column (1), the results are consistent with our previous findings. However, in the period following GameStop, the sentiments of peers and recent market moves appear to lose significance, however, users appear more stubborn in their views as visible from the coefficient on $\Phi_{i,j,t-1}$ in Column (2) – user i 's own historical sentiment about stock j . The findings underscore the fact that the GME frenzy led to a change in the quality of informational content shared on WSB, as also documented in Bradley et al. (2021). It also appears that users became more stubborn and unwilling to update their views - consistent with the later narrative about the forum as a social movement, rather than a place to discuss meaningful insights about assets.

2.C.5 Evidence of identification strategy – Commenter Network

Distribution of sentiments A potential concern with our approach is whether the sentiments expressed by individuals who are part of the commenters network follow the same distribution as all submissions on the forum. Figure 2.C.2b presents the distribution of sentiments for those who comment on other's posts, which provides evidence that the sentiment distributions are similar to that of other posters on WSB. This supports the hypothesis that our analysis offers insight into how all individuals on WSB form opinions.

Random network rewiring A concern is whether we effectively control for unobserved ticker characteristics. Similarly to Patacchini & Zenou (2016), we run 'placebo tests', where we replace the composition of an author's peers with a random cohort of people who post on WSB about the same ticker. The random cohort is chosen as follows. We observe how many peers an individual author has. We then select a random sample of the same number of individuals, without replacement, through a random network rewiring (we select posts randomly about the same ticker before the current post). The results are presented

Table 2.C.5: Peer influence: *Frequent Posters* – IV estimates pre- and post-GME

		Dependent Variable: $\Phi_{i,j,t}$	
		Pre-GME (1)	Post-GME (2)
Independent Variables	$\Phi_{i,j,t-1}$	0.228 (0.052) ***	0.297 (0.015) ***
	$\hat{\Phi}_{-i,j,(t-1,t)}$	0.024 (0.008) ***	0.008 (0.012)
	$r_{j,t}$	0.006 (0.003) **	0.002 (0.003)
	$\bar{r}_{j,t}$	0.001 (0.003)	0.005 (0.004)
	$\sigma_{j,t}^2$	0.001 (0.003)	0.009 (0.005) *
	Ticker Fixed Effects	Yes	Yes
No. Observations:		17,833	9,084
R^2 :		0.13	0.13
R^2_{adj} :		0.12	0.12

Notes: The dependent variable is individual investor sentiment about an asset, scaled continuously between $(-\infty, \infty)$, and is estimated by the individual's previously expressed sentiment about the same asset ($\Phi_{i,j,t-1}$) and a set of market control variables ($r_{j,t}, \bar{r}_{j,t}, \sigma_{j,t}^2$), using OLS. The sentiment of peers ($\hat{\Phi}_{-i,j,(t-1,t)}$) is estimated using data ending immediately after GME hit (spanning through January 31, 2021) in Column (1) and then using data after GME hit its peak in Column (2). The dataset of [Semenova et al. \(2024\)](#) is used for these experiments. Variables are normalised.

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

in Tables 2.3, column (3). We observe that all the coefficients remain close to their original values, except for the peer effect, which becomes insignificant. This lends credibility to our peer identification strategy and shows that unobserved factors that influence within ticker variation are not confounding our estimates.

Overidentifying restrictions We cannot directly calculate the J-statistic for our *Commenter Network* approach, since we estimate our IV using observations on several neighbours. We, therefore, take an average of the neighbours' past sentiments (transforming the categorical variable into a continuous one) and the average across their neighbour's neighbours' sentiments. We use this to compute a J-Statistic with two degrees of freedom.

2.D Additional results for market impact

2.D.1 GIV

Extracting idiosyncratic social shocks In Table 2.D.1, we present the coefficients from Eq. 2.8. We observe that, consistently with our previous findings and the proposed structure in Figure 2.C.1, returns and past user sentiments all impact the expressed sentiment of user i about asset j at time t .

GIV - First Stage Table 2.D.2 presents our First Stage regression, where we regress the popularity-weighted idiosyncratic sentiment, as our dependent variable, on the difference

Table 2.D.1: Idiosyncratic social shocks estimation

<i>Dependent variable: $\Phi_{i,j,t}$</i>	
$r_{j,t}$	0.324 (0.086) ***
r_{j,t_w}	0.370 (0.090) ***
Φ_{j,t_w}	0.212 (0.043) ***
Ticker FE	Yes
Number of obs.	46,694
R ²	0.077
F-statistic	41.67

Notes: The dependent variable is the log-odds of a given submission by author i at time t on stock j to express bullish over bearish sentiment. Explanatory variables include: the log return on day t , $r_{j,t}$, the log-return in week including day t denoted as t_w , r_{j,t_w} , and the average past sentiment of peers, Φ_{i,j,t_w-1} . Accompanying standard errors, displayed in brackets, are clustered at the stock level.

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Table 2.D.2: GIV - First Stage

<i>Dependent variable: $\bar{e}_{j,t_w}$</i>	
GIV	0.932 (0.019) ***
Number of obs.	2,441
R ²	0.510
F-statistic	2,292

Notes: The dependent variable is the popularity-weighted average idiosyncratic sentiment expressed about asset j in week t_w , $\bar{e}_{j,t_w}$. The explanatory variable is the difference between the popularity-weighted and raw average of idiosyncratic sentiments expressed about asset j in week t_w .

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

between the popularity-weighted and regular average idiosyncratic sentiment. We observe that our GIV is highly predictive of the popularity-weighted sentiment, $\bar{e}_{j,t_w}$, however, explains only part of the variation. In this way, we are able to extract the element of the popularity-weighted sentiment measure driven by social preferences, versus those driven by other factors.

We illustrate the effect of the GIV through the following two scenarios. In *Scenario One*, a very popular and an unpopular post both express a positive, idiosyncratic sentiment. In *Scenario Two*, a very popular expresses the positive sentiment, while the unpopular posts expresses no idiosyncratic sentiment. Without employing the GIV, our original popularity-weighted average idiosyncratic sentiment measure $\bar{e}_{j,t_w}$ would be very similar in both scenarios. However, by using the GIV, we would *predict* zero idiosyncratic sentiment in *Scenario One*, but a large positive idiosyncratic sentiment in *Scenario Two*. In this way, our GIV allows us to distinguish the signal coming from popularity.

Influential posts – examples We present some examples of popular content during weeks with a with a large $\hat{u}_{j,t_w}$. Figure 2.D.1 presents the most popular posts about AMZN in the week of June 24-26, 2020 – the posts were identified as having large, positive, idiosyncratic

sentiment within the week. AMZN exhibited a return of 7.8% the following week. Figure 2.D.3 presents the most popular posts about WYNN in the week of March 16-20, 2020 – the post had a positive, idiosyncratic sentiment within the week. WYNN exhibited a return of 21.8% the following week. Figure 2.D.2 presents the most popular post about GM in the week of March 9-13, 2020 – the post was identified as having large, negative, idiosyncratic sentiment within the week. The following week, GM exhibited a 27% decline.

We observe that the content has very different style. Some discussions, such as the GM post in Figure 2.D.2, present a thorough, substantiated discussion backing the author's claim. Other popular content, such as the WYNN post in Figure 2.D.3, attracts attention through the large risk that the author has taken, potentially inspiring others to follow in their footsteps.

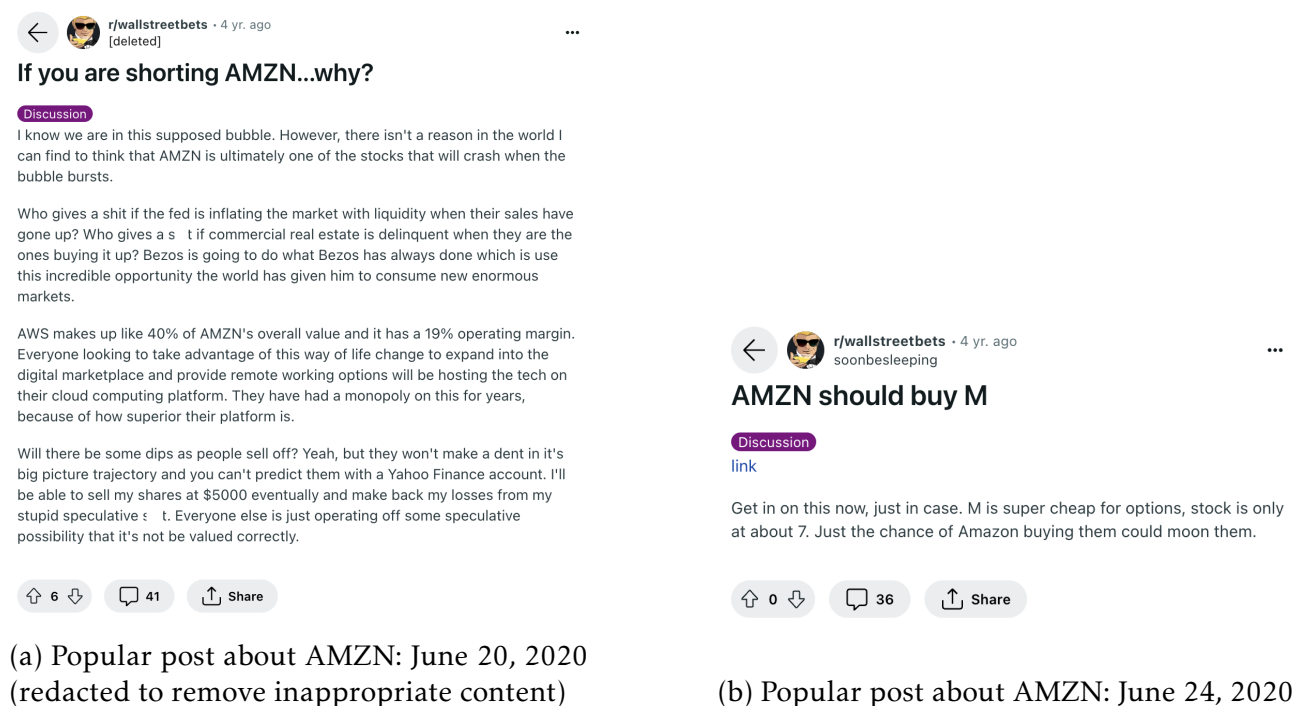


Figure 2.D.1: Popular post about AMZN, June, 2020.

2.D.2 Additional evidence of price impact

We are interested in finding variation in current sentiments which are exogenous with respect to current returns. The goal is to detect a positive effect carried by retail investor sentiment, proxied by WSB activity. First, we formulate the linear relationship between returns and current sentiments. We then propose a 2SLS estimation strategy to quantify the impact of social contagion on stock market variables.

Independent variable Our independent variable estimated from WSB sentiment data measures sentiment change $\Delta\Phi_{j,t}$, the first-difference of stock j 's mean daily sentiment between calendar weeks t and $t - 1$. The purpose for this variable is to gauge the stock-specific response to a change in WSB's associated attention and sentiments on a week-by-week basis.

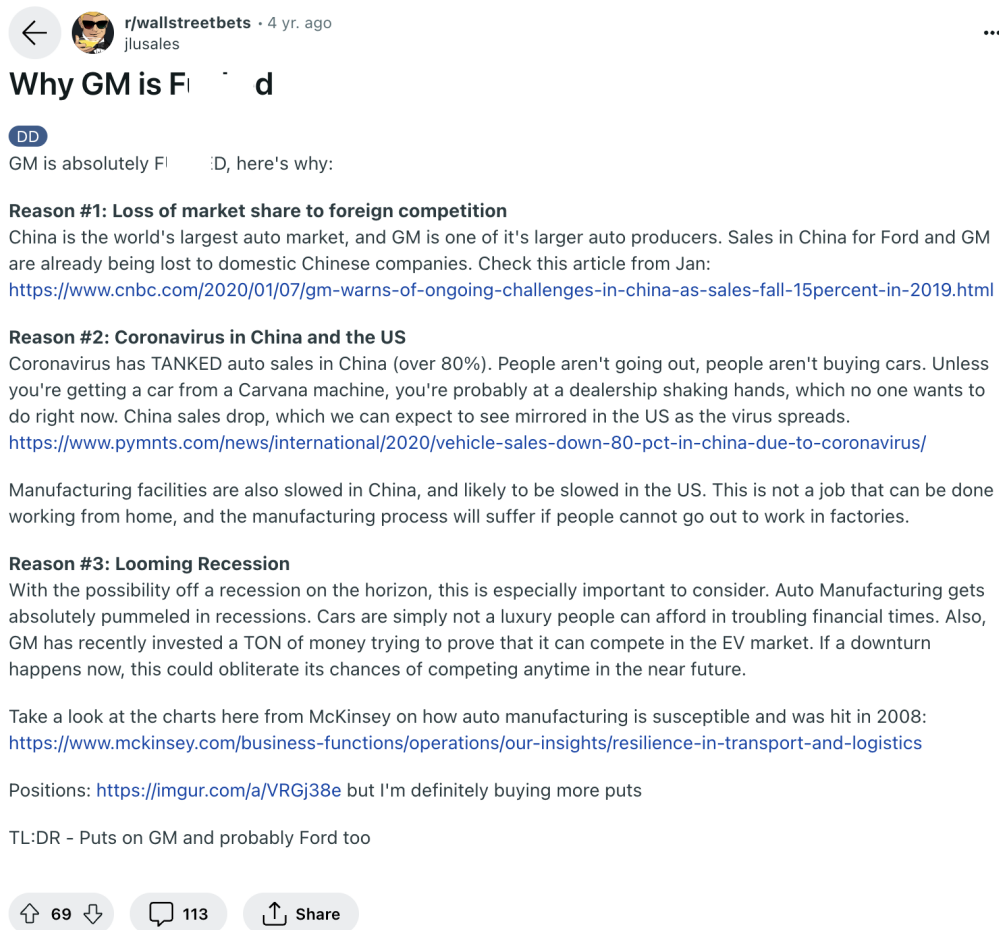


Figure 2.D.2: Popular post GM: March 10, 2020 (redacted).

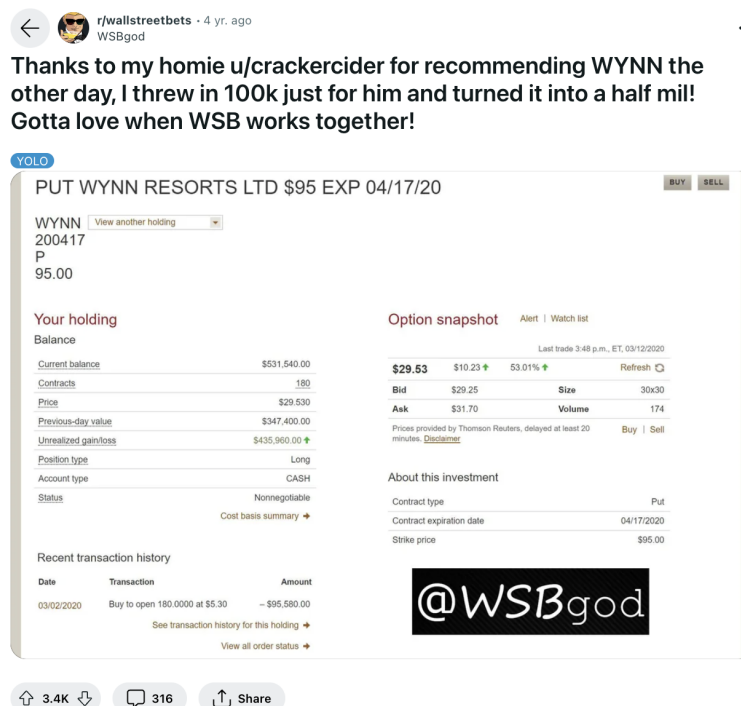


Figure 2.D.3: Popular post WYNN: March 13, 2020.

Measuring the difference in average sentiments between two periods proxies for the change in asset demand due to changes in the intensity of corresponding sentiments.

Reduced Form We regress changes in weekly log-returns on changes in weekly sentiments:

$$\Delta \bar{r}_{j,t} = \omega \Delta \Phi_{j,t} + \eta_t + \varepsilon_{j,t}, \quad (2.34)$$

where ω is the coefficient of interest, η_t denotes week fixed effects, and $\varepsilon_{j,t}$ an idiosyncratic error. We specifically use the first difference in weekly returns to account for stock-specific heterogeneity. The Reduced Form approach does not provide conclusive evidence of a causal relationship between social investor activity and stock market activity. The narratives and sentiments expressed on social media platforms are influenced by real-time news, events, and stock market fluctuations, which can result in reverse causality. In other words, positive sentiments may be expressed during weeks of high returns, regardless of previous sentiment. This makes it challenging to establish a clear link between social investor activity and stock market activity.

First Stage We use variation in sentiments that can be explained by past activity on WSB and past stock performance to identify our parameter of interest. We predict sentiment $\Phi_{j,t}$ using past stock price behaviour, as well as past sentiments:

$$\Phi_{j,t}^+ = \log \left(\frac{P(\phi_{j,t} = +1)}{P(\phi_{j,t} = 0)} \right) = \lambda_r^+ \bar{r}_{j,t-1} + \lambda_\sigma^+ \sigma_{j,t-1}^2 + \lambda_1^+ \Phi_{j,t-1}^+ + \lambda_2^+ \Phi_{j,t-1}^- + \eta_t^+ + \varepsilon_{j,t}^+, \quad (2.35)$$

$$\Phi_{j,t}^- = \log \left(\frac{P(\phi_{j,t} = -1)}{P(\phi_{j,t} = 0)} \right) = \lambda_r^- \bar{r}_{j,t-1} + \lambda_\sigma^- \sigma_{j,t-1}^2 + \lambda_1^- \Phi_{j,t-1}^+ + \lambda_2^- \Phi_{j,t-1}^- + \eta_t^- + \varepsilon_{j,t}^-, \quad (2.36)$$

where superscripts differentiate between the average log-odds of a submission in week t expressing bullish (+) versus negative (−) sentiments, over neutral sentiments. Week fixed effects remain in the sentiment models, so that the full estimation strategy rests on within-week variation in all explaining, as well as explained, variables.

The approach outlined above relies on coarse aggregates for sentiments: the probabilities here are not estimated on data for individual submission sentiments, as is the case in Section 2.3. Rather, the probabilities are calculated by averaging the probabilities for *all* submissions in week t , discussing ticker j , to be bullish ($P(\phi_{j,t} = +1)$), bearish ($P(\phi_{j,t} = -1)$), or neutral ($P(\phi_{j,t} = 0)$). Predicted values for our sentiment measure follow from the fitted sentiment model:

$$\Delta \widehat{\Phi}_{j,t} = \frac{1}{2} (\widehat{\Phi}_{j,t}^+ - \widehat{\Phi}_{j,t}^-) - \Phi_{j,t-1}, \quad (2.37)$$

where a hat denotes the values fitted from the first stage regressions.

Results In all our estimates, we restrict ourselves to a sub-sample spanning January 2016 to July 2020. This choice serves to limit the amount of missing data in times when activity

on WSB was relatively sparse.

Table 2.D.3: First Stage estimates for consensus and contagion in WSB

	<i>Dependent variable:</i>	
	$\Phi_{j,t}^+$	$\Phi_{j,t}^-$
$\bar{r}_{j,t-1}$	-0.0002 (0.52)	-1.53 (0.65) **
$\sigma_{j,t-1}^2$	-3.78 (0.98) ***	-3.75 (0.68) ***
$\Phi_{j,t-1}^+$	0.09 (0.02) ***	-0.06 (0.01) ***
$\Phi_{j,t-1}^-$	-0.03 (0.01) ***	0.16 (0.01) ***
Week FE	Yes	Yes
Number of obs.	6,711	6,711
F-statistic	17.63	49.53

Notes: The dependent variable in Column (2) is the average log-odds of a given submission in week t on stock j to express bullish over neutral sentiment, and in Column (3) – bearish over neutral sentiments. Explanatory variables include: the average log-return $\bar{r}_{j,t-1}$, and the variance in log-returns $\sigma_{j,t-1}^2$. The logit-transformed sentiments are regressed on the lag of the weekly mean and variance of log-returns, as well as the lag in logit-transformed sentiments. Each specification includes week-specific fixed effects. Accompanying standard errors, displayed in brackets, are clustered at the stock level, and calculated in the manner of [MacKinnon & White \(1985\)](#).

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Table 2.D.3 helps assess the instruments' strength in predicting sentiments on WSB. The high F-statistics justify that the explanatory variables are not weak instruments. In both columns, we find that lagged weekly mean and variance in returns, combined with lagged sentiments, are significant predictors for the current log-odds in average weekly submissions expressing bullish and bearish sentiments. This is in line with our findings in Section 2.3.

Table 2.D.4 presents our main results. Panel A regresses changes in average returns against *observed* measures for sentiment changes $\Delta\Phi_{j,t}$. Panel B in Table 2.D.4 presents causal evidence for the impact of sentiments among WSB users on stock market variables, using *predicted* sentiments from the model presented in Table 2.D.3. The effect in Panel B from our Reduced Form appears larger than our First Stage - we explain this through the fact that the significance of the first stage is decreased due to our weekly fixed effects, which impact returns and sentiments simultaneously. The estimated average effect is small, which is not surprising given that many of the stocks discussed on WSB have large market capitalisation.

We do not argue that WSB alone affects the markets, but rather that WSB data offers a rich sample of retail investor behaviour. Variation in sentiments explained by the past offers a measure for the intensity by which retail investor asset demand propagates from one week to the next. Even though sentiments reflect current returns, prior beliefs are expected to change prices, thus returns, beyond the market average accounted for by time fixed effects.

Table 2.D.4: Market impact of WSB discourse

Panel A: Reduced Form relationship between WSB and market activity	
<i>Dependent variable:</i>	
	$\Delta \bar{r}_{j,t}$
$\Delta \Phi_{j,t}$	0.002 (0.0003) ***
Week FE	Yes
Number of obs.	6,671
F-statistic	24.32
Panel B: structural relationship between WSB and market activity	
$\Delta \widehat{\Phi}_{j,t}$	0.004 (0.001) ***
Week FE	Yes
Number of obs.	6,671
F-statistic	12.63
J-statistic	8.108

Notes: This table presents OLS estimates for stock j 's change in average log-return, $\Delta \bar{r}_{j,t}$, in week t . We filter the sample to stocks mentioned in at least 31 distinct submissions on WSB, and exclude any ETFs. Explanatory variables include a measure for sentiment change, $\Delta \Phi_{j,t}$, which tracks the change in average sentiments on WSB. Each specification includes week-specific fixed effects. Accompanying standard errors, displayed in brackets, are clustered at the stock level, and calculated in the manner of [MacKinnon & White \(1985\)](#). Panel A computes the coefficients using values directly from WSB data, whereas Panel B employs sentiments and stock discussion predicted by past sentiments, stock discussions, as well as returns and return volatility, for which results are in Table 2.D.3. The associated J-statistics are recorded at the bottom of Panel B, which are computed by regressing the residuals from the Second Stage on all variables used for predicted $\Delta \widehat{\Phi}_{j,t}$.

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Chapter 3

Is social media a game-changer for financial stability?

(With Valentina Semenova)

Abstract

A growing audience turns to social media for promising investment ideas. This paper presents and calibrates a dynamic system to study the qualitative behaviour of asset prices when investors are susceptible to social contagion in their investment decisions. These ‘hype’ investors invest more(less) when the asset performs well(poorly), and when their peers invest more(less). Different parameters produce eight distinct regimes for the system’s behaviour, with varying degrees of stability, cyclicity and multiple steady states for expressed sentiments and returns. The paper subsequently leverages text data from the WallStreetBets forum, using natural language processing tools, to estimate these parameters. We find that social contagion and trend following are sufficiently small that asset prices remain stable on average. However, the magnitude of those parameters can vary enough to produce potential bifurcations in stock price returns and sentiments, as was the case for a few stocks. Finally, we reconcile our predictions for system stability with observed returns and sentiments during the GameStop short squeeze.

3.1 Introduction

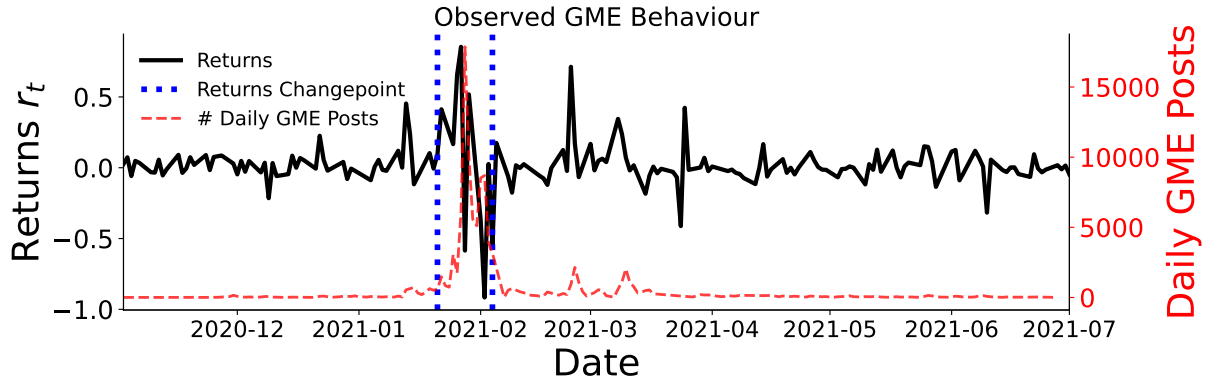


Figure 3.1.1: **GameStop Asset Returns and WallStreetBets Interest.**

GameStop (GME) had been popularised as a short-squeeze target for months before January 2021 – so why did a sudden influx of interested retail investors change the stock’s dynamics? The efficient markets hypothesis might suggest that the stock should have reflected all available information, and a movement on social media gaining momentum should have had little impact. However, the stock’s returns exhibited a detectable changepoint in late January 2021 (Figure 3.1.1) and subsequently become stable (marked by a second changepoint) as hype investor interest dissipated.

As social media gains a more prominent role on the virtual trading floor, we argue that understanding sudden changes in the asset dynamics requires adopting new, complex systems methods into the established financial toolkit. An important characteristic of complex systems is the potential for sudden changes and tipping points when the system moves from one dominant dynamic pattern to another. Despite there being a long-standing history of studying social movements with respect to their tipping points ([Lamberson et al. 2012](#), [Moser & Dilling 2007](#)), the connection to financial markets is under-explored. This paper’s contribution rests on bringing relevant complex systems techniques to data from social media and the financial markets, as well as using modern natural language processing (NLP) methods and network techniques to empirically estimate key parameters.

We model a market with ‘hype’ investors. Their decision to buy or sell an asset is governed, in part, by trend-following as well as social contagion. Prices adjust to satisfy their demand for the asset. In this model, parameters that govern the degree of trend-following and social contagion, as well as the financial ‘capacity’ of the hype investor population, can produce eight different regimes for the joint behaviour of sentiments and returns. These range from convergence to a steady state at zero, to periodic and quasi-periodic dynamics (bordering on chaos) with multiple steady states. Even in cases where dynamics stabilise in the long-term, assets can experience short-term, unexpected bursts of instability driven by a change in parameters. Such a change can be as simple as the hype investor population increasing. The complex systems approach, therefore, complements market efficiency to better explain short-term price fluctuations.

The paper uses text and network data from Reddit’s WallStreetBets (WSB) forum to categorise different assets by the type of regime they fall under within our model. We extract sentiments from the WSB subreddit to quantify the degree of trend-following and social contagion among users for the forum, both as a whole and for specific assets. We demonstrate that for average parameter values, WSB users would need to obtain a very large amount of capital to push a regime towards instability. However, these parameters can vary substantially between different assets, such that hype investors are capable of destabilising prices even with a fraction of the asset’s market capitalisation in available funds.

In general, the main manner in which asset prices experience a phase transition is when the number of hype investors increases. In this case, returns can quickly move from their stable steady state to exhibiting bifurcations, then reverting back to stability once hype investors lose interest. This finding is of particular relevance, since we find that asset interest in the hype investor population is well-modelled through classic epidemic contagion models. These imply that interest peaks, and subsequently declines over time. Finally, we reconcile our model with empirical parameter estimates in the context of the GameStop short squeeze.

Models in behavioural finance and narrative economics produce important insights for how psychological factors can impact asset prices (Shiller 1984, Black 1986, Hirshleifer 2001, Kahneman & Tversky 1972, Bursztyn et al. 2014, Lahno & Serra-Garcia 2015, Banerjee et al. 2013), with several recent papers focusing on traders exchanging information on social media (Heimer 2016, Cookson et al. 2022, Chen & Hwang 2022, Kuchler & Stroebel 2021, Hu et al. 2021, Bradley et al. 2021) as well as the predictive power of online forums for market moves (Antweiler & Frank 2004, Sabherwal et al. 2011, Chen et al. 2014, Azar & Lo 2016, Agrawal et al. 2018). A different strand of literature in the computational social sciences aims to model online collective dynamics (Bacaksizlar Turbic & Galesic 2023, Jusup et al. 2022, Nguyen et al. 2012, Becker et al. 2017, Kubin & Von Sikorski 2021), with several works focusing on investor forums specifically (Mancini et al. 2022, Lucchini et al. 2022). Researchers have also scrutinised the interactions of heterogeneous agents in stock markets, following many, sometimes adaptable, rules (Lux 1998, Hommes 2021), with some papers focusing on social dynamics specifically (Kirman 1993, Avery & Zemsky 1998, Cont & Bouchaud 2000). A highly relevant branch of literature considers market evolution and market ecology (Scholl et al. 2021, Farmer 2002). However, no studies have attempted to calibrate a complex systems model to understand the emergent financial dynamics driven by online investor discussions in real time. The current paper presents an analytic approach and empirical methodology to help address this gap.

3.2 Model setup

We build a dynamic model for the relationship between asset price returns and sentiments expressed by investors on social media. The main goal is to study the importance of social contagion, or peer effects, in sentiment for returns. The key challenge is that we do not observe actual trading decisions for investors. We therefore resort to a few assumptions to

pin down a functional form between sentiments observed on social media and price returns.

Consider a market for one asset with two types of participants: ‘hype’ investors, who buy Y shares, and ‘non-hype’ investors, who buy quantity S . Their total demand is equal to the number of shares outstanding,

$$Q = Y + S.$$

We propose a model in discrete time t where hype investors observe the behaviour of others, as well as the asset’s returns, before updating their demand for the asset to maximise their expected utility. The goal is to study the stability of a market with social dynamics, given that hype investors are able to move price. The key endogenous variable is asset demand, which governs the propensity for hype investors to buy or sell the asset, thus shifting demand Y_t .

Total hype investor demand Y_t is composed of i) the average demand across all active hype investors, $d_t \in [-1, 1]$, ii) the average purchasing power of an individual hype investor M/p_t , where M are funds available to an average hype investor, and iii) the total number of hype investors N :

$$Y_t = \frac{M}{p_t} N d_t. \quad (3.1)$$

We largely treat total funds available to hype investors MN as exogenous, and focus on their decision to buy or sell the asset, reflected by their average demand d_t . We endogenise demand d_t to change with the asset’s price, as well as the degree of coordination among hype investors.

Individual investor demand We formalise a functional form for a given hype investor’s demand $d_{i,t+1}$, indexed by i . The utility an investor derives from buying, or selling, the asset is a function of other hype investors’ average expressed *sentiment* ϕ_t , asset return $r_t \equiv p_t/p_{t-1} - 1$, as well as an error ϵ_i :

$$U_i(d_{i,t+1} = +1) = \epsilon_i^+ + \alpha\phi_t + \beta r_t - \gamma r_t^2, \quad U_i(d_{i,t+1} = -1) = \epsilon_i^- - \alpha\phi_t - \beta r_t - \gamma r_t^2. \quad (3.2)$$

where β is a trend following parameter, α is the network spillover effect from peers, and γ captures aversion to large swings in price. In this manner, the utility of buying(selling) the asset increases(decreases) with peer *sentiment*, as the investor seeks to align their position with that of other hype investors. They must rely on sentiment, because we assume that hype investors are unable to communicate the size of their position, and instead only share their decision to buy or sell the asset. Utility also increases(decreases) with recent price returns, and decreases(decreases) with return volatility. This imparts a trend-following tendency for hype investor’s decisions, but also an aversion to large swings in price. Residuals ϵ_i^+ and ϵ_i^- capture unobserved variation in utility among various hype investors when buying or selling the asset.

While we do not observe demands on WSB, we observe sentiments. We interpret sentiments as a binary random variable with which a user expresses their preference to buy

($\phi_{i,t+1} = 1$) or sell ($\phi_{i,t+1} = -1$) the asset. This allows us to determine the probability for a hype investor to express a given sentiment, based on observed peer sentiment and asset returns:

$$\phi_{i,t+1} = \begin{cases} +1, & \text{with probability } P(\phi_{i,t+1} = +1) = \frac{\exp[(\alpha\phi_t + \beta r_t)/\lambda]}{\exp[(\alpha\phi_t + \beta r_t)/\lambda] + 1}, \\ -1, & \text{with probability } P(\phi_{i,t+1} = -1) = \frac{1}{\exp[(\alpha\phi_t + \beta r_t)/\lambda] + 1}. \end{cases} \quad (3.3)$$

Investor i 's sentiment $\phi_{i,t+1}$ is what we observe on social media, and is assumed to be an honest evaluation of their investment decision $d_{i,t+1}$. Parameter λ is a scale parameter for the breadth of the error distributions, assumed to follow a Gumbel distribution, and specifies the amount heterogeneity in individual investment preferences [Bouchaud \(2013\)](#). Appendix [3.A.1](#) offers further details.

Aggregate hype investor buying intensity We aggregate sentiments into an average measure ϕ_{t+1} using Eq. [3.3](#):

$$\begin{aligned} \phi_{t+1} &= \frac{1}{N} \sum_{i=1}^N P(\phi_{i,t+1} = +1) - P(\phi_{i,t+1} = -1) \\ &= \frac{\exp[(\alpha\phi_t + \beta r_t)/\lambda] - 1}{\exp[(\alpha\phi_t + \beta r_t)/\lambda] + 1} \\ &= \tanh\left[\frac{\beta r_t + \alpha\phi_t}{\lambda}\right], \end{aligned} \quad (3.4)$$

with a detailed derivation presented in Appendix [3.A.1](#). Thus, Eq. [3.3](#) determines the evolution of average sentiments as a function of past sentiments, in addition to past returns. We describe the properties of this function in the **Results** section, as well as Appendix [3.A.1](#).

Hype investor demand and market moves We complete our stylised model by linking hype investor sentiments to price returns r_{t+1} . Recall that, under market clearing, the shares held by hype and non-hype investors must equal outstanding shares Q . For our purposes, we make a simplifying assumption that non-hype investors keep the nominal value invested in the asset fixed, so that

$$\Delta p_t S_t = 0.$$

This enables us to study the emergent price dynamics due to social dynamics in isolation (even though our findings hold when this assumption is relaxed). Under this assumption, the dynamics for price returns are:

$$r_{t+1} = \underbrace{\frac{MN}{p_t Q}}_{(1) \text{ Capacity}} \times \underbrace{\Delta\phi_{t+1}}_{(2) \text{ Sentiment}}. \quad (3.5)$$

Component (1) in Eq. [3.5](#), labelled ‘Capacity’, is a market depth term, measuring how much market power hype investors hold, as a share of the total market capitalisation of the

asset. Term (2) captures the change in their aggregate sentiment, which reflects underlying trades of hype investors as they respond to lagged returns and sentiments. Appendix 3.A.1 details the derivations of each of the above expressions and discusses properties of the hyperbolic tangent function.

Linking asset returns to sentiments is the main feature of the model, for which we make an assumption for the process that generates sentiments as a function of individual hype investor asset demands. This is relevant to our use of social media data, because demands are unobserved.

3.3 Results

Market stability in the presence of social dynamics

How will returns and sentiments evolve in the presence of hype investors? In the absence of any fundamental news, our model with hype investors generates the following dynamic system:

$$\phi_{t+1} = \tanh \left[\frac{\beta r_t + \alpha \phi_t}{\lambda} \right], \quad (3.6)$$

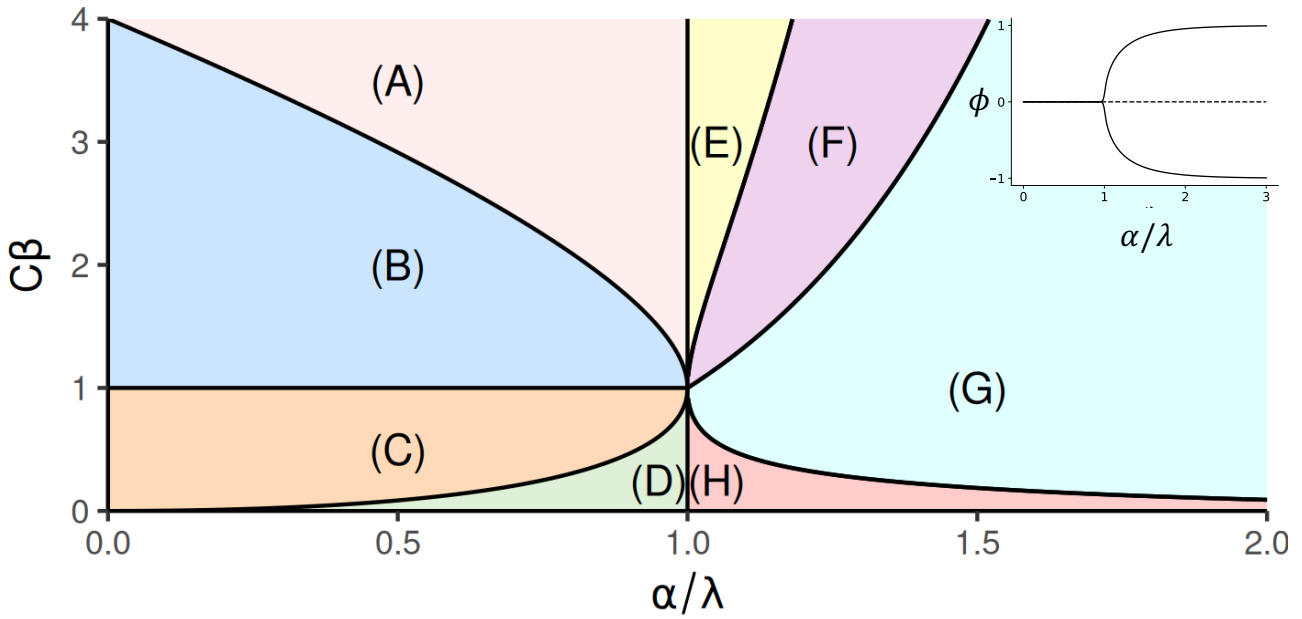
$$r_{t+1} = C(\phi_{t+1} - \phi_t), \quad (3.7)$$

where we simplify our initial expression for returns by defining the relative capacity of hype investors $C = MN/p_t Q$. The key dynamic of interest is the strategic coordination observed among hype investors at a given point in time, ϕ_t , and its impact on asset returns. Therefore, we focus on parameters α , β , and C .

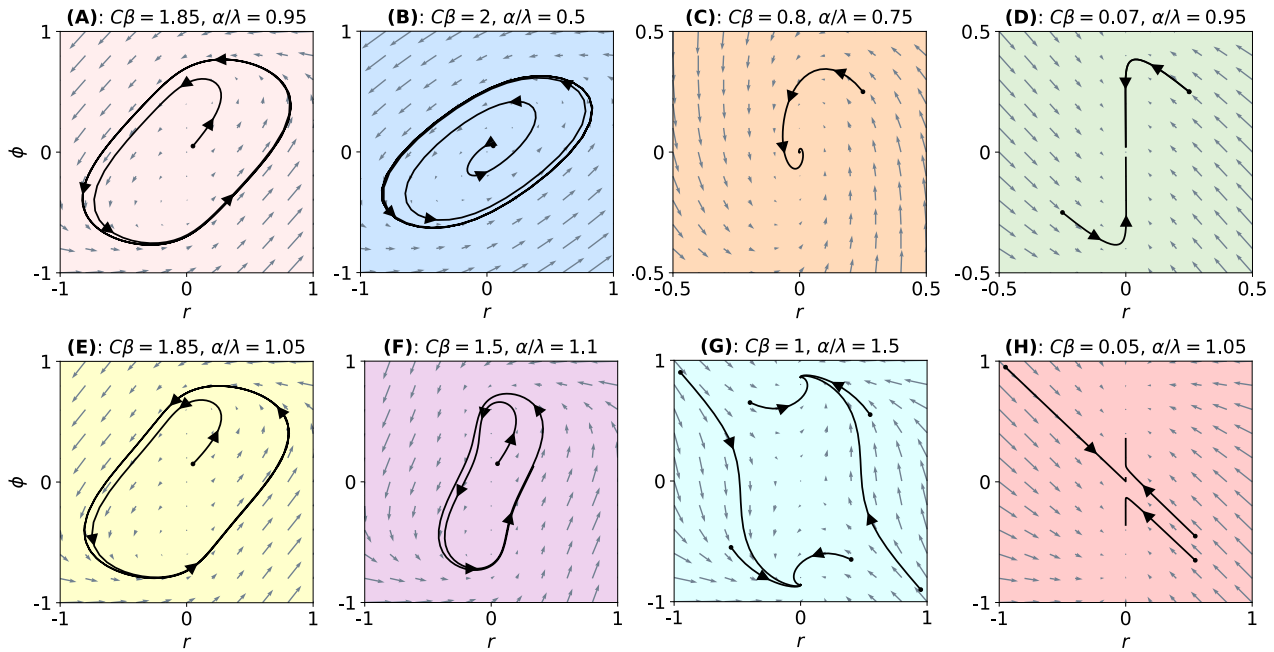
Eqs. 3.6 and 3.7 are two simultaneous difference equations, with sentiment ϕ_t and returns r_t as state variables. Importantly, the evolution of aggregate sentiment ϕ_{t+1} is non-linear. The system produces either one or three steady states, depending on parameter values. In any steady state, returns $r_{t+1} = r_t$ must be zero. The steady state values for sentiments ϕ are governed by α/λ , illustrated by the inset in Figure 3.3.1a: one at zero, then two more at $\{\phi^+, \phi^-\}$ when $\alpha/\lambda > 1$, which do not have a closed form solution. This yields three steady states for the system as whole when $\alpha/\lambda > 1$, which we denote as (sentiment, return) tuples: i) $(0, 0)$, ii) $(\phi^+, 0)$, and iii) $(\phi^-, 0)$.

Determining the steady states in a dynamic system is critical for understanding how the system evolves over time. As the name implies, when a dynamic system is in its steady state, it will remain there: in our case, sentiments and asset returns will not change. However, it is not obvious how the system behaves outside its steady state(s). For example, if the asset's return is non-zero due to extenuating circumstances, or if a rumour rips through the hype investor population and exogenously changes sentiments. To tackle this, we examine the stability of each steady state individually.

We show that the key parameters governing the system are i) the consensus parameter α , the degree to which investors' sentiments mimic those of their peers, and ii) the capacity parameter C , iii) the degree of trend following β and iv) the noise parameter λ . We examine



(a) Stability regions for returns and sentiments: in the region where $\alpha/\lambda < 1$, we observe one steady state where $r_{ss} = 0$ and $\phi_{ss} = 0$; in the region where $\alpha/\lambda > 1$, we observe three steady state where $r_{ss} = 0$ and ϕ_{ss} can take on values of ϕ^+ , 0 , ϕ^- . As noted in the inset bifurcation diagram (upper right) the ϕ^+ and ϕ^- are stable steady states and are indicated with a solid line, whereas $\phi_{ss} = 0$ is unstable and is indicated by a dotted line.



(b) Phase diagrams with different parameters: this plot shows the different trajectories that sentiments and returns can take when the system begins outside of a steady state.

Figure 3.3.1: Stability of Returns and Prices: The behaviour of price returns changes qualitatively with the intensity of social contagion (α/λ) and the financial capacity of hype investors ($C\beta$). Figure 3.3.1a distinguishes regions in the parameter space in which the behaviour of the system changes qualitatively. Three particular features are i) convergence to a fixed point, ii) the number of equilibria and iii) the presence of cycles. Figure 3.3.1b presents phase diagrams to illustrate these different behaviours, colour-coded to match the specified parameters' location in Figure 3.3.1a. The key property of interest to this paper is the emergence of cycles as the capacity of hype investors increases.

the stability of each steady state in Appendix 3.A.2, using the eigenvalues of the Jacobian evaluated at the steady states.

Figure 3.3.1a shows the different steady states that emerge within our system from the Jacobian for different parameter values of α/λ and $C\beta$ (with the precise derivation and a more detailed discussion on various types of steady states presented in the **Methods** section). Qualitatively different behaviours emerge within the different stability regions of Figure 3.3.1a. These are illustrated in Figure 3.3.1b, where the paths display cyclicity in all regions except (D) and (H), quasi-periodic dynamics in regions (A), (B), (E), (F), or direct convergence in regions (D), (H).

The recent impact of WSB

What has happened to the markets with the influx of eager retail traders onto investment-centric discussion forums? Despite investor discussion platforms having some traction from the early 2000s (with niche groups of investors using myForexBook, StockTwits, or SeekingAlpha), the rise to prominence of the WSB subreddit forum constituted the first widespread coordination attempt among retail traders. By the end of 2023, the forum had gained 15M followers.¹ Shortly after the GameStop short squeeze, news outlets noted the shift in the investment climate, such as the fact that one in five investors claimed that they had used Reddit to drive an investment decision.²

For this paper, we estimate the key parameters in our dynamic system at various points in time and for different assets discussed on WSB. Several parameters we expect to remain relatively stable over time and are, therefore, of low interest to this paper. These are not driven by social media, but rather characteristics of the investor population. Using survey data, we assume that the average hype investor holds \$10,000 in funds.^{3,4} We estimate λ to be approximately one.

Other parameters, such as the size of the hype investor population N , the degree of trend-following β , and social contagion α , are characteristics that change across time and vary between different assets in our data. We use a network approach to estimate them from WSB data, as described in our **Methods** section.

We are able to observe α and β for users active on WSB, as well as for specific stocks. We take into consideration the evolution in the number of users N discussing an asset as a function of returns and sentiments, endogenously increasing capacity C . An adapted epidemic contagion model serves particularly well in predicting the number of new discussants for an asset, based on the number of users previously discussing it and its returns. These findings are detailed in Appendix 3.C. Variation in user count compels us to vary C , and gauge what dynamics dominate market stability using our model.

¹<https://subredditstats.com/r/wallstreetbets>

²<https://markets.businessinsider.com/news/stocks/reddit-retail-investor-trend-survey-social-media-wallstreetbets-wsb-stock-2021-4-1030366600>

³<https://andriymulyar.com/blog/how-a-subreddit-made-millions-from-covid19>

⁴https://www.reddit.com/r/wallstreetbets/comments/a2qvj3/wsb_demographics_survey_results_a_portrait_of/

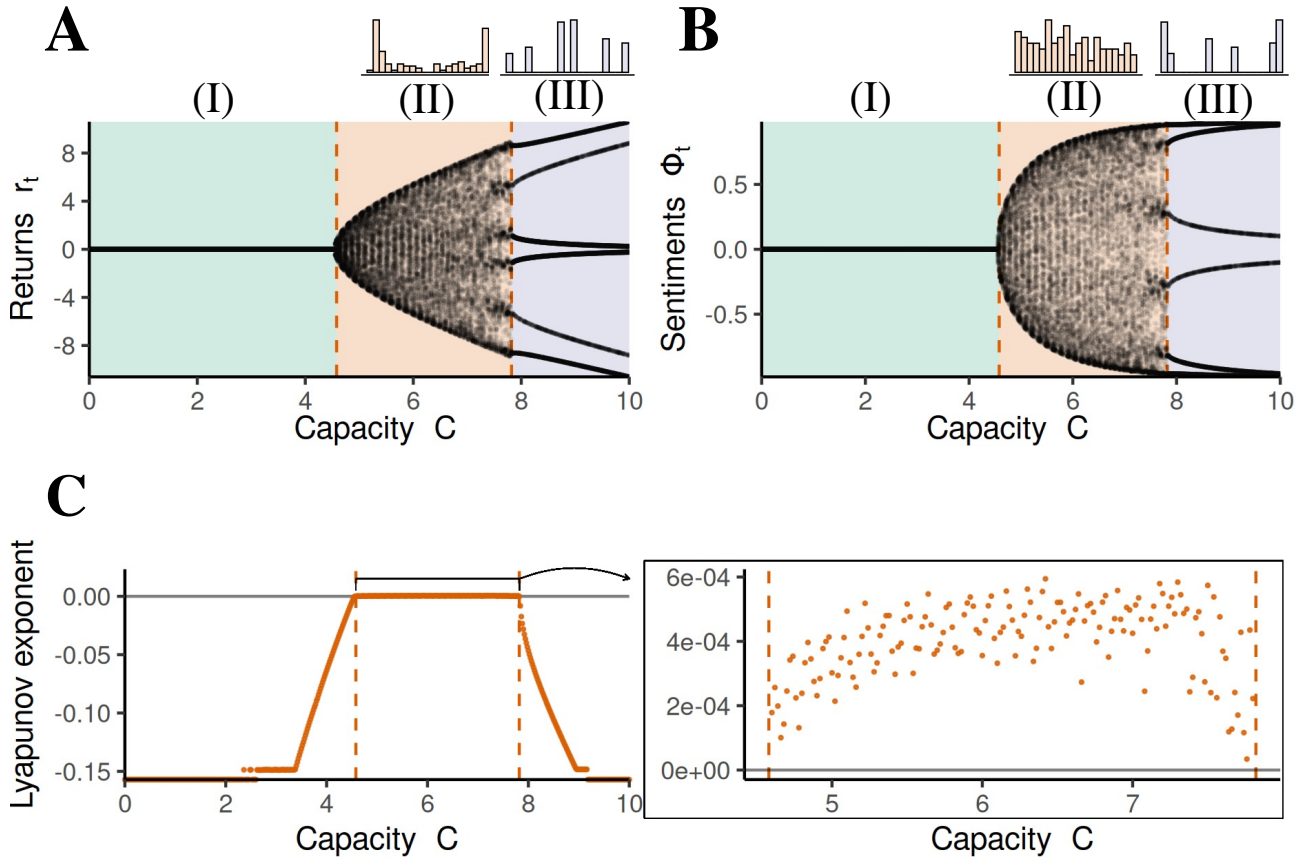


Figure 3.3.2: **Sentiment and Return Bifurcation Diagrams:** The final value from the 1000th timestep of the dynamic system in Eqs. 3.6-3.7 is plotted for the average α and β values observed on WSB ($\alpha = 0.05$ and $\beta = 0.22$) with different values for C in plots A and B. Plot A displays final values for returns, while B displays sentiments. Capacity C varies between zero and ten. For each C value, we perform 100 simulations with initial values for r drawn from a normal distribution with mean zero and standard deviation of 0.05, and with ϕ initiated at zero. We plot the distribution of the returns and sentiments for all 100 simulations at the 1000th time step for $C = 6$ above regions (II) and $C = 8$ above region (III). Plot C displays the highest Lyapunov exponent estimated using the Jacobian Matrix algorithm, as described in Wu & Baleanu (2015).

Typical system behaviour Across all assets in our sample, we find an average value of $\alpha = 0.05$ and $\beta = 0.22$ (both coefficients are significant at the 1% level). Figure 3.3.2 displays the resulting price and sentiment dynamics that occur as we vary C . For each value of C , we display the final returns r_t and sentiments ϕ_t from 100 different simulations at the 1000th time step. The simulations are initialised with r_0 drawn from a normal distribution with mean zero and standard deviation of 0.05. The goal is to better understand how a small perturbation from the known steady-state of $r = 0$ and $\phi = 0$ may impact price and sentiments when social investors are present.

Figure 3.3.2 demonstrates that capacity C would have to be unreasonably large – total funds of WSB users must be over four times greater than the asset’s market capitalisation – for the system to fall out of stable region (I), and into chaotic region (II). These parameters reside largely in regions (II) and (III) from Figure 3.3.2.

Figure 3.3.2’s region (II) displays quasi-periodic / chaotic dynamics. The system’s be-

haviour does not precisely repeat itself, and the return time-series appears irregular. However, the concentration of final values of r_{1000} and ϕ_{1000} demonstrates a higher probability for us to observe extreme values for r_t and ϕ_t (a uniform distribution for r_{1000} and ϕ_{1000} would have suggested chaotic dynamics). The degree of periodicity in this region varies for different values of C .

In region (III) of Figure 3.3.2, the system enters a stable cycle. It moves between several different states, without visiting intermediate values for r and ϕ .

Following our model, hype investors can destabilise markets. However, in the typical case this would require them to have implausibly large capacity, in the order of four times an asset's market capitalisation. As we show below, that does not rule out exceptional circumstances in which WSB users displayed large degrees of social contagion, or trend following.

System behaviour for specific assets Have hype investors actually destabilised the markets? From our average estimates, it is clear that WSB users would have to have access to vast amounts of capital to cause market fluctuations. However, the picture is more nuanced on a granular level. To demonstrate this heterogeneity, we estimate α and β for specific stocks.

We present a sample of our estimates in Table 3.3.1. Within our sample, we observe that WSB users exhibit varied degrees of trend following and social contagion. In the case of AAPL, for example, trend following appears to play a significant role. For GME, both trend following and social contagion have modest, but persistent effect. The outsized market capitalisation of AAPL relative to GME, as well as the number of active discussants, made an important difference.

Table 3.3.1: Estimates for α and β for different assets

Ticker	Name	Market Cap. (Dec. 2020, \$ Millions)	α	β
VXRT	Vaxart, Inc.	837	2.05(0.21)***	0.77(3.62)
TLRY	Tilray Brands, Inc.	1,078	0.35(0.07)***	0.38(0.70)
PCG	PG&E Corp.	25,184	0.29(0.10)***	2.35(2.03)
GME	GameStop Corp.	1,030	0.07(0.01)***	0.13(0.04)***
EA	Electronic Arts, Inc.	36,881	0.42(0.12)***	27.78(10.51)***
AMD	Advanced Micro Devices, Inc.	111,407	0.09(0.03)***	1.62(0.91)*
AMC	AMC Entertainment Holdings, Inc.	355	0.07(0.02)***	-0.08(0.14)
AAPL	Apple, Inc.	2,086,461	0.12(0.04)***	9.44(3.40)***
MU	Micron Technology, Inc.	74,922	0.01(0.72)	4.95(0.00)***
RAD	Rite Aid Corp.	730	0.06(0.47)	3.40(0.01)***

Notes: This table lists the parameter estimates for α and β for specific stocks in our WSB sample. Standard errors (calculated while accounting for autocorrelation with one lag and heteroskedasticity) are presented in brackets. For implementation please reference [Seabold & Perktold \(2010\)](#).

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

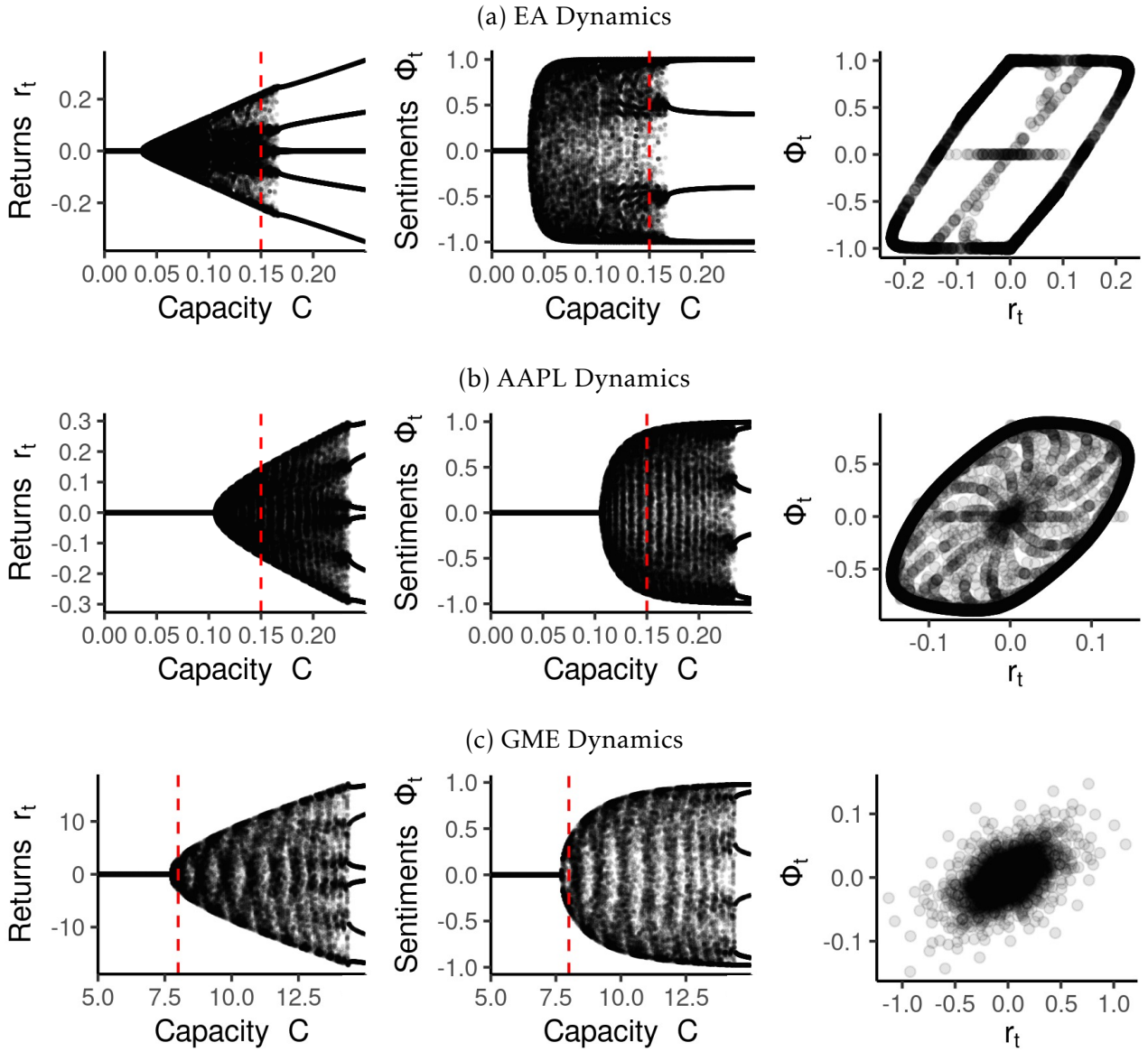


Figure 3.3.3: Specific stock returns and sentiments entered unstable territory: This figure illustrates the relationship between sentiments and returns generated by the model, using parameter values estimated for stocks EA, AAPL and GME. Parameter values are reported in Table 3.3.1. The high values estimated for capacity C produce a cyclical relationship between sentiments and returns, which never converge to a fixed point despite the absence of stochastic noise.

Figure 3.3.3 demonstrates how specific values for α and β can lead to substantially different dynamics. In the case of AAPL and EA, for example, the combination of a high degree of trend-following and social contagion results in a bifurcation occurring when social investors possess much lower capacity. In the case of EA, we observe destabilising impacts when the total capacity C of social investors is less than 5% of the stock's market cap – a quantity not beyond the realm of imagination, given that WSB reached 15M followers by 2023.

In the case of GME, the lower estimates for α and β imply stable asset price dynamics until hype investors reached a certain capacity. However, after a certain capacity threshold is exceeded for hype investors, the asset returns and investor sentiments undergo a bifur-

cation. This hypothesis is consistent with the data, where GME returns appeared relatively stable until the WSB discussion forum increased in size and focused its attention on GME. At this point, the time-series of returns experienced a changepoint, as displayed in Figure 3.1.1. These returns again reverted to a more stable regime once hype investors lost interest in the asset, and their overall capacity diminished.

3.4 Discussion

This paper uses a complex systems approach to understand market instability in the presence of social forces. This approach is of particular relevance as a greater proportion of retail investors turns to social media for stock market advice. The market, in the presence of such investors, can be effectively modelled through the interactions of different actors, resulting in potentially destabilising collective dynamics.

This approach helps complement traditional models in finance, by demonstrating how various parameters can produce different equilibria. We isolated the key parameters driving the behaviour of an asset when hype investors are present. The hype investors' capacity, with the degree of trend following and social contagion all play an important role. One component we did not address in our empirical evaluation is the amount of noise in the investor decision-making process, although it is also central to the model in amplifying the effect of social contagion. We identify eight distinct stability regions all exhibiting different emergent return and sentiment behaviours – from stable nodes to periodic and quasi-periodic dynamics.

The second contribution of our paper is to reconcile our model with data from the infamous WallStreetBets forum. We use estimates for those key investor characteristics, such as the approximate amount of funds each user holds, as well as the degree of trend-following and social contagion. While trend-following and social contagion play a significant role in sentiment formation, the magnitude of these effects is relatively small on average. As a result, WSB users required significant capacity in order to destabilise trading in a given asset, which emerges as bifurcations for returns and sentiments in the model.

However, we distinguish between different assets to demonstrate that the parameters governing sentiments can vary significantly. In light of our model, we conclude that WSB users occasionally destabilised trading in specific assets despite relatively limited capacity, thanks to a greater tendency to coordinate their asset purchases through sharing sentiments.

3.5 Methods

3.5.1 Finding the steady states

We outline the steps taken to find our system steady states. By definition, r_t must be zero at the steady state. This means that steady states of the dynamic system in Eqs. 3.6-3.7 are those for which ϕ_t is a solution to Eq. 3.6, which is the hyperbolic tangent function displayed

in the upper right hand corner of Figure 3.3.1a. Zero is a unique steady state when $\alpha/\lambda < 1$, and two further steady states emerge when $\alpha/\lambda > 1$. Those two additional steady states are solved numerically in all simulation exercises, using a solving algorithm.

In a discrete time system, the type of stability around a steady state is dependent on the eigenvalues of the Jacobian matrix at the steady state (general stability patterns with a single eigenvalue are discussed in greater detail in Appendix 3.A.2). When a system has multiple eigenvalues, different behaviours can emerge. We use the Jacobian matrix at the different steady states $(0, 0)$, $(\phi, 0)$.

The Jacobian matrices at the different steady states are

$$J(0, 0) = \begin{bmatrix} \alpha & \beta \\ C(\alpha - 1) & C\beta \end{bmatrix}, \quad J(\phi, 0) = \begin{bmatrix} \alpha \operatorname{sech}^2(\alpha\phi) & \beta \operatorname{sech}^2(\alpha\phi) \\ C(\alpha \operatorname{sech}^2(\alpha\phi) - 1) & C\beta \operatorname{sech}^2(\alpha\phi) \end{bmatrix}, \quad (3.8)$$

where $\phi \in \{\phi^+, \phi^-\}$. The corresponding eigenvalues for steady state $(0, 0)$ are

$$x_1 = \frac{C\beta + \alpha + \sqrt{(C\beta + \alpha)^2 - 4C\beta}}{2}, \quad x_2 = \frac{C\beta + \alpha - \sqrt{(C\beta + \alpha)^2 - 4C\beta}}{2}, \quad (3.9)$$

and those for steady states $(\phi, 0)$ are

$$x_1 = \frac{\operatorname{sech}^2(\alpha\phi)(C\beta + \alpha) + \sqrt{\operatorname{sech}^4(\alpha\phi)(C\beta + \alpha)^2 - 4C\beta \operatorname{sech}^2(\alpha\phi)}}{2}, \quad (3.10)$$

$$x_2 = \frac{\operatorname{sech}^2(\alpha\phi)(C\beta + \alpha) - \sqrt{\operatorname{sech}^4(\alpha\phi)(C\beta + \alpha)^2 - 4C\beta \operatorname{sech}^2(\alpha\phi)}}{2}. \quad (3.11)$$

These expressions trace out different regions of stability as a function of $C\beta$ and α in Figure 3.3.1a.

Figure 3.3.1a is drawn by solving numerically for the points $C\beta$ for which the eigenvalues in Eq. 3.9, on the left, and Eqs. 3.10-3.11, on the right, switch between the different stability regions. Specifically, we solve for when the complex part of the eigenvalues are equal to zero or not equal to zero, and are within or outside the unit circle. These points are computed by taking a range of values for α/λ , then applying a solving algorithm. These solutions are then used to distinguish the different regions of stability in Figure 3.3.1a. We discuss the different types of stability regions in greater detail in Appendix 3.A.2.

3.5.2 Estimating Parameters from WSB

Data We use the raw data from [Semenova et al. \(2024\)](#) as our data source of submissions published on WSB from the forum's creation in 2012 through the period following the GameStop short-squeeze in mid-2021. We label sentiments using the FinBERT large language model (LLM), fine-tuned on finance dialogue [Araci \(2019\)](#) achieving an accuracy of 92% on our test set (various other methods were tested, including more recent LLMs and the less computationally intense RegEx model, yielding little improvement); assets are extracted using the three / four capital letter combination (often prefaced with a '\$' sign)

representing tickers. A unique author id for the user creating a specific post is available in our data. Using these techniques, we are able to extract the log-likelihood of a post by author i about asset j at time t expressing bullish over bearish sentiments:

$$\Phi_{i,j,t} = \frac{1}{2} \log \left(\frac{P(\phi_{i,j,t} = +1)}{P(\phi_{i,j,t} = -1)} \right), \quad (3.12)$$

where the probabilities for a post to be bullish ($P(\phi_{i,j,t} = +1)$) or bearish ($P(\phi_{i,j,t} = -1)$) are a direct output of our LLM.

Daily asset returns and prices are sourced from WRDS CRSP: the log-return of asset j on day t are $r_{j,t} = \log(p_{j,t}/p_{j,t-1})$, where $p_{j,t}$ is the price of asset j when markets close on day t . Asset returns and WSB posts are matched by trading day: if the post occurs before 4:00PM EST (American financial market closing time) on a given day, the post is considered to have occurred within that trading day; if it occurs after 4:00PM EST – it is matched with the following trading day.

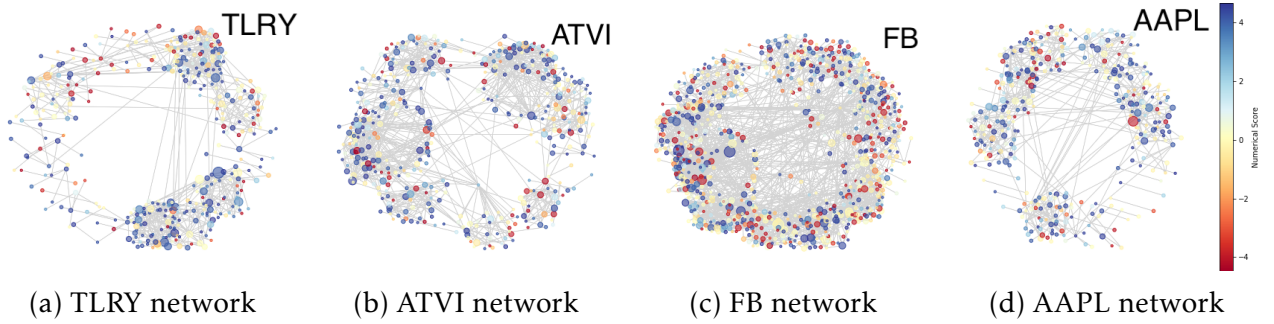


Figure 3.5.1: Sentiment Spillover Networks: We display the sentiments expressed within submissions on WSB (nodes) and the connections between nodes, representing whether the author of the later submission commented on the previous one. The TLRY and AAPL networks have high network spillovers where the sentiments of earlier submissions appear to heavily influence those of future, linked submissions. The ATVI network, on the other hand, has a strong trend-following component: we see distinct clusters of submissions with different sentiments, however, these are likely driven by movements in the underlying asset price. The FB network does not exhibit strong network spillovers or trend following – submissions with different sentiments are homogeneously mixed. The colour bar displays that red submissions correspond to bearish sentiments, blue to bullish and yellow to neutral.

Estimating α and β Two parameters that are of key importance for our dynamic system are α (the degree of sentiment spillovers) and β (the degree of trend following). We estimate both using market data in combination with data from WSB.

First, we construct a directed network tracking how older submissions might influence future submissions. An older submission about a specific asset is linked to a future one if the author of the future submission comments on the older submission – this network construction method allows us to observe which submissions an author interacts with and might influence their opinion before creating their own post. The author's own, previous submissions about the asset are considered as a separate control variable. As described above and presented in Figure 3.5.1, the networks are asset-specific.

We estimate α and β with the following linear model:

$$\Phi_{i,j,t} = \alpha \bar{\Phi}_{j,t} + \beta_1 r_{j,t} + \beta_2 \bar{r}_{j,t-1} + X_{i,j,t}K + \epsilon_{i,j,t}, \quad (3.13)$$

where $X_{i,j,t}$ represents the vector of control variables and is composed of stock-specific fixed effects, author i 's past sentiment about asset j (recorded as a categorical variable), and the variance in log returns on the five days prior to day t ; K is a vector of corresponding coefficients. We estimate this regression across all assets in our sample together for which we find $\alpha = 0.05$ and $\beta_1 = 0.22$ (both significant at the 1% level across different standard error estimation techniques), as well as for individual assets to compute the results presented in Table 3.3.1. The full regression estimates are presented in Appendix 3.B.

Bibliography

- Agrawal, S., Azar, P. D., Lo, A. W. & Singh, T. (2018), 'Momentum, mean-reversion, and social media: Evidence from Stocktwits and Twitter', *The Journal of Portfolio Management* 44(7), 85–95.
- Antweiler, W. & Frank, M. Z. (2004), 'Is all that talk just noise? the information content of internet stock message boards', *The Journal of finance* 59(3), 1259–1294.
- Araci, D. (2019), 'FinBERT: Financial sentiment analysis with pre-trained language models', *arXiv preprint arXiv:1908.10063*.
- Avery, C. & Zemsky, P. (1998), 'Multidimensional uncertainty and herd behavior in financial markets', *American Economic Review* pp. 724–748.
- Azar, P. D. & Lo, A. W. (2016), 'The wisdom of Twitter crowds: Predicting stock market reactions to FOMC meetings via Twitter feeds', *The Journal of Portfolio Management* 42(5), 123–134.
- Bacaksizlar Turbic, N. G. & Galesic, M. (2023), 'Group threat, political extremity, and collective dynamics in online discussions', *Scientific Reports* 13(1), 2206.
- Banerjee, A., Chandrasekhar, A. G., Duflo, E. & Jackson, M. O. (2013), 'The diffusion of microfinance', *Science* 341(6144).
- Banerjee, A. V. (1993), 'The economics of rumours', *The Review of Economic Studies* 60(2), 309–327.
- Becker, J., Brackbill, D. & Centola, D. (2017), 'Network dynamics of social influence in the wisdom of crowds', *Proceedings of the national academy of sciences* 114(26), E5070–E5076.
- Black, F. (1986), 'Noise', *The Journal of Finance* 41(3), 528–543.
- Bouchaud, J.-P. (2013), 'Crises and collective socio-economic phenomena: simple models and challenges', *Journal of Statistical Physics* 151(3), 567–606.

- Bradley, D., Hanousek Jr., J., Jame, R. & Xiao, Z. (2021), Place your bets? the market consequences of investment research on Reddit's Wallstreetbets, Technical report, SSRN.
- Brock, W. A. & Durlauf, S. N. (2001), 'Discrete choice with social interactions', *The Review of Economic Studies* **68**(2), 235–260.
- Bursztyn, L., Ederer, F., Ferman, B. & Yuchtman, N. (2014), 'Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions.', *Econometrica* **82**(4), 1273.
- Chen, H., De, P., Hu, Y. J. & Hwang, B.-H. (2014), 'Wisdom of crowds: The value of stock opinions transmitted through social media', *The Review of Financial Studies* **27**(5), 1367–1403.
- Chen, H. & Hwang, B.-H. (2022), 'Listening in on investor's thoughts and conversations', *Journal of Financial Economics* **145**(2, Part B), 426–444.
- Cont, R. & Bouchaud, J.-P. (2000), 'Herd behavior and aggregate fluctuations in financial markets', *Macroeconomic Dynamics* **4**(2), 170–196.
- Cookson, J. A., Engelberg, J. E. & Mullins, W. (2022), 'Echo Chambers', *The Review of Financial Studies* **36**(2), 450–500.
- Farmer, J. D. (2002), 'Market force, ecology and evolution', *Industrial and Corporate Change* **11**(5), 895–953.
- Heimer, R. Z. (2016), 'Peer pressure: Social interaction and the disposition effect', *The Review of Financial Studies* **29**(11), 3177–3209.
- Hirshleifer, D. (2001), 'Investor psychology and asset pricing', *The Journal of Finance* **56**(4), 1533–1597.
- Hommes, C. (2013), *Behavioral Rationality and Heterogeneous Expectations in Complex Economic Systems*, Cambridge University Press.
- Hommes, C. (2021), 'Behavioral and experimental macroeconomics and policy analysis: A complex systems approach', *Journal of Economic Literature* **59**(1), 149–219.
- Hu, D., Jones, C. M., Zhang, V. & Zhang, X. (2021), The rise of Reddit: How social media affects retail investors and short-sellers' roles in price discovery, Technical report, SSRN.
- Jusup, M., Holme, P., Kanazawa, K., Takayasu, M., Romić, I., Wang, Z., Geček, S., Lipić, T., Podobnik, B., Wang, L. et al. (2022), 'Social physics', *Physics Reports* **948**, 1–148.
- Kahneman, D. & Tversky, A. (1972), 'Subjective probability: A judgment of representativeness', *Cognitive Psychology* **3**(3), 430–454.
- Kirman, A. (1993), 'Ants, rationality, and recruitment', *The Quarterly Journal of Economics* **108**(1), 137–156.

- Kubin, E. & Von Sikorski, C. (2021), 'The role of (social) media in political polarization: a systematic review', *Annals of the International Communication Association* **45**(3), 188–206.
- Kuchler, T. & Stroebe, J. (2021), 'Social finance', *Annual Review of Financial Economics* **13**, 37–55.
- Lahno, A. M. & Serra-Garcia, M. (2015), 'Peer effects in risk taking: Envy or conformity?', *Journal of Risk and Uncertainty* **50**(1), 73–95.
- Lamberson, P., Page, S. E. et al. (2012), 'Tipping points', *Quarterly Journal of Political Science* **7**(2), 175–208.
- Lucchini, L., Aiello, L. M., Alessandretti, L., De Francisci Morales, G., Starnini, M. & Baronchelli, A. (2022), 'From Reddit to Wall Street: The role of committed minorities in financial collective action', *Royal Society Open Science* **9**(4), 211488.
- Lux, T. (1998), 'The socio-economic dynamics of speculative markets: interacting agents, chaos, and the fat tails of return distributions', *Journal of Economic Behavior & Organization* **33**(2), 143–165.
- MacKinnon, J. G. & White, H. (1985), 'Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties', *Journal of Econometrics* **29**(3), 305–325.
- Mancini, A., Desiderio, A., Di Clemente, R. & Cimini, G. (2022), 'Self-induced consensus of Reddit users to characterise the Gamestop short squeeze', *Scientific Reports* **12**(1), 13780.
- Moser, S. C. & Dilling, L. (2007), 'Toward the social tipping point: Creating a climate for change', *Creating a climate for change: Communicating climate change and facilitating social change* pp. 491–516.
- Nguyen, L. T., Wu, P., Chan, W., Peng, W. & Zhang, Y. (2012), Predicting collective sentiment dynamics from time-series social media, in 'Proceedings of the first international workshop on issues of sentiment discovery and opinion mining', pp. 1–8.
- Sabherwal, S., Sarkar, S. K. & Zhang, Y. (2011), 'Do internet stock message boards influence trading? Evidence from heavily discussed stocks with no fundamental news', *Journal of Business Finance & Accounting* **38**(9-10), 1209–1237.
- Scholl, M. P., Calinescu, A. & Farmer, J. D. (2021), 'How market ecology explains market malfunction', *Proceedings of the National Academy of Sciences* **118**(26), e2015574118.
- Seabold, S. & Perktold, J. (2010), statsmodels: Econometric and statistical modeling with python, in '9th Python in Science Conference'.
- Semenova, V., Gorduza, D., Wildi, W., Dong, X. & Zohren, S. (2024), 'Wisdom of the crowds or ignorance of the masses? a data-driven guide to WallStreetBets.', *Journal of Portfolio Management* **50**(4).

- Shiller, R. J. (1984), Stock prices and social dynamics, Technical Report 2, The Brookings Institution.
- Shiller, R. J. (2005), *Irrational Exuberance: (Second Edition)*, Princeton University Press.
- Shiller, R. J. (2017), 'Narrative economics', *American Economic Review* **107**(4), 967–1004.
- Wu, G.-C. & Baleanu, D. (2015), 'Jacobian matrix algorithm for Lyapunov exponents of the discrete fractional maps', *Communications in Nonlinear Science and Numerical Simulation* **22**(1-3), 95–100.

Appendix

3.A Model appendix

3.A.1 Deriving aggregate buying intensity and market impact

Individual investor buy versus sell decisions The utility an investor derives from a specific sentiment is a function of other hype investors' average buying intensity ϕ_t , asset returns r_t , as well as an error ϵ_i :

$$U_i(d_{i,t+1} = +1) = \epsilon_i^+ + \alpha\phi_t + \beta r_t - \gamma r_t^2, \quad U_i(d_{i,t+1} = -1) = \epsilon_i^- - \alpha\phi_t - \beta r_t - \gamma r_t^2.$$

where β is a trend following component, α is the social network spillover parameter among hype investors, and γ captures aversion to large swings in price. For simplicity, these parameters are assumed common to all investors in the bullish and bearish case. We abstract away from cognitive biases that may, for example, drive greater herding behaviour in a downturn.

ϵ_i^+ and ϵ_i^- refer to a random, exogenous element in individual preferences. We make Assumption 3.1 about ϵ_i^+ and ϵ_i^- , which will allow us to find an expression for aggregate sentiment.

Assumption 3.1. ϵ_i^+ is independent and identically distributed, following a type-I Extreme Value (EV) distribution. The same holds for ϵ_i^- .

Assumption 3.1 is standard in the literature [Bouchaud \(2013\)](#), and justifiable in our present context since investors make a choice governed by a maximisation process.

Aggregating hype investors An assumption on the distribution of the idiosyncratic terms ϵ_i^+ and ϵ_i^- (Assumption 3.1) allows us to derive an expression for the probability of bullish

versus bearish sentiment. We write the probability of bullish sentiment as

$$P[U_i(+1) > U_i(-1)] = P(\epsilon_i^+ + \alpha\phi_t + \beta r_t - \gamma r_t^2 > \epsilon_i^- - \alpha\phi_t - \beta r_t - \gamma r_t^2), \quad (3.14)$$

$$= P(\epsilon_i^- < \epsilon_i^+ + 2\alpha\phi_t + 2\beta r_t), \quad (3.15)$$

$$= \int_{-\infty}^{+\infty} P(\epsilon_i^- < \epsilon_i^+ + 2\alpha\phi_t + 2\beta r_t | \epsilon_i^+) \times p(\epsilon_i^+) d\epsilon_i^+, \quad (3.16)$$

$$= \int_{-\infty}^{+\infty} \exp\left[-\exp\left(-\frac{\epsilon_i^+ + 2\alpha\phi_t + 2\beta r_t}{2\lambda}\right)\right] \times \frac{1}{2\lambda} \exp\left[-\frac{\epsilon_i^+}{2\lambda} - \exp\left(-\frac{\epsilon_i^+}{2\lambda}\right)\right] d\epsilon_i^+, \quad (3.17)$$

$$= \frac{\exp\left(\frac{\alpha\phi_t + \beta r_t}{\lambda}\right)}{\exp\left(\frac{\alpha\phi_t + \beta r_t}{\lambda}\right) + 1}, \quad (3.18)$$

where step 3.17 follows from the fact that we are integrating the cumulative density of ϵ_i^- , with scale 2λ , over the support of ϵ_i^+ , with marginal density with marginal density $p(\epsilon) = \exp[-\epsilon/2\lambda - \exp(-\epsilon/2\lambda)]/2\lambda$. The final simplification comes from the fact that $U(d_{i,t+1} = +1)$ and $U(d_{i,t+1} = -1)$ are symmetric in parameters α, β and γ .

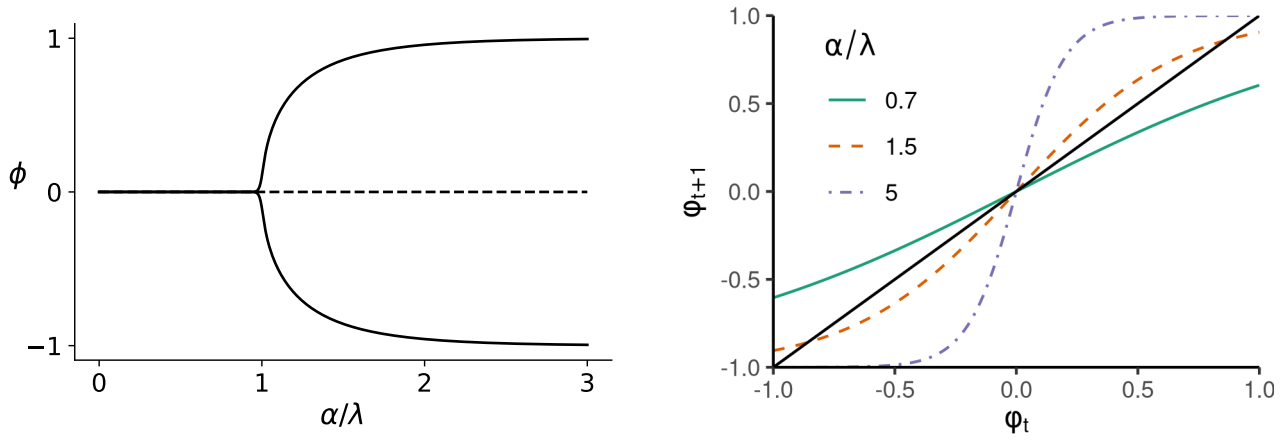
If we want to keep them distinct, we arrive at the quantal response function for the probability of positive or negative sentiment:

$$\phi_{i,t+1}^D = \begin{cases} +1, & \text{with probability } \frac{\exp[(\alpha\phi_t + \beta r_t - \gamma r_t^2)/2\lambda]}{\exp[(\alpha\phi_t + \beta r_t - \gamma r_t^2)/2\lambda] + \exp[(-\alpha\phi_t - \beta r_t - \gamma r_t^2)/2\lambda]}, \\ -1, & \text{with probability } \frac{\exp[(-\alpha\phi_t - \beta r_t - \gamma r_t^2)/2\lambda]}{\exp[(\alpha\phi_t + \beta r_t - \gamma r_t^2)/2\lambda] + \exp[(-\alpha\phi_t - \beta r_t - \gamma r_t^2)/2\lambda]}. \end{cases}$$

Therefore, we can model the aggregate buying intensity by summing $\phi_{i,t+1}^D$ across the N hype investors, which yields the well-known hyperbolic tangent function:

$$\phi_{t+1} = \tanh\left[\frac{\beta r_t + \alpha\phi_t}{\lambda}\right]. \quad (3.19)$$

Properties of the hyperbolic tangent function The function in Eq. 3.19 is known as the *hyperbolic tangent* function, which has received significant attention in the area of strategic decision-making [Brock & Durlauf \(2001\)](#), [Bouchaud \(2013\)](#). We highlight the key properties of the hyperbolic tangent function in our application, as they are important components of our model's stability. When α/λ is between zero and one, the function produces one stable steady state for $\phi_{t+1} = \phi_t$ at zero. Two stable steady states, and one unstable one at zero, emerge when α/λ is greater than one, as shown in Figure 3.A.1a. Figure 3.A.1 presents a visualisation of the pitchfork bifurcation, as well as the temporal evolution, of the hyperbolic tangent function [Hommes \(2013\)](#). The stability of the system is critically dependent on α/λ . When the social signal $\alpha\phi_t$ is high relative to noise λ , investor coordination is effective and all hype investors converge on a similar investment strategy. On the other hand, if α/λ is low, then noise overwhelms the social signal – coordination fails and the zero steady state is the only one that is ever reached.



(a) Steady states of the hyperbolic tangent function; we plot ϕ when the system has converged after 100 time steps for different values of α/λ and ϕ_0 .

(b) Hyperbolic tangent function in ϕ_t ; we plot $\phi_{t+1} = \tanh(\alpha\phi_t/\lambda)$ for three values of α/λ . The function crossing the 45 degree line (solid black) denotes a steady state.

Figure 3.A.1: Properties of the hyperbolic tangent: Figure 3.A.1a shows the bifurcation of the hyperbolic tangent at $\alpha/\lambda = 1$. When $\alpha/\lambda < 1$, $\phi = 0$ is the unique, stable steady state. When $\alpha/\lambda > 1$, two new stable steady states emerge, ϕ^+ and ϕ^- , and $\phi = 0$ becomes unstable. Figure 3.A.1b shows the temporal behaviour of the hyperbolic tangent function for different α/λ . When $\alpha/\lambda = 0.7$, the function only crosses the 45 degree line (solid black) once at the origin. For values of 1.5, in dashed orange, and five, in dot-dashed purple, it crosses the 45 degree line in three places, indicating the emergence of two new steady states.

Deriving market impact Eq. 3.5 follows from the first difference of the market capitalisation:

$$\Delta(p_{t+1}Q) = \Delta(p_{t+1}Y_{t+1}) + \Delta(p_{t+1}S_{t+1}) = MN\Delta\phi_{t+1}, \quad (3.20)$$

where equality holds because $\Delta(p_{t+1}S_{t+1}) = 0$ and the expression for $p_{t+1}Y_{t+1}$ in Eq. 3.1. Rearranging, and substituting $r_{t+1} = \Delta p_{t+1}/p_t$ yields the desired result.

3.A.2 Stability Regions

Common system behaviours Table 3.A.1 shows some types of common behaviours around a single steady state, as they relate to the Jacobian eigenvalues. When a system has multiple eigenvalues, different behaviours can emerge. For example, if we consider two eigenvalues x_1 and x_2 and $|x_1| > 1$ while $|x_2| < 1$, we would observe a saddle.

Stability at (0,0) Steady state (0,0) displays five different types of behaviours as a function of parameters $C\beta$ and α . In region (A), both eigenvalues are real and greater than one. This means that (0,0) is an unstable node. In region (B), both eigenvalues are complex, and their modulus is greater than one; in this region, (0,0) is an unstable focus. In regions (C) and (D), both moduli are less than one, indicating that (0,0) is stable. The eigenvalues in (C), unlike those in (D), are complex, meaning that (0,0) is a focus instead of a node in that region. When $\alpha/\lambda > 1$, both eigenvalues are real, but one is greater than one, meaning that (0,0) is a

Table 3.A.1: System behaviour around steady states and Jacobian Eigenvalues

$ x_i < 1$	$\mathbb{I}(x_i) = 0$	stable node / sink
$ x_i < 1$	$\mathbb{I}(x_i) \neq 0$	stable focus / spiral sink
$ x_i > 1$	$\mathbb{I}(x_i) = 0$	unstable node / source
$ x_i > 1$	$\mathbb{I}(x_i) \neq 0$	unstable focus / spiral source

Notes: The behaviour of our system around our steady states is dependent on the eigenvalues of the Jacobian at the steady states. Here we describe some of the most common eigenvalue combinations and behaviours. For a more comprehensive review, see [Hommes \(2013\)](#).

saddle.

A change in capacity C can therefore be significant, in two qualitative ways. Assuming that consensus formation is relatively weak, such that $\alpha < 1$, an exogenous increase in capacity from a state initially close to zero introduces oscillations in the path to the stable node at $(0,0)$. The change in observed dynamics between regions (D) and (C) can be attributed to hype investors overreacting to large returns, thus overshooting the steady state in region (D). The ensuing change in sentiments fails to cause returns to diverge from their steady state in perpetuity, and investors correct their positions in the opposite direction in the subsequent period. A more consequential bifurcation happens when the change in capacity is large enough, so that $C\beta > 1$ and the parameters move from region (C) to region (B). This type of bifurcation is known as a ‘Hopf bifurcation’, as two complex eigenvalues cross the unit circle [Hommes \(2013\)](#).

Stability at $(\phi^+, 0)$ and $(\phi^-, 0)$ When consensus formation is relatively high ($\alpha > 1$), two new steady states emerge; one with positive sentiment $(\phi^+, 0)$ and one with negative sentiment $(\phi^-, 0)$. Both display identical behaviours in terms of stability, as a function of parameters $C\beta$ and α . These steady states only exist when $\alpha > 1$. In regions (E) and (F), the moduli of both eigenvalues are greater than one, but are real only in region (E). In contrast, both moduli are less than one in regions (G) and (H), and real only in region (H). This means that both steady states are stable focuses in region (G), and stable nodes in region (H).

Generally, the stronger consensus formation is in the dynamic system, the larger the combined trend following component and investor capacity has to be in order to destabilise these two steady states. Functionally, the historical signal is too weak to overcome the strong consensus effect, and hype investors hardly stray from their equilibrium sentiment. In this sense, social dynamics in the form of consensus formation may actually serve to stabilise markets.

3.B Hype investor sentiment spillovers

We present our full parameter estimates from Eq. 3.13 using several different methods for calculating standard errors. Our estimates are highly significant across the different methodologies used.

Table 3.B.1: Sentiment Spillover – full regression estimates

		Dependent Variable – $\Phi_{i,j,t}$		
		(1)	(2)	(3)
Independent Variables	$\phi_{i,j,t-1}^{-1}$	-0.85 (0.11)***	-0.85 (0.06)***	-0.85 (0.06)***
	$\phi_{i,j,t-1}^0$	-0.19 (0.08) **	-0.19 (0.03)***	-0.19 (0.03)***
	$\phi_{i,j,t-1}^{+1}$	0.67 (0.03) ***	0.67 (0.03)***	0.67 (0.03)***
	$\bar{\Phi}_{j,t}$	0.05 (0.01) ***	0.05 (0.01)***	0.05 (0.01)***
	$r_{j,t}$	0.22 (0.07) ***	0.22 (0.04)***	0.22 (0.04)***
	$\bar{r}_{j,t-1}$	0.44 (0.08) ***	0.44 (0.09)***	0.44 (0.09)***
	$\sigma_{j,t}^2$	0.04 (0.12)	0.04 (0.09)	0.04 (0.09)
	Ticker Fixed Effects	Yes	Yes	Yes
	No. Observations:	24,902	16,514	25,220
R^2 :		0.09	0.07	0.09
R_{adj}^2 :		0.06	0.06	0.06

Notes: We present the full regression estimates for Eq. 3.13. The dependent variable is individual investor sentiment about an asset expressed in a single submission, scaled continuously between $(-\infty, \infty)$. We estimate it using the individual's previously expressed sentiment about the same asset ($\phi_{i,j,t-1}$) as a categorical variable, with the author not having posted previously ($\phi_{i,j,t-1}^{NA}$) as the baseline. We control for a set of market control variables ($r_{j,t}, \bar{r}_{j,t}, \sigma_{j,t}^2$). The standard error is estimated in several ways. In column (1), we present the non-robust standard errors (which also match those estimated with a t-distribution), clustered at the ticker level. In column (2), we use heteroskedasticity-autocorrelation robust standard errors, clustered at the ticker level with one lag (the previous sentiment expressed about the same ticker on the forum in our network sample). In column (3), we use MacKinnon & White (1985) heteroskedasticity robust standard errors. Observations with incomplete data are dropped. Implementation details are provided by Seabold & Perktold (2010).

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

3.C Hype investor interest

In the WSB context, we would expect awareness about specific assets to spread from one user to another, in line with the observations of Shiller (2005), Banerjee (1993) and Banerjee et al. (2013). The emphasis of this section is not on identifying a causal relationship, but rather understanding the dynamics which govern asset interest among investors. These insights, combined with a mechanism for investors' joint sentiment adoption, allow us to paint a more complete picture of retail investor decision-making and the resultant stock market dynamics.

We model the log-odds of an author posting about stock j over a baseline using the following linear model:

$$l(a_{j,t}) = \log\left(\frac{a_{j,t}}{s_t}\right) = ca_{j,t-1}(1 - a_{j,t-1}) + da_{j,t-1} + \beta_1 \bar{r}_{j,t-1} + \beta_2 \sigma_{j,t-1}^2 + X_j \beta_4 + \zeta_{j,t}, \quad (3.21)$$

where t denotes time (in weeks), the baseline s_t is the probability of posting about a stock that is not widely discussed within the forum (a stock that is mentioned in fewer than 31 submissions within our sample), $a_{j,t-1}$ is the share of all active investors who post about

ticker j at times $t - 1$ ($a_{j,t} \in [0, 1]$ for all j and t), $\bar{r}_{j,t-1}$ is the average log-return in $t - 1$, and $\sigma_{j,t-1}^2$ is the variance of the same log-returns. X_j is a vector of stock dummies.

Our framework is inspired by [Banerjee et al. \(2013\)](#) – individuals become interested in an asset due to their peers and thanks to a public signal of the asset’s performance. Parameter c captures the rate of independent mixing between investors aware of stock j , $a_{j,t-1}$, with unaware investors, $1 - a_{j,t-1}$. Parameter d captures the rate at which aware investors lose interest. The latter terms control for the asset’s perceived profitability and riskiness. Parameter β_1 is a ‘quality of signal’ term capturing how well the asset has performed in the past, and β_2 a ‘noise of signal’ term, measuring the asset’s recent volatility. We propose that coefficients c and β_1 are positive – implying that these dynamics contribute to increased interest in a stock – while d and β_2 are negative.

The choice to aggregate over weeks is done to address the sparsity of submissions, especially pre-2017. In addition, we categorise stocks mentioned fewer than 31 times since January 2012 into an ‘other stocks’ group, which forms our benchmark s_t .

We also consider a different formulation where we test for the direct impact of historical peer sentiments and the interactions between historical sentiments and returns / volatility: $\bar{\phi}_{j,t-2}\bar{r}_{j,t-1}$, $\bar{\phi}_{j,t-2}$ and $\bar{\phi}_{j,t-2}\sigma_{j,t-1}^2$. This formulation allows us to evaluate whether WSB users are more likely to discuss a stock if the predictions of their peers have been correct, and accurate, in the past.

3.C.1 Results

The OLS estimates of our model in Eq. 3.21, presented in Table 3.C.1, demonstrate that WSB users follow each other in their choice of stocks. There is strong evidence that the homogeneous mixing property partially explains the uptake of new assets: using estimates in column (1), an increase in the share of authors discussing stock j from 0.1 to 0.2 increases the ratio of authors discussing j over ‘other stocks’ in the following week by approximately sevenfold. The effect tapers as attention saturates: a further increase from 0.2 to 0.3 raises the ratio of authors discussing j over ‘other stocks’ by only around 75% – the tapering is driven by the large negative coefficient on $a_{j,t-1}$. This is strongly reminiscent of epidemic contagion models, adapted to the spread of narratives [Banerjee \(1993\)](#), [Shiller \(2017\)](#).

When we consider the impacts of stock-specific variables in isolation, presented in columns (1), (3) in Table 3.C.1, volatility and returns appear to be leading factors for authors deciding what asset to discuss. Average, historical returns are statistically significant at the 1% level in columns (1) and (3), indicating that discussion sizes are stimulated by large, notably positive, returns. Examining the coefficient in column (3), a stock that experienced a 5% greater return in one week is the subject of about 10% more submissions than usual. Volatility appears to play a greater role in our formulation without ticker-specific effects, with its significance declining from column (1) to (3) – a factor perhaps explained by the choice of hype investors to overlook recent volatility in certain assets, but not others. Our alternative formulation presented in columns (2) and (4), estimating the effect of the correctness and consistency of past WSB predictions in an asset, appears to have some significance

Table 3.C.1: Stocks discussed on WSB

	<i>Dependent variable: $l(a_{j,t})$</i>			
	(1)	(2)	(3)	(4)
$a_{j,t-1}(1 - a_{j,t-1})$	71.58*** (5.62)	71.52*** (6.48)	41.32*** (7.63)	40.45*** (8.34)
$a_{j,t-1}$	-30.20*** (4.65)	-29.99*** (5.32)	-15.19** (6.04)	-14.75** (6.62)
$\bar{r}_{j,t-1}$	0.92*** (0.29)		1.83*** (0.26)	
$\sigma_{j,t-1}^2$	-1.52*** (0.30)		0.25 (0.36)	
$\bar{\phi}_{j,t-2}\bar{r}_{j,t-1}$		2.09** (0.83)		3.55*** (0.68)
$\bar{\phi}_{j,t-2}^2\sigma_{j,t-1}^2$		-4.69*** (1.77)		2.14 (1.61)
Constant	-4.02*** (0.01)	-4.04*** (0.01)		
Ticker FE	No	No	Yes	Yes
Number of obs.	25,637	12,708	25,637	12,708
Adjusted R ²	0.25	0.35	0.10	0.14
F-statistic	2,166.05	1,699.32	825.40	660.10

Notes: This table presents OLS estimates for the log-odds of users discussing stock j in week t , over a collection of stocks that are mentioned fewer than 31 times. Explanatory variables include: the lag in the share of authors discussing j , $a_{j,t-1}$, the interaction with the share of authors not discussing j , $a_{j,t-1}(1 - a_{j,t-1})$, as well as the lag in stock j 's weekly average log-return, $\bar{r}_{j,t-1}$, and variance, $\sigma_{j,t-1}^2$. In columns (2) and (4), the average log-return is multiplied by the two period lag in the average sentiment expressed among WSB submissions on stock j , $\bar{\phi}_{j,t-2}$, and the variance in log-returns by the same sentiment's square, $\bar{\phi}_{j,t-2}^2$. Columns (3) and (4) include stock-specific fixed effects. Accompanying standard errors, displayed in brackets, are clustered at the stock level, and calculated in the manner of [MacKinnon & White \(1985\)](#).

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

in explaining asset interest.

Chapter 4

The heavy tails in labour productivity

(With Jangho Yang, Torsten Heinrich, François Lafond, Pantelis Koutroumpis and J. Doyne Farmer)

*This chapter is an excerpt from a paper titled ‘Measuring productivity dispersion: a parametric approach using the Lévy alpha-stable distribution’, published in *Industrial and Corporate Change* (Yang et al. 2024).*

This excerpt documents the persistent shape of the distribution for labour productivity. I include a passage I was mostly responsible for, and which best fit the narrative of the overall thesis. It also features a software package I developed with Torsten Heinrich, *finit*. All further details are available in [Yang et al. \(2024\)](#).

This paper motivates much of this thesis, for two reasons. First, the shape of a distribution has meaningful implications for our interpretation of economic data. While economic research is mostly concerned with mean estimates – strategies to measure point elasticities, treatment effects, and so on – the frequency and magnitude of extremely productive firms suggest that aggregate patterns may not necessarily be related to the performance of some average, representative firm. Second, the shapes of distributions are too common to be random. Given that these shapes can matter, studying their emergence in economics constitutes a primary research goal of this thesis, one that I begin to address in Chapters 7 and 8.

4.1 Introduction

In the last two decades, the availability of micro-data has revealed the tremendous productivity differences between firms, even within sectors at a fairly detailed level (Bartelsman & Doms 2000, Syverson 2011). Measuring productivity dispersion is important in order to understand misallocation (Gopinath et al. 2017), innovation and diffusion (Berlingieri, Blanchenay & Criscuolo 2017, Andrews et al. 2019), business dynamism (Foster et al. 2021), and ultimately the aggregate productivity slowdown (Goldin et al. 2021).

In this paper, we focus on the distribution of value added (VA) per worker at the firm level, a common measure of productivity in the literature¹. We employ a commercial dataset, *Orbis Europe*. We have access to a comprehensive version, that includes around 23 million European firm-year observations for the period 2006–2017. The distribution of productivity levels, even within country-year pairs, exhibits heavy tails, and therefore an infinite variance. In practice, this means that measuring the variance in a given sample is not meaningful, as the result is driven by the sample size, rather than reflecting a true moment of the underlying population (see Fig. 4.1.1).

A common solution to this problem is to compute the variance of the *logarithm* of productivity (instead of the variance of productivity). For variables that have positive support, such as revenue per worker, this is an acceptable solution since the distribution of the log values likely exhibits finite variance. However, this is problematic for VA per worker, because firm-level datasets often contain a substantial proportion of firms that have *negative* value added - typically of only a few percent, but up to 23% for country-year pairs in our dataset (Table 4.2.1). To state the obvious, as far as measuring dispersion is concerned, removing the firms on the left tail is problematic. The prevalence of firms that stay alive despite very poor performance should be a key indicator in studies of misallocation, creative destruction, and business dynamism.

How can we measure dispersion in a dataset with negative values and heavy tails? In this paper, we propose a transparent and straightforward approach: first, finding a good parametric model for the distribution of productivity, and second, using the fitted parameters to evaluate dispersion. For instance, if the data was normally distributed, an excellent way to evaluate how dispersion differs across time or countries would have been to estimate the scale parameter of the normal distribution, that is, the variance. Here, however, we find

¹See for instance Souma et al. (2009), Aoyama et al. (2015), Andrews et al. (2019), Berlingieri, Blanchenay & Criscuolo (2017), Gu (2019), Ilzetzi & Simonelli (2017), Campbell et al. (2019), Gouin-Bonenfant (2020) and De Loecker et al. (2022). Baily et al. (1996) note that it is a ‘conceptually preferable measure of labour productivity’ to gross output per worker, even though they have to remove negative VA establishments in some of their analyses. Oulton (2000) discards roughly 1% of his sample of UK firms, which have negative value added. Aradanaz-Badia et al. (2017) note that only a third of the UK firms with negative VA in 2010 had exited the market by 2015, which is only 10 percentage points higher than other firms. De Loecker et al. (2022) find a proportion of negative VA UK firms in both Orbis and administrative data. While issues with VA-based measures of productivity have been noted (Campbell et al. 2019, Cunningham et al. 2021), measures based on gross output are less comparable across different sectors, as intermediate consumption is naturally larger in some industries than in others, so that using gross output-based measures creates a need to use industry means to renormalise the values and compare them; this itself comes with its own issues, if only because classification systems are imperfect, lack granularity, and change over time.

that the data exhibits heavy tails and substantial asymmetry, so we turn to more complex models. We find that the Lévy alpha-stable distribution gives an excellent fit to the data, and we argue that its parameters - particularly the scale and the tail parameters - provide appropriate metrics of dispersion.

The scale parameter of the Lévy distribution is similar to the standard deviation in a normal distribution. Roughly speaking, it provides information on the span of the support of the distribution where most of the data lies. In contrast, the tail exponent provides information on the prevalence of extremely productive or unproductive firms.

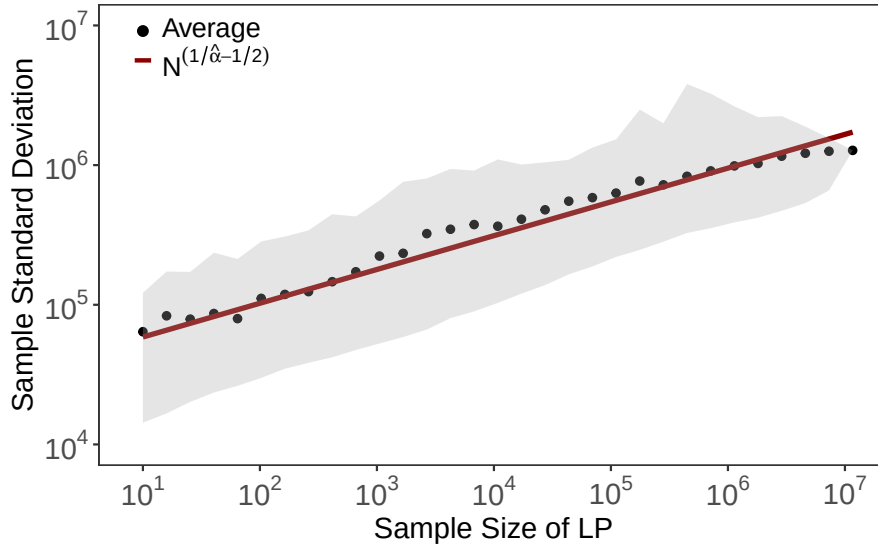


Figure 4.1.1: **Measured standard deviation in sub-samples of firm labour productivity:** We construct a distribution by pooling together the productivity levels of firms in France, Italy, Germany, and Spain (which have the same currency), for all years, expressed in thousands. Then, for each sub-sample size N , we compute the standard deviation of 100,000 sub-samples, and report the average (black dots), and 5th and 95th percentiles. In Appendix 4.E, we explain that on i.i.d. data with a power law tail, the theoretical scaling between the sample standard deviation and the sample size would be $N^{\frac{1}{\alpha}-\frac{1}{2}}$, where α is the tail exponent. The red line shows that this scaling holds empirically here (we use our estimate of the tail parameter $\hat{\alpha} = 1.33$, and the intercept is calculated by simulating 100,000 i.i.d. random samples of size 10 from a Lévy alpha-stable with the same parameters as those we estimate on empirical data). Data from Orbis Europe.

While the Lévy alpha-stable distribution elegantly addresses the issues of heavy tails and negative values, the problem with parametric measures is that the underlying model can be wrong. We address this in two ways. First, we show that the Lévy alpha-stable offers a surprisingly good fit to our productivity data, and that it is clearly better than the Subbotin (Asymmetric Exponential Power, AEP) distribution (Bottazzi & Secchi 2011), and as good as or slightly better than the Asymmetric Student-t (AST) distribution (Zhu & Galbraith 2010), both of which have one more parameter. We add credibility to the model by using independent tests to confirm that the second moment is infinite. Second, we argue that the Lévy alpha-stable model is plausible *a priori* simply because it is the result of the

Generalized Central Limit Theorem (GCLT) (Nolan 2020). If VA per worker can be thought of as the aggregation of several micro-level variables such as the productivity of individual employees, contracts, tasks or routines, then under fairly general conditions the GCLT predicts that firm-level productivity will follow a Lévy alpha-stable distribution.

Related literature Our results relate to three strands of literature. The first is the large body of literature on heavy tails in economics and finance (Mitchell 1915, Mandelbrot 1960, 1963, Fama 1965, Samuelson 1967, Embrechts et al. 1997, Fagiolo et al. 2008, Gabaix 2011, Acemoglu et al. 2017, Axtell forthcoming), more specifically on the distributions of firm sizes (Ijiri & Simon 1977, Axtell 2001) and firm growth rates (Bottazzi & Secchi 2003, Bottazzi et al. 2007, Bottazzi & Secchi 2011, Schwarzkopf et al. 2010, Holly et al. 2013, Moran et al. 2020). Most related to our study is the work on estimating productivity distributions such as Aoyama et al. (2010), who estimate a Generalized Beta model for labour productivity levels², Souma et al. (2009) who find a power law tail for labour productivity levels³, and Gaffeo (2008, 2011) who estimates a Lévy alpha-stable distribution for the growth rates of sector-level total factor productivity. In the present paper, we focus on *firm-level* data, productivity *levels*, and productivity measured as *value added per worker*.

Second, our results on productivity levels are relevant to the literature that discusses how productivity dispersion may reflect the misallocation of factors of production (Hsieh & Klenow 2009, Bartelsman et al. 2013, Foster et al. 2021, Gopinath et al. 2017, Haltiwanger et al. 2018). In this paper, we critically discuss measures of dispersion in detail, showing how some statistics may be misleading and what parts of the distribution can affect dispersion. We also suggest that theoretical models should aim to derive a Lévy alpha-stable distribution, where ideally the parameters would be interpreted in terms of misallocation, or other sources⁴.

Third, our results relate to the work of statistical agencies around the world, who are developing procedures to make summary statistics of the micro data publicly available, and in particular industry-level productivity dispersion (Cunningham et al. 2021, Berlingieri, Blanchenay, Calligaris & Criscuolo 2017). A key contribution of this paper is to make a suggestion regarding what statistics should be released, namely 5 quantiles necessary to perform McCulloch (1986)'s quantile estimation of the four Lévy alpha-stable parameters. Typically, readily available software packages use the 5th, 25th, 50th, 75th and 95th quantiles, as in McCulloch (1986)'s original paper. A remarkable consequence of the quality of the Lévy fit and the availability of the McCulloch (1986) quantile estimator is that we can in principle obtain a good estimate of the tail parameter by using only 5 quantiles, the largest of which is only the 95th quantile. No substantial risk of disclosing any confidential data

²Generalized Beta models are defined only over a positive support and thus are not suitable for our data, which contains negative productivity firms.

³See also Aoyama et al. (2009), Mizuno et al. (2012) and Aoyama et al. (2015).

⁴There is an important theoretical literature on productivity dispersion, which recognises that the right tail of the distribution of productivity levels is a power law, and explains this mostly as the result of innovation and imitation processes, see Ghiglino (2012), Lucas & Moll (2014), Perla & Tonetti (2014) and König et al. (2016). The presence of extreme values in empirical data also implies that data cleaning choices may affect substantive research results, as shown for the misallocation literature by Rotemberg & White (2021).

would result from this. This is important because estimators of the tail exponent generally rely on order statistics, which are by definition very disclosive and would not be made available by statistical offices. We suggest these five quantiles are good candidates, and could be supplemented by more quantiles which can then be used for testing the quality of the fit.

4.2 Data and descriptive statistics

4.2.1 Data sources

In this section, we present the dataset and some basic patterns of the empirical data. We use data from the Orbis Europe database, compiled by Bureau van Dijk, which includes the balance sheets and profit-loss statements for 7 million unique firms across Europe, yielding approximately 23 million firm-year observations from 2006 to 2017. We keep all types of organisations, and use the words ‘firms’ throughout to refer to all kinds of organisations in our dataset.

Unlike other widely used firm-level data such as Compustat and Worldscope, Orbis Europe records a large number of small and medium sized firms that are often not publicly traded.⁵ Table 4.A.2 in Appendix 4.A reports the number of observations per country-year.

To avoid double counting, we only use unconsolidated data.⁶ We then remove duplicated firm-year pairs, and we exclude self-employed firms, where the distinction between wage and profit is ambiguous. Finally, we regard negative values for sales, wage, and employment as missing values. For a detailed discussion on data cleaning, see Appendix 4.A.

4.2.2 Construction of variables

For each firm i in industry j and year t , we define real value added ($Y_{i,t}$) as the sum of real labour income ($W_{i,t}$) and real capital income ($\Pi_{i,t}$), that is

$$Y_{i,t} \equiv W_{i,t} + \Pi_{i,t} = \frac{\omega_{i,t}}{p_{j,c,t}^v} + \frac{\pi_{i,t}}{p_{j,c,t}^v}, \quad (4.1)$$

where ω and π are nominal wage and nominal profit, and $p_{j,c,t}^v$ is the value added deflator of industry j in country c at time t . In Orbis, nominal wage and nominal profit variables are recorded in Cost of Employees (STAF) and Earnings Before Interest, Taxes, Depreciation & Amortization (EBITDA)⁷. The firms’ VA is deflated using the industry-level VA deflator from the EUKLEMS database (VA_P in Jäger & The Conference Board (2018)). See Appendix 4.A.3 for more details.

⁵Orbis Europe includes many different types of firms, including Private limited company, Joint-stock company, Partnership, Cooperative, Consortium, Foundation, and Public agency/corporation. For a more detailed discussion on the advantages and drawbacks of the Orbis Europe database, see Kalemli-Ozcan et al. (2015).

⁶For some corporate groups that consist of multiple firms, both consolidated and unconsolidated data is provided, i.e. data is listed for the entire group and again for each of the firms.

⁷VA can also be calculated by subtracting intermediate costs (MATE in Orbis) from gross output (TURN in Orbis). This is the output-based approach to computing VA. However, material costs are less frequently observed compared to wages and earnings in our sample. 27.67% of firms in our sample do not report material costs.

Denoting the number of employees by $L_{i,t}$ (EMPL in Orbis), our measure of firm-level labour productivity ($LP_{i,t}$) is defined as the ratio of value added to the number of employees,

$$LP_{i,t} = \frac{Y_{i,t}}{L_{i,t}}, \quad (4.2)$$

and is expressed in units of currency of the country where the firm is located.

Handling negative value added Researchers in productivity analysis typically proceed to a data transformation: they analyse the natural logarithm of productivity ([Bartelsman & Wolf 2018](#)). However, when output is measured as VA, firm-level labour productivity can be negative. If a firm's intermediate cost is greater than revenues in a given fiscal year, the firm has negative VA and productivity. From the income perspective of the definition of VA (Wages + Earnings), the sum of earnings and wages can be negative since firms' losses (negative profits) can be greater than wages.

Empirically, a sizeable share of firms has negative labour productivity. Table 4.2.1 shows the proportion (%) of negative observations per country-year in our data. Overall, 5% of the firms in our sample have negative productivity. As a more extreme case, more than a fifth of UK firms had negative productivity in 2006–2008⁸.

Table 4.2.1: Proportion (%) of negative observations per country-year

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Belgium	2.58	1.33	1.56	1.62	1.49	1.44	1.55	1.59	1.53	1.49	1.53	1.58
Bulgaria		9.33	4.02	5.79	5.86	7.45	7.60	7.02	6.35	6.10	5.47	5.13
Croatia	6.91	4.46	4.60	5.66	6.18	5.27	5.79	4.54	4.00	3.55	2.86	
Czech Republic	7.10	4.89	5.32	6.49	6.26	6.35	6.33	6.19	5.48	4.52	4.30	3.69
Denmark							2.51	2.98	2.83	2.51	2.91	3.04
Estonia		2.90	4.09	6.09	4.65	3.25	3.11	3.12	3.37	3.38	3.30	2.75
Finland	2.09	1.60	1.83	1.99	1.84	2.03	1.91	1.95	1.94	2.08	2.08	1.90
France	1.19	1.10	1.31	1.27	1.13	1.16	1.32	1.46	1.43	1.60	1.81	1.89
Germany	2.56	1.96	2.48	2.43	2.14	2.17	2.16	2.00	2.39	2.48	2.31	
Hungary		4.11	3.11	7.01	6.61	6.68	7.18	5.76	5.16	4.87	4.47	3.62
Italy	4.64	3.08	3.64	4.16	3.68	3.78	5.00	4.74	4.32	3.84	3.16	2.79
Poland	3.77	4.01	4.78	6.11	3.72	3.95	8.12	7.42				2.49
Portugal	8.62	6.06	6.51	6.12	5.81	6.87	8.34	7.87	7.86	6.47	5.57	4.80
Romania		7.58	8.52	10.73	11.73	11.46	12.31	12.47	10.40	7.34	6.39	
Slovakia		4.44	4.55	7.64	7.11	7.23	7.20	7.13	6.05	4.10	4.21	
Slovenia				1.50	1.69	1.81	1.88	1.56	1.44	1.35	1.25	1.15
Spain	3.78	2.92	3.60	3.90	3.65	4.04	4.41	4.05	3.33	2.94	2.57	2.07
Sweden	3.82	2.16	2.41	2.27	2.13	2.18	2.10	2.10	2.07	2.08	2.15	2.35
United Kingdom	20.98	21.72	23.49	11.31	4.14	4.04	3.84	3.52	3.57	3.73	3.92	

Notes: This table records the proportion (%) of negative observations of value added per country-year for all 19 countries.

⁸Other authors find smaller percentages for the UK, but there is systematically at least a few percents of negative VA firms. This is close to 10% in [Haldane \(2017, Chart 17\)](#). In employment weighted distributions, this is around 4–5% ([Office for National Statistics 2022](#)). [De Loecker et al. \(2022\)](#) find only 'a few' in Orbis, but remove many sectors and firms with less than 10 employees.

4.2.3 Characteristic patterns of the distributions

To motivate Lévy alpha-stable distributions as a model for LP , we first show qualitatively that the distributions are heavy-tailed and asymmetric.

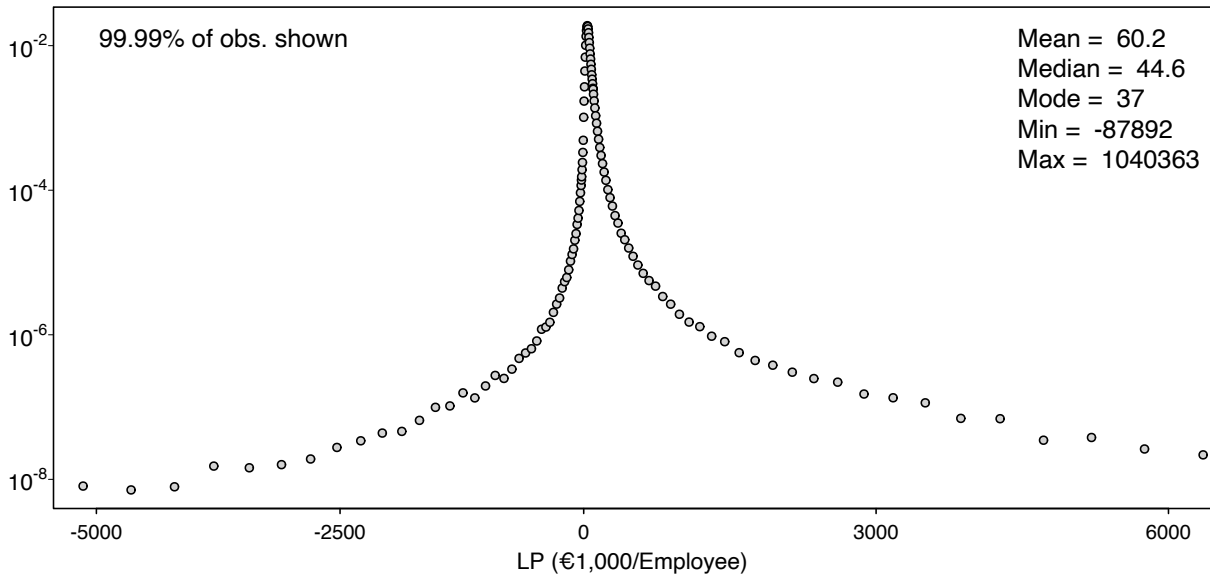


Figure 4.2.1: **Distribution of labour productivity in France, 2006–2017:** To display as many data points in the extreme tails as possible, we use logarithmic binning for the bottom and top 10% productivity. The plot shows 99.99% of the entire data. Units are in Euros, deflated to 2015.

Fig. 4.2.1 shows the distribution of labour productivity for France for the years covered by the data on a semi-log scale.⁹ An important observation is that even though the chart shows 99.99% of observations, the tails appear relatively ‘well behaved’. Once the bin’s mid-points are chosen appropriately, there are no large fluctuations and one could draw a fairly smooth line through the dots. This suggests that the common practice of winsorising (removing e.g. 1% of observations in each tail) removes observations that are actually well-behaved statistically, rather than being unexpected or strange ‘outliers’ (see e.g. [Nolan \(2020, p.196–197\)](#) for a discussion).

We note the following five general characteristics of the empirical distribution of labour productivity. First, the distribution appears unimodal. Second, the support of the distribution is very large. Third, the distribution is asymmetric, with a pronounced right skewness. Fourth, the distribution exhibits slowly decaying tails. And fifth, comparing the results for

⁹Visualising histograms for heavy-tailed data presents unique challenges. When all data are visualised, the expansive support can obscure meaningful patterns in high-density modal areas. Additionally, this can result in many missing bins in the tails. Simply truncating the tails in a visualisation is not viable, as doing so contradicts the very essence of studying data with extreme values. To strike a balance and effectively visualise both the tail and body behaviour, a dual binning approach is adopted. For the body of the data, specifically between the 10% and 90% quantiles, uniform binning is used to ensure equal-sized bins. However, for the bottom and top 10% of values, or the tails, we employ logarithmic binning. This method places data points in the tails into bins spaced at exponential distances, meaning the extreme values of larger magnitude in the tails are grouped together into larger bins. We then use the middle value of each bin to visualise the histogram.

different years (not shown here), we find that the shape of the distribution is very persistent, as one would expect.

We now focus on one of the most important of these features: heavy tails.

4.2.4 Testing for power law tails and infinite variance

We say that the tail of a distribution follows a power-law if its cumulative distribution function $F(x)$ takes the form

$$1 - F(x) \sim L(x)x^{-\alpha}, \quad x \rightarrow \infty, \quad (4.3)$$

where $L(x)$ is a slowly varying function, and α is the tail exponent. Intuitively, the smaller the tail exponent, the slower the frequency of a large event decreases as we consider increasingly extreme values. Very extreme values will be relatively more frequent than in a scenario with a higher tail exponent. In practice, it also determines what moments are finite; any moment greater than α will be infinite, while any moment less than α remains finite. Some authors use the term ‘heavy’ tails for the case where $\alpha < 2$, as in this case the variance is infinite (Nolan 2020, Resnick 2007).

Here we propose four methods to determine whether labour productivity distributions have heavy tails: (modified) QQ-plots, scaling of the sample standard deviation with sample size, direct estimates of the tail exponent based on extreme value distributions, and Trapani (2016)’s test for infinite moments.

QQ-plots Resnick (2007) shows that QQ (Quantile-Quantile) plots for location-scale families can be used not only to provide visual evidence on the compatibility of a specific functional form with the data, but also to estimate the parameters. This is convenient because if X is a random variable drawn from a Pareto distribution,

$$P(X > x) = \left(\frac{x}{k^*}\right)^{-\alpha}, \quad x > k^*, \quad (4.4)$$

where k^* is the threshold, and α is the tail exponent, the distribution of $\log X$ is a location-scale family. Resnick (2007) shows that the values

$$\left\{ \left(-\log \left(1 - \frac{i}{N+1} \right); \log X_{(i)} \right), 1 \leq i \leq N \right\}, \quad (4.5)$$

where $X_{(i)}$ is the i^{th} order statistic, would lie on a line with intercept $\log k$ and slope $1/\alpha$ under the null (Eq. 4.4). To implement this procedure, we need to choose a cutoff k^* , or, alternatively, the sample size N . If we believe that the data has a Pareto tail, but only as $x \rightarrow \infty$, we would prefer to choose the highest possible cutoff. However, this implies less data and thus more variance in the resulting estimates, so there is a trade-off.

Fig. 4.2.2 shows the log of top and bottom 0.5% observations of firm labour productivities. The datapoints by country roughly follow straight lines, which is good evidence that the tails are Pareto. Moreover, the slopes appear to lie between 1 and 1/2, suggesting $1 < \alpha < 2$.

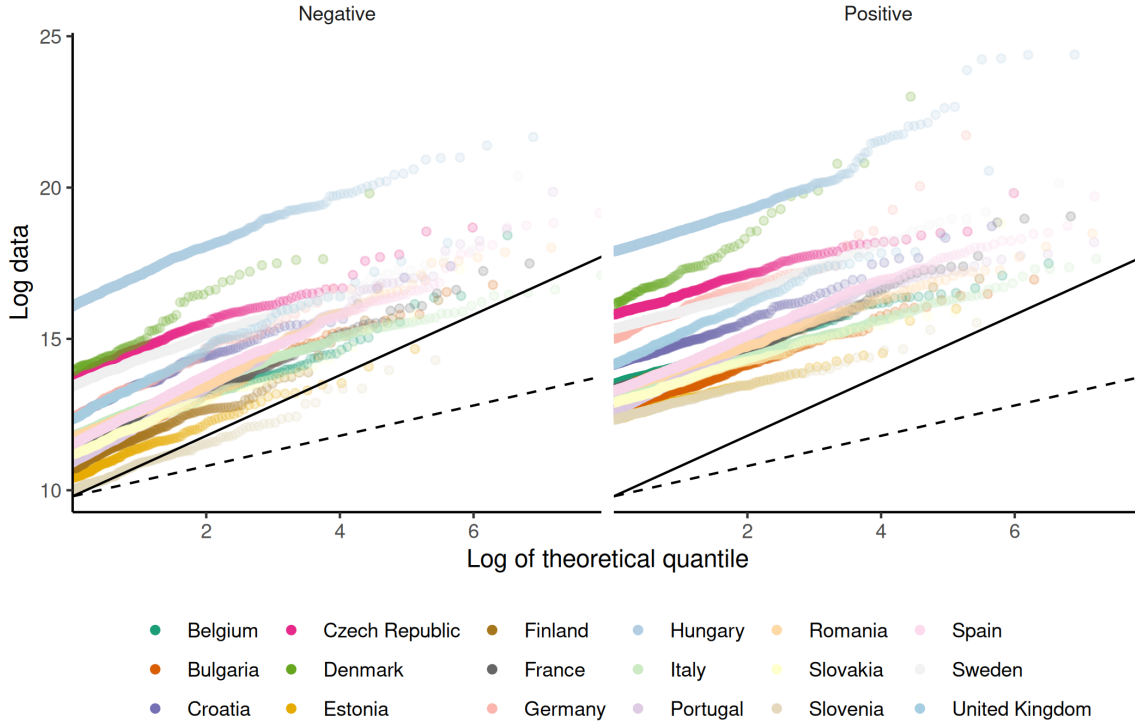


Figure 4.2.2: **QQ-plots for positive and negative labour productivity tails, 2015:** Quantiles of the labour productivity for the top and bottom 0.5% of firms in 18 European countries are very close to the quantiles of a Pareto distribution. A Pareto tail with an exponent α would appear as a line with slope $1/\alpha$ (Resnick 2007). A line with a slope of one (solid black), and a line with a slope of $1/2$ (dashed) are plotted for visual comparison.

One point to note is that the slopes for negative productivity appear slightly steeper than for positive productivity, suggesting a degree of tail asymmetry that is not in line with the Lévy hypothesis, where the tail exponent is the same on each side (Section 4.B.1).

Scaling of the sample standard deviation with sample size If the distribution has power-law tail with $\alpha < 2$, the variance and higher-order moments are infinite. Due to this property, in finite samples, the larger the sample size, the higher the chance that an extreme event is drawn, leading to a larger sample standard deviation. More precisely, the sample standard deviation scales with sample size N as $N^{\frac{1}{\alpha}-\frac{1}{2}}$ (see Appendix 4.E for a heuristic derivation).

Figure 4.2.3 shows the scaling of the sample standard deviation of labour productivity for different sample sizes. It is clear that for most countries, the sample standard deviation gets larger as the sample size increases. For four selected countries, we show the theoretical scaling derived by estimating the power law exponent as the tail parameter of a Lévy alpha-stable distribution. The scaling of the sample standard deviation holds well empirically, although with a possible plateauing as $n \rightarrow \infty$ for some countries, such as Spain. Overall, this provides good evidence not only for power law tails, but also for the idea that the Lévy alpha-stable distribution can be a good model to retrieve the value of the tail index.

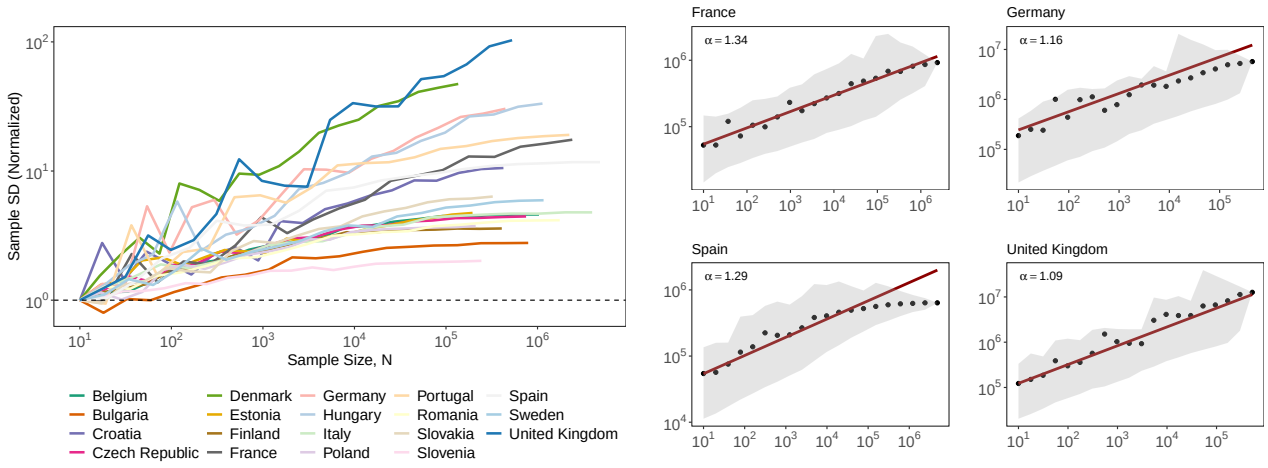


Figure 4.2.3: Measured standard deviation in sub-samples of firm labour productivity, country samples: We construct a distribution by pooling together the productivity levels of firms in each of 19 countries, for all years. Then, for each sub-sample size N , we compute the standard deviation of each of 1,000 sub-samples. The plot on the left shows the average of standard deviations for each of all 19 countries, while the plot on the right shows the same scaling for four selected countries, with the linear scaling calculated as $N^{\frac{1}{\hat{\alpha}} - \frac{1}{2}}$ (see Appendix 4.E), where $\hat{\alpha}$ is estimated by fitting the Lévy distribution. The shaded area corresponds to the values of the sample standard deviation that fall between the 5th and 95th percentiles.

Estimating tail exponents Another way to test for infinite second moment is to estimate the tail exponent and see if it is lower than 2. Typically, one considers data from the tail only; that is, order statistics of up to order k , which makes it possible to estimate tail behaviour and determine the finiteness of moments, independently of the behaviour of the rest of the distribution. An important issue is that one has to choose a value k^* , which determines which data is used to estimate the tail parameter. Usually, the tails are influenced in a non-negligible way by the slowly varying function $L(x)$ (see Eq. 4.3), and therefore the choice of k^* may be difficult and lead to biased estimates of α .

An early and popular method for estimating tail exponents is the Hill estimator, but as noted in Resnick (2007), the Hill estimator provides very different estimates of the tail depending on which value of k^* we choose (‘Hill Horror Plots’). Because of these well-known issues for regularly varying distributions (Eq. 4.3) that are somewhat far from pure power laws, several estimators of the tail exponents have been developed, and tested for cases where the slowly varying function is non-negligible. Here we use the estimators described and implemented by Voitalov et al. (2019). Voitalov et al. (2019)’s implementation includes an automatic double bootstrapping procedure for picking k^* . We use their package with default values. In addition, we also use the popular Hill estimator described in Clauset et al. (2009), which finds k^* by minimising the Kolmogorov-Smirnov statistic (computed assuming that the true model is a *pure* power law). We add the estimate of α based on the Lévy distribution as well.

Figure 4.2.4 (left) shows the results using data from France for each year separately, and strongly suggests that $\alpha < 2$. Three out of 5 estimators suggest similar values of $\alpha \approx 1.4$

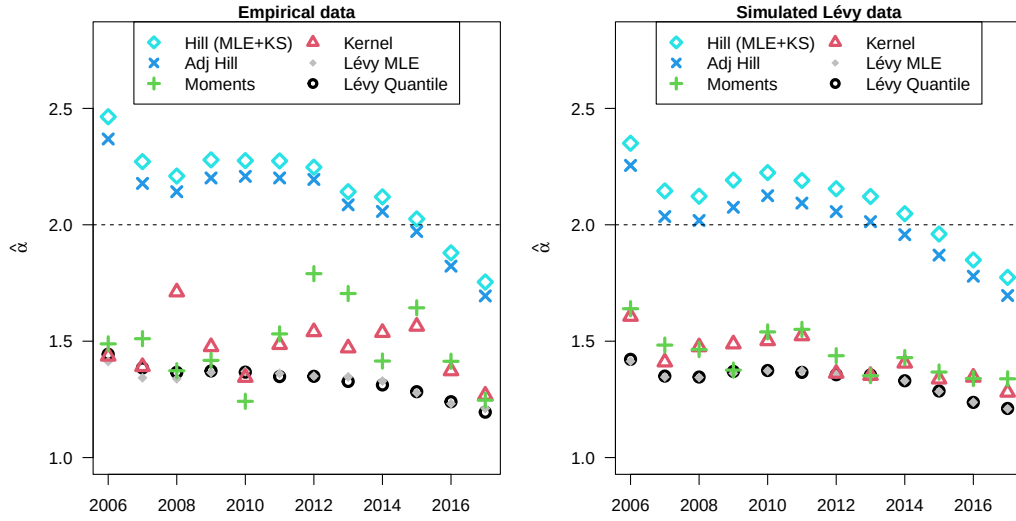


Figure 4.2.4: **Estimated values of the tail exponent $\hat{\alpha}$** : Using pooled data from France for each year separately, and using 6 different estimators: The Hill estimator from [Clauset et al. \(2009\)](#) ('Hill (MLE+KS)'), three estimators implemented by [Voitalov et al. \(2019\)](#) ('Adj Hill', 'Moments' and 'Kernel'), and estimates of α from the Lévy distribution, either using Maximum Likelihood (our preferred approach, see Appendix 4.C) or the quantile estimator from [McCulloch \(1986\)](#)). The left chart shows $\hat{\alpha}$ for the empirical data. The right chart shows $\hat{\alpha}$ for data drawn from a Lévy-stable distribution with parameters equal to those estimated on the empirical data using the MLE method. The left chart shows that 4 estimators suggest infinite variance ($\hat{\alpha} < 2$), while the two Hill-based estimators sometimes suggest finite variance on empirical data, but would also fail to capture infinite variance if the data were indeed Lévy-distributed (right). We conclude that labour productivity has indeed a power law tail with $\alpha < 2$.

(dots for 2010 are missing for two estimators as the routines sometimes fail). The two estimators that suggest $\alpha > 2$ for the earlier years are Hill-based estimators, which are known to produce poor results for Lévy distributed data ([Resnick 2007](#)). To further confirm the plausibility of these estimates, for each year, we simulate one sample of Lévy distributed data using the empirical sample size, and using parameters for the Lévy distribution estimated from the empirical data. In this case (right panel), the 5 estimators behave almost exactly as in the empirical data, suggesting that the Lévy distribution is a plausible model, and reinforcing the previous conclusion that the second moment is infinite.

The left panel of Figure 4.2.4 also helps to make an important point: if the data is *not* Lévy distributed, but is a regularly varying distribution with a tail exponent that is correctly estimated by the Kernel or Moments estimators, then estimating the exponent using simply 5 quantiles and assuming a Lévy distribution gives relatively good results in the case of our data. This is important, because statistical offices holding detailed micro-data may not want to release the order statistics that are necessary to run the Kernel or Moments estimators. They may, however, be willing to release 5 quantiles, since the lowest/highest quantiles (5th and 95th) are likely to be uninformative about the situation of specific firms (in contrast to, say, the 99.9th quantile).

Trapani's test As a final check, we implemented a more formal statistical test due to [Trapani \(2016\)](#). The testing procedure exploits the fact that non-finite sample moments diverge with sample size.

[Trapani \(2016\)](#) suggests a test for the divergence of arbitrary moments of order p , including fractional (non-integer) moments. As the moment of order p may be infinite, it is unknown whether a limiting distribution of the moments exists. The test can therefore not be applied directly. Following a randomised testing approach, the test therefore adds artificial randomness to manipulate the quantity in question, the sample moment of order p , to yield a known distribution if the moment is infinite. The resulting test statistic can then be compared to this distribution to obtain a p-value for whether or not the null hypothesis, that the moment of order p is infinite, is correct.

Test procedure The test statistic is derived such that under the null hypothesis H_0 , the p th moment of the distribution is infinite. For this, the approach starts with the absolute sample moment of order p ,

$$A_p = \frac{1}{n} \sum_{k=1}^n |X_k|^p.$$

To make the resulting test statistic scale-invariant and comparable, the absolute moment must be rescaled,

$$A_p^* = \frac{A_p}{(A_\psi)^{p/\psi}} \times \frac{(A_\psi^\mathcal{N})^{p/\psi}}{A_p^\mathcal{N}},$$

with $\psi \in (0, p)$. $A_p^\mathcal{N}$ denotes the p th absolute fractional lower order moment of the standard normal distribution, $A_\psi^\mathcal{N}$ the ψ th absolute fractional lower order moment of the standard normal, etc. Next, an artificial random sample ξ of size r ¹⁰ is generated from a standard normal distribution and rescaled,

$$\varphi_r = \sqrt{e^{A_p^*}} \times \xi_r.$$

The intuition here is that φ_r follows a normal distribution with mean zero and a finite variance, as $n \rightarrow \infty$, if A_p^* is finite itself. Thus the problem has been reduced from testing for any moment p , to testing the existence of the variance of the transformed random variable φ_r .

The next step is to generate a sequence ζ_r , given by

$$\zeta_r(u) \equiv I[\varphi_r \leq u],$$

where $I[\cdot]$ is an indicator function, and $u \neq 0$ is any real number. Under H_0 , $\zeta_r(u)$ will have a Bernoulli distribution with mean $\frac{1}{2}$ and variance $\frac{1}{4}$. This is not the case under the alternative, where $A_p^* < \infty$, as $e^{A_p^*}$ converges to a finite value.

¹⁰Trapani suggests $r = n^{0.8}$ where n is the number of observations used to compute A_p .

Values for u are picked from some density, but for simplicity it can be taken from a uniform distribution $U(-1,1)$. As a result, the test statistic of interest is obtained as

$$\Theta_r \equiv \int_{-1}^1 \frac{1}{2} \vartheta_r^2(u) du,$$

$$\vartheta_r(u) \equiv \frac{2}{\sqrt{r}} \sum_{j=1}^r \left[\zeta_j(u) - \frac{1}{2} \right].$$

Under the null hypothesis that moment p is infinite, $\vartheta_r(u)$ should reduce to zero given that $\zeta_j(u)$ has a mean of one half. Θ_r is thus shown to follow a χ^2 distribution with $df = 1$ if moment p is infinite.

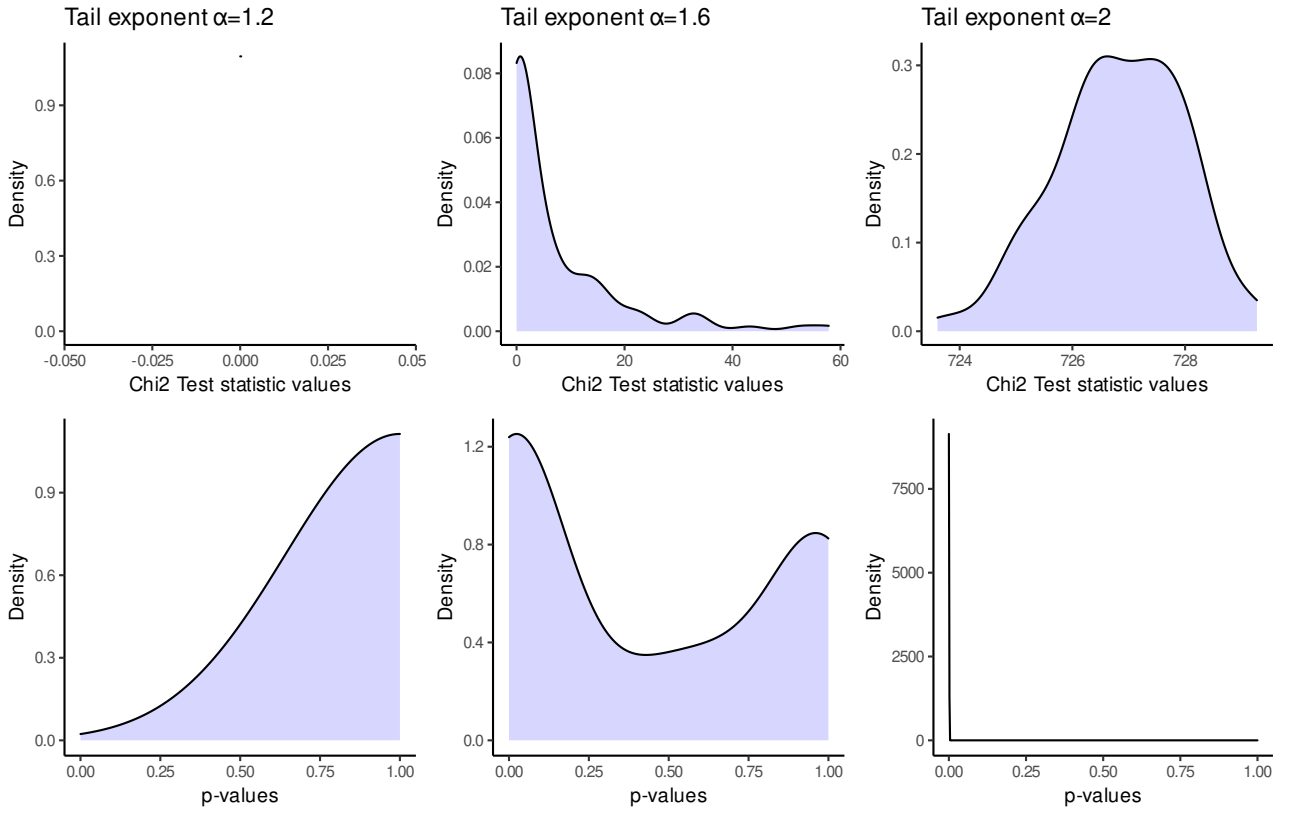


Figure 4.2.5: **Performance test for Trapani's finite moment test conducted using the R package finity:** Distribution of χ^2 values and p-values given by the test for the divergence of the second moment ($p = 2$) for samples of size $N = 100,000$ drawn from a Lévy alpha-stable distribution $S(\alpha, 0.5, 1.0, 0.0; 0)$ for three different values of $\alpha \in \{1.2, 1.6, 2.0\}$.

Performance of the test Trapani's test depends on a number of parameters; the choice of ψ , r , and u can influence the test results. We largely follow [Trapani \(2016\)](#) in our choices of the parameter values and choose

$$u = 1, r = n^{0.8}, \psi = \begin{cases} p - 1 & \text{if } p > 1 \\ p/2 & \text{if } p \leq 1 \end{cases}.$$

The test results should be expected to be somewhat noisy. Especially for samples drawn from a Lévy alpha-stable distribution $S(\alpha, \dots)$ in the vicinity of the true tail parameter $p \approx \alpha$, the test loses accuracy.

For the tests performed here, we use the R package `finty` (Heinrich & Winkler 2020). Fig. 4.2.5 shows the distribution of χ^2 statistics and p-values for a test of divergence of the second moment ($p = 2$), for samples of size $N = 100,000$ drawn from a Lévy alpha-stable distribution $S(\alpha, 0.5, 1.0, 0.0; 0)$ for three different values of α . The results in Fig. 4.2.5 reflect a typical pattern of performance of our implementation of the test at relatively large sample sizes ($N > 1000$), where the test tends to seriously over-reject the null even for values of $2 > \alpha > 1.6$, reinforcing the robustness of our conclusion that productivity has an infinite variance.

Table 4.2.2: Testing for infinite moments of firm-level labour productivity

Country	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Belgium	1.00	0.30	0.72	0.00	0.00	0.00	0.27	0.08	0.95	0.98	0.99	0.98
Bulgaria		0.98	0.86	0.91	1.00	1.00	1.00	0.99	0.93	0.99	0.92	1.00
Croatia	1.00	1.00	1.00	1.00	0.99	0.99	0.98	0.99	0.92	0.00	0.32	
Czech Republic	0.99	0.97	0.99	1.00	1.00	0.99	0.83	1.00	0.65	0.52	1.00	0.99
Denmark							1.00	1.00	1.00	–	0.99	0.99
Estonia		1.00	0.03	0.01	1.00	0.42	0.00	0.09	0.00	0.00	0.00	0.00
Finland	0.10	0.08	0.63	0.08	0.28	0.59	0.63	0.82	0.92	0.62	0.76	1.00
France	0.08	0.32	1.00	0.99	1.00	1.00	0.99	0.99	0.24	0.99	0.99	1.00
Germany	1.00	0.99	1.00	1.00	1.00	0.99	0.99	1.00	–	–	1.00	
Hungary		1.00	0.98	1.00	0.99	0.99	–	–	0.99	0.99	1.00	0.99
Italy	0.99	0.44	0.93	0.99	0.67	0.51	1.00	0.00	0.00	0.00	0.87	0.00
Poland	1.00	1.00	0.59	0.99	0.99	0.98	0.94	1.00				0.96
Portugal	0.95	0.95	1.00	1.00	0.99	0.99	1.00	1.00	1.00	1.00	0.99	0.99
Romania		1.00	0.99	0.99	0.90	0.35	1.00	1.00	0.74	0.99	1.00	
Slovakia		0.98	0.99	1.00	0.16	0.37	1.00	0.43	0.78	0.00	0.86	
Slovenia					0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00
Spain	0.28	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Sweden	0.00	0.95	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.93	0.00	0.00
United Kingdom	0.98	0.99	0.99	0.99	0.99	0.99	–	0.99	–	1.00	0.99	

Notes: P-values of Trapani's finite moment test for the nonexistence of the second moment of the distribution of labour productivity, for country-year samples. In the vast majority of cases, the p.value is above any sensible threshold (e.g., 1%, 5%, or 10%), suggesting that the nonexistence of the second moment can not be rejected. For example, 84% of the P-values are greater than 0.05, meaning that, for 84% of country-year samples, the finite moment test failed to reject the null hypothesis of the infinite second moment at the 5% significance level.

Finite moment test results for country-year sample In the vast majority of both the country-year and country-industry samples, we cannot reject the hypothesis that the second moment of labour productivity is infinite. For example, for 84% of country-year samples, Trapani's finite moment test failed to reject the null hypothesis of the infinite second moment at the 5% significance level.

4.3 Models and estimation

This section provides a detailed discussion of the key properties of the Lévy alpha-stable distribution, and introduces the estimation methods. For details on the competing distributional models (the 5-parameter Asymmetric Student-t (AST) distribution and the 5-parameter Asymmetric Exponential Power (AEP)), see Appendix 4.D and Bottazzi & Secchi (2011).

4.3.1 Models

The Lévy alpha-stable distribution Here we briefly describe the key properties of the Lévy alpha-stable distribution (see Appendix 4.B for details).

The Lévy alpha-stable distribution is a four-parameter distribution with the following parameters: The parameter $\alpha \in (0, 2]$ is also called the tail exponent because the tail of the distribution decays as a power law with exponent α , that is, $P(X > x) \approx Cx^{-\alpha}$. The lower α , the thicker the tail (the more ‘tail dispersion’ there is). However, note that while a smaller α indicates more dispersion because of a heavier tail, a smaller α also leads to a higher concentration of values close to the mode, and thus, in this sense, to a lower dispersion in the body of the distribution.

The ‘skew’ or ‘asymmetry’ parameter $\beta \in [-1, 1]$ is such that for $\beta = 0$, the distribution is symmetric, for $\beta > 0$ right-skewed, and for $\beta < 0$ it is left-skewed. In practice, we will find that productivity distributions are right skewed.

The higher the scale parameter γ , the wider the body of the distribution. The parameter γ is not dimensionless; it is expressed in the same units as the data.

Finally, the location parameter $\delta \in (-\infty, +\infty)$ shifts the distribution, even though it is not in general exactly related to key quantities such as the mode, mean, or median.

The Lévy distribution arises from the Generalized Central Limit Theorem (GCLT), which is the appropriate generalisation of the classical Central Limit Theorem (CLT) when the distribution of the variables being summed up does not necessarily have a finite variance. Like for the CLT, there are extensions showing conditions where the GCLT still holds for non i.i.d. variables, although these results are more complicated.

Estimation method To estimate the parameters of the Lévy alpha-stable distribution, we use the maximum likelihood estimation method based on the Limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS-B) algorithm (Byrd et al. 1995). Despite its superior performance for the Lévy parameter estimation (Nolan 2020, Chap. 4), the MLE is computationally highly prohibitive, as the density function of the Lévy alpha-stable distribution does not have a closed form and thus its evaluation takes long. For this reason, when using the MLE, we resort to using a sub-sample of the data. As long as the size of the sub-sample is large enough ($\geq 30,000$), the estimation results remain robust, irrespective of the specific sub-samples used in the MLE estimation. See Appendix 4.C for a detailed discussion of the estimation method.

Benchmarks There are two benchmark models used in this paper: the 5-parameter Asymmetric Student-t (AST) distribution (Zhu & Galbraith 2010) and the 5-parameter Asymmetric Exponential Power (AEP) distribution (Bottazzi & Secchi 2011). Similar to the Lévy alpha-stable distribution, these distributions generalise the Gaussian distribution to allow for skewness and more probability density in the tails. The AST distribution provides a particularly good benchmark as it has the power law tail similar to the Lévy alpha-stable distribution, but it allows for asymmetry in the tails with two distinct tail parameters (ν_1 and ν_2). Therefore, it can help us examine potentially different patterns in each of the left and right tails and determine whether accounting for asymmetry results in an overall better fit for the data. In contrast, the AEP distribution does not feature power law tails and can serve as an effective benchmark to demonstrate how crucial accounting for power-tails is in understanding the distribution of firm-level labour productivity.

Both of these distributions have been suggested as models for quantities that are known to have heavier tails than Gaussians, such as return volatility (Zhu & Galbraith 2011), changes in currency exchange rates (Ayebo & Kozubowski 2003), and (logarithmic) firm growth rates (Bottazzi et al. 2007). Appendix 4.D gives the functional forms of the 5-parameter AST and 5-parameter AEP models, the details of the fitting procedure, and goodness comparisons to the Lévy alpha-stable model fit.

4.3.2 Distribution fit and model comparison

Fig. 4.3.1 shows the empirical densities and the fits for the Lévy alpha-stable, AST and AEP distributions¹¹ for the France sample.¹² The Lévy alpha-stable distribution is able to fit the entire domain of the empirical distribution extremely well,¹³ possibly with small shortcomings in dealing with asymmetric tail behaviour (as noted in Section 4.2.4).

The AST is broadly able to fit the body of the empirical distribution but in the case of France is less effective in capturing the tail part of the distribution. The AEP distribution not only fails to fit the tails, but also offers an unsatisfactory fit for the distribution’s body¹⁴.

As an additional validation, Fig. 4.3.2 shows the scaling of the estimated standard deviation in randomly generated samples from the Lévy alpha-stable, AST, and AEP models, respectively. This comparison uses a simulated sample of 2.4 million observations, equivalent to the size of our empirical dataset for France. The distributions were generated using the parameters estimated from the respective model based on our empirical data for France. The Lévy distribution random sample exhibits a power law scaling with the standard deviation.

¹¹We use the R-package **SkewtDist** package (Xia 2022) to estimate the AST and Bottazzi (2014)’s **Subbotools** package for the interval-constrained likelihood optimisation of the 5-parameter AEP.

¹²We selected France as the most representative case given its significant country sample size, alongside a firm size composition that closely aligns with the sample mean. For a detailed overview of the number of observations and size composition by country, please refer to Table 4.A.3.

¹³There is a mild discontinuity around zero, as is shown by a slight peak in the histogram between zero and approximately 10,000 Euros. Upon thorough examination, this discontinuity is observed in several country samples. We were not able to find a clear cause for this. For the France sample, the distortion around zero is predominantly due to the medium-sized firms even though small and large firms do have some degree of distortion around zero.

¹⁴See Moran et al. (2020) for a discussion of the inability of the Laplace distribution to fit growth rates due to the sharp change of behaviour around the mode.

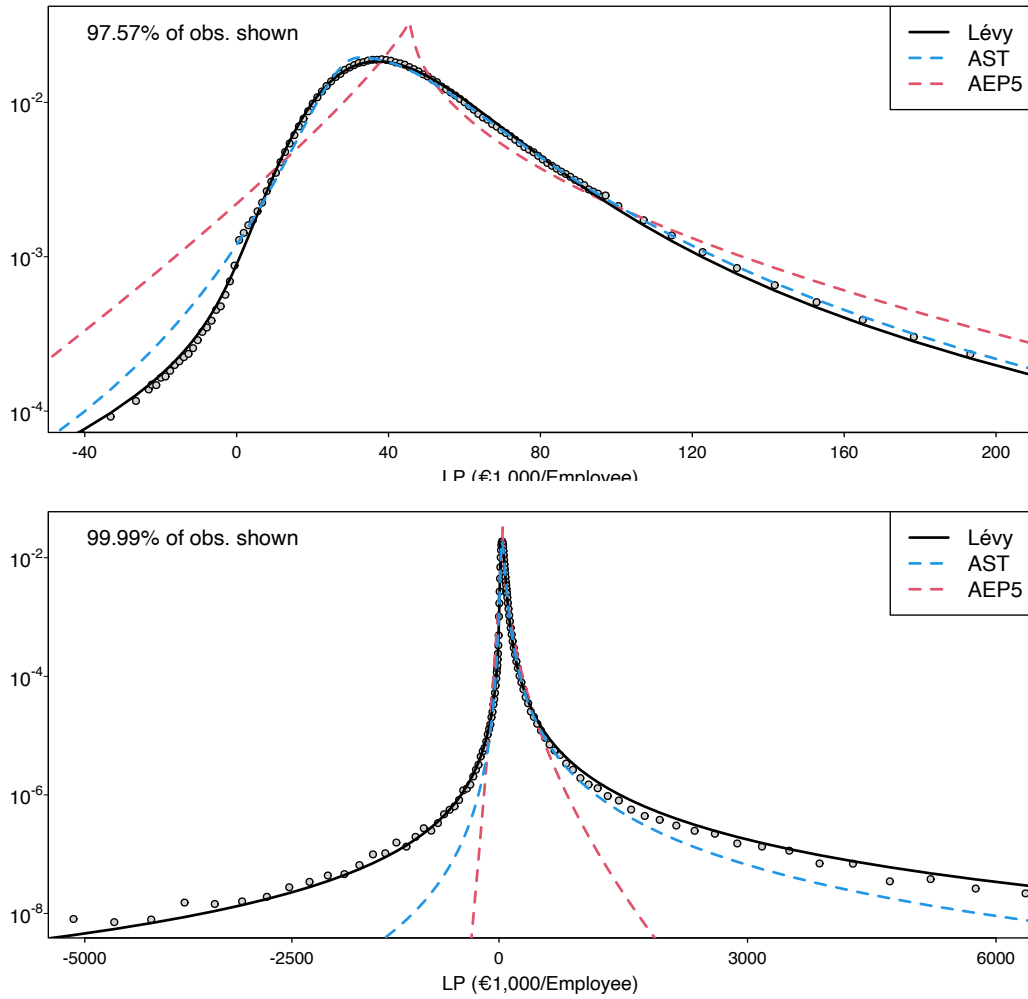


Figure 4.3.1: **Distribution of labour productivity in France with Lévy alpha-stable and AST and AEP fits:** The solid black line indicates the Lévy alpha-stable, the dashed blue line indicates the 5-parameter AST fit and the dashed red line indicates the 5-parameter AEP fit.

tion, and aligns closely with the empirical data. By contrast, while the AST random sample does show scaling with the standard deviation, it does not match well with the empirical data.¹⁵ Predictably, an AEP random sample fails to reproduce the pattern of the scaling due to its lack of power law tails.

4.4 Measuring productivity dispersion

To sum up from the earlier discussion, we cannot measure productivity dispersion using the empirical standard deviation because it is likely to be a measurement of a moment that is infinite, so its measured value depends on sample size and is not meaningful. We also cannot measure dispersion as the standard deviation of the log of VA per worker, because too many values are negative, and would need to be removed when taking the log. As a result, we are left with only two serious families of contenders: non-parametric metrics

¹⁵The same result applies to most of the 19 countries in our sample. Notable exceptions are Romania and Slovenia, where the AST is a better fit than the Lévy.

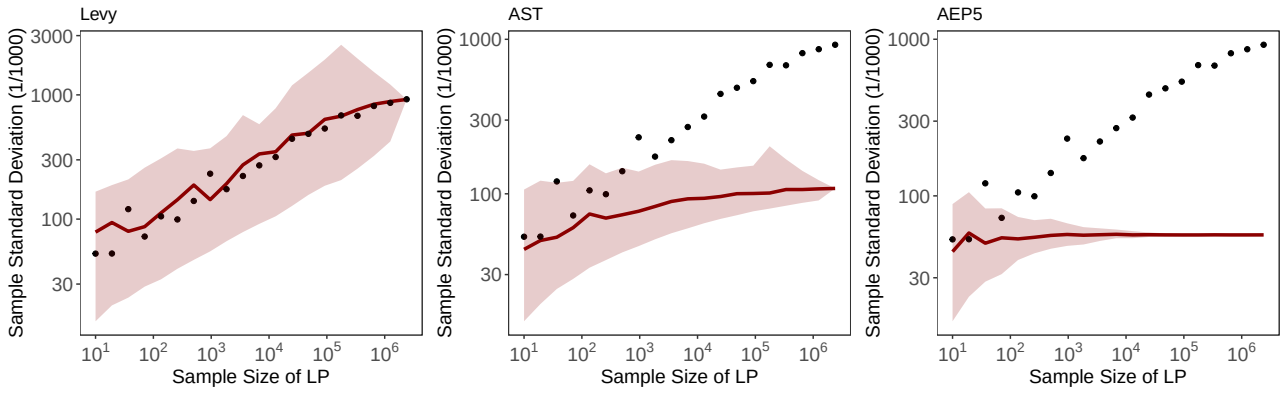


Figure 4.3.2: **Simulation of the scaling of the sample standard deviation with sample size for Lévy alpha-stable, AST and AEP distributed data:** We generate surrogate datasets that follow either the Lévy alpha-stable (left), the 5-parameter AST (middle), and the 5-parameter AEP distribution (right), using the same sample size as the empirical data for France (2.4 million), and the parameters estimated on the empirical data. We then estimate the standard deviation in sub-samples of the surrogate datasets, using many sub-samples for each sub-sample size. The figures show mean and 5%-95% quantiles for estimates of the standard deviation. The black dots show the mean (across sub-samples of a given size) of the sample standard deviation in the empirical dataset.

based on quantiles, and parametric metrics based on distributions with heavy tails and a support on both negative and positive values.

We first discuss our choices of indicators of dispersion in the body of the distribution, and our indicators for tail fatness. Of course, these choices are somewhat arbitrary, as one can choose different quantiles, and sometimes different parametrisations for a given distribution. Table 4.4.1 provides an overview of all our dispersion metrics.

Table 4.4.1: Summary of dispersion metrics

	Non param.	LogN	Lévy	AST
Location	Q_{50}	μ_{LN}	δ	μ_{AST}
Abs. Scale	$Q_{75} - Q_{25}$	σ_{LN}	γ	σ_{AST}
Rel. Scale	$(Q_{75} - Q_{25})/Q_{50}$ or Q_{90}/Q_{10}	σ_{LN}/μ_{LN}	γ/δ	σ_{AST}/μ_{AST}
Right tail	$(Q_{95} - Q_{50})/Q_{50}$	Q_{95}/Q_{50}	α	ν_2
Left tail	$(Q_{50} - Q_{05})/Q_{50}$	Q_{05}/Q_{50}	α	ν_1
Skewness	$\frac{(Q_{90}-Q_{50})-(Q_{50}-Q_{10})}{Q_{90}-Q_{10}}$	-	β	α_{AST}
Comments	Q_{10} can be < 0	Excludes non-positive values		

Notes: This table compares common dispersion measure, in the first two columns, against parametric measures from our paper (Lévy) and the Asymmetric Student-T (AST).

‘Body’ dispersion: Absolute and relative measures of scale Let us start by discussing Inter-Quantile Ratios (IQRs), which are very popular in part because they are easy to inter-

pret. Let us denote the value of the p^{th} quantile for VA per worker (V) by $Q_p(V)$. Considering for instance firms at the 90th and 10th quantiles, we define the IQR as the ratio

$$IQR_{90/10} = \frac{Q_{90}(V)}{Q_{10}(V)}, \quad (4.6)$$

which tells us how many times more productive the firm at the 90th percentile is compared to the firm at the 10th percentile. Usually, the variables are initially log transformed. If all values are positive, the firm sitting at the p^{th} quantile of VA per worker is also the firm sitting at the p^{th} quantile of the $\log(\text{VA per worker})$, so we have $\log(Q_p(V)) = Q_p(\log(V))$, and therefore Eq. 4.6 can be rewritten

$$IQR_{90/10} = \exp\left(Q_{90}(\log V) - Q_{10}(\log V)\right).$$

In other words, the Inter-Quantile Ratio is the exponential of the Inter-Quantile *range* of the log-transformed values. While one usually log transforms the data before computing an Inter-Quantile range, this is by no means necessary, because one could just directly compute an Inter-Quantile Ratio.

This makes a key advantage of IQRs clear: even when there are negative values, IQRs can still be computed, as long as the bottom quantile is positive. For example, Spain has roughly 3–4% of negative labour productivity observations (Table 4.2.1), so we do not need to drop the negative values if we want to compare, say, the bottom 5% to the top 5%. Unfortunately, the practice of automatically taking log as a first step of the analysis has sometimes led to removing non-positive values before computing IQRs, leading to a bias in the measurement of dispersion.¹⁶

An issue with Inter-Quantile Ratios is that they do not make sense when the lower quantile is negative, which can happen relatively frequently. In our sample of 206 country-year observations, this happens 10 times, almost 5%. An alternative is to compute the Inter-Quantile Range $IQR_{90-10} \equiv Q_{90}(V) - Q_{10}(V)$, which faces no particular issues when $Q_{10}(V) < 0$. This is an ‘absolute’ measure of dispersion, expressed in the same units as the original data. A value of around 70,000 for the early years in France, for instance, means that the firm at the 90th quantile has a value added per worker that is €70,000 higher than the firm at the 10th quantile. This can be compared with the parametric absolute measures of scale, namely γ in the Lévy distribution, σ_{AST} in the AST, and σ_{LN} , the standard deviation of the log values (removing non-positive observations). In fact, when α is small and β is large, γ is very close to half the interquartile range (Nolan (2020), see Appendix 4.B). To facilitate the comparison between quantile-based and Lévy, we choose IQR_{75-25} as a measure of absolute dispersion.

Since $IQR_{90/10}$ is a very popular and intuitive way to measure dispersion, however, we

¹⁶If one wants to take the log but keep all the observations (to keep the ranks, at least for the positive values (Campbell et al. 2019)), one can code the log of negative values as some value strictly lower than the lowest value of the log values of the positive values available in the sample. Note that coding log labour productivity as 0 would be incorrect as there could be values of VA per worker that are between 0 and 1 (particularly if expressed in thousands or millions of currency units), leading to log values that are negative.

would like to devise similar *relative* measures of dispersion from the parameters of the Lévy distribution. A natural candidate is simply to divide by the estimated location parameter $\hat{\gamma}/\hat{\delta}$, since δ also depends on the units of the original data. We can devise a similar metric for the AST and for the standard deviation of the log-transformed data. Note, however, that the non-parametric ‘relative’ scale dispersion $\text{IQR}_{90/10}$ (Eq. 4.6) takes the 10th percentile as the basis, while the three parametric indicators above take a location parameter, presumably much closer to the mean or median than to a lower quantile such as Q_{10} . This suggests that, to make better comparisons, a better non-parametric measure of relative scale dispersion would be taking an absolute metric of dispersion and dividing by a measure of central tendency. Thus, we choose $(Q_{75} - Q_{25})/Q_{50}$.

‘Tail’ dispersion A key advantage of the Lévy model is that it provides an easy way to interpret the parameter for the tail, the tail exponent α . The AST is even better in this respect, as it allows a separate estimate for the left tail, ν_1 , and the right tail, ν_2 . We consider that a fatter tail (α closer to zero) indicates more dispersion, since fatter tails indicate that there is a higher mass of super-productive firms. In almost all cases, the skewness parameter β is substantially above 0, indicating right-skewness. So, as we will confirm, α tends to be an indicator of the fatness of the *right* tail. We often use $-\alpha, -\nu_1, -\nu_2$ so that these indicators, like all the others, indicate more dispersion when they increase.

To compare these indicators of tail dispersion to non-parametric metrics, we follow the literature and focus on the 95th quantile (e.g. De Loecker et al. (2022), Andrews et al. (2019)), which we express as relative distance to the median, $\text{IQR}_{95/50} = Q_{95}/Q_{50}$. This measures the extent to which superstar firms deviate from a ‘typical’ firm¹⁷. Our choice is driven by previous work that has identified an increase in productivity dispersion mostly due to the divergence of superstar firms (Andrews et al. 2019, De Loecker et al. 2022). By symmetry, we could use Q_{05}/Q_{50} for the left tail (Oliveira-Cunha et al. (2021) use Q_{10}/Q_{50}), but Q_{05} is often negative during crisis years, so we use $(Q_{50} - Q_{05})/Q_{50}$.

4.5 Discussion and conclusion

The distributions of firm-level VA per worker have extremely large support, are asymmetric, and are heavy-tailed. A major consequence of this is that measuring dispersion is not straightforward: standard deviations are poor metrics because second moments do not exist, and log transformations are not recommended due to the proportion of the negative values.

We propose the Lévy alpha-stable distribution as a sensible distributional model for labour productivity, motivated by empirical evidence of an infinite variance, and by the fact that the Lévy distribution should be fairly common, as it emerges from the Generalized

¹⁷Berlingieri, Blanchenay & Criscuolo (2017) and Oliveira-Cunha et al. (2021) look at the evolution of $\text{IQR}_{90/50}$ as a measure of ‘divergence of the upper tail’. As discussed in Section 4.2.4, tail estimators determine a threshold value below which the data is not used for estimation. Typically they make use of only a few percent of top observations. For instance, the Moment and Kernel estimators reported in Fig. 4.2.4 would typically use between 0.1 and 2.5% of observations, but sometimes more (the precise numbers depend on the random seed, as the double bootstrap procedure to find an optimal threshold is stochastic).

Central Limit Theorem. We show that the Lévy distribution provides a better fit to the data than the five-parameter AEP distribution, and at least as good a fit as the less parsimonious AST distribution.

Good distributional models make it possible to offer a richer picture of dispersion. While the scale parameter captures the overall width of the distribution, the tail parameter captures the occurrence of the extreme events. These are qualitatively distinct aspects of dispersion. While existing research does attempt to distinguish between body and tail dispersion using quantile ratios, we argue that parametric measures make this distinction clearer and more objective, as there is no need to log-transform the data or to choose specific quantiles.

To illustrate this point, we have presented time series patterns in the UK, where our sample has a dramatic number of negative VA firms, making lower quantiles negative in crisis years. We have found that the Lévy parameters are useful to describe the strong time series patterns: an increase in skewness, a decline in body dispersion, and an increase in tail fatness.

To go further, the research on productivity dispersion needs administrative data, which typically cannot be made available. Our findings have implications for statistical offices releasing moments of micro-data. Because the Lévy alpha-stable distribution can be estimated using 5 quantiles (more conveniently the 5th, 25th, 50th, 75th and 95th), if statistical offices release more than these 5 quantiles, then the Lévy alpha-stable can be estimated on 5 quantiles and the fit can be tested on the other quantiles. If the fit is good, researchers can then compute any other statistic about the distribution almost as if they had access to the raw data.

Finally, our paper establishes a number of facts that are useful for the burgeoning literature on misallocation and productivity dispersion, suggesting that models should be able to reproduce Lévy-distributed VA per worker, with tail parameters $\alpha \in [1, 1.5]$ (although possibly with a broader range when considering country-industry-year samples), substantially positive skewness ($\beta \in [0.5, 1]$), and a ‘body’ dispersion $\gamma/\delta \in [0.2, 1]$.

Bibliography

- Acemoglu, D., Ozdaglar, A. & Tahbaz-Salehi, A. (2017), ‘Microeconomic origins of macroeconomic tail risks’, *American Economic Review* **107**(1), 54–108.
- Andrews, D., Criscuolo, C. & Gal, P. N. (2019), ‘The best versus the rest: Divergence across firms during the global productivity slowdown’, *CEP Discussion Paper* **1645**.
- Aoyama, H., Fujiwara, Y. & Gallegati, M. (2015), ‘Micro-macro relation of production: double scaling law for statistical physics of economy’, *Journal of Economic Interaction and Co-ordination* **10**(1), 67–78.
- Aoyama, H., Yoshikawa, H., Iyetomi, H. & Fujiwara, Y. (2009), ‘Labour productivity superstatistics’, *Progress of Theoretical Physics, Supplement* **179**, 80–92.

- Aoyama, H., Yoshikawa, H., Iyetomi, H. & Fujiwara, Y. (2010), 'Productivity dispersion: facts, theory, and implications', *Journal of Economic Interaction and Coordination* 5(1), 27–54.
- Aradanaz-Badia, A., Awano, G. & Wales, P. (2017), 'Understanding firms in the bottom 10% of the labour productivity distribution in Great Britain: "the laggards", 2003 to 2015', *Office for National Statistics* .
- Axtell, R. (forthcoming), 'Dynamics of firms from the bottom up: Data, theories and models'.
- Axtell, R. L. (2001), 'Zipf distribution of U.S. firm sizes', *Science* 293(5536), 1818–1820.
- Ayebo, A. & Kozubowski, T. J. (2003), 'An asymmetric generalization of Gaussian and Laplace laws', *Journal of Probability and Statistical Science* 1(2), 187–210.
- Baily, M. N., Bartelsman, E. J. & Haltiwanger, J. (1996), 'Downsizing and productivity growth: myth or reality?', *Small Business Economics* 8(4), 259–278.
- Bartelsman, E., Haltiwanger, J. & Scarpetta, S. (2013), 'Cross-country differences in productivity: The role of allocation and selection', *American Economic Review* 103(1), 305–34.
- Bartelsman, E. J. & Doms, M. (2000), 'Understanding productivity: Lessons from longitudinal microdata', *Journal of Economic Literature* 38(3), 569–594.
- Bartelsman, E. J. & Wolf, Z. (2018), Measuring productivity dispersion, in E. Grifell-Tatjé, C. K. Lovell & R. C. Sickles, eds, 'Oxford Handbook of Productivity Analysis', Oxford University Press, United Kingdom.
- Berlingieri, G., Blanchenay, P., Calligaris, S. & Criscuolo, C. (2017), 'The multiprod project: A comprehensive overview', *OECD Science, Technology and Industry Working Papers* 2017(4), 1.
- Berlingieri, G., Blanchenay, P. & Criscuolo, C. (2017), 'The great divergence(s)', *OECD STI Policy Papers* 39.
- Bottazzi, G. (2014), Subbotools user's manual: For version 0.9.7.1, 8 september 2004, Technical report, LEM Working Paper Series. Originally published 2004, updated 2014, <http://www.lem.sssup.it/WPLem/files/2004-14.pdf>.
- Bottazzi, G., Cefis, E., Dosi, G. & Secchi, A. (2007), 'Invariances and diversities in the patterns of industrial evolution: Some evidence from Italian manufacturing industries', *Small Business Economics* 29(1-2), 137–159.
- Bottazzi, G. & Secchi, A. (2003), 'Common properties and sectoral specificities in the dynamics of US manufacturing companies', *Review of Industrial Organization* 23(3), 217–232.

- Bottazzi, G. & Secchi, A. (2011), 'A new class of asymmetric exponential power densities with applications to economics and finance', *Industrial and Corporate Change* **20**(4), 991–1030.
- Bouchaud, J.-P. & Potters, M. (2003), *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management*, 2 edn, Cambridge University Press.
- Byrd, R. H., Lu, P., Nocedal, J. & Zhu, C. (1995), 'A limited memory algorithm for bound constrained optimization', *SIAM Journal on Scientific Computing* **16**(5), 1190–1208.
- Campbell, S., Nguyen, T., Sibelle, A. & Soriano, F. (2019), 'Measuring productivity dispersion in selected Australian industries', *Treasury-ABS Working Paper* .
- Clauset, A., Shalizi, C. R. & Newman, M. E. (2009), 'Power-law distributions in empirical data', *SIAM review* **51**(4), 661–703.
- Cohen, J. E., Davis, R. A. & Samorodnitsky, G. (2020), 'Heavy-tailed distributions, correlations, kurtosis and Taylor's Law of fluctuation scaling', *Proceedings of the Royal Society A* **476**(2244), 20200610.
- Cunningham, C., Foster, L., Grim, C., Haltiwanger, J., Pabilonia, S. W., Stewart, J. & Wolf, Z. (2021), 'Dispersion in dispersion: Measuring establishment-level differences in productivity', *IZA Discussion Paper* **14459**.
- De Loecker, J., Obermeier, T. & Van Reenen, J. (2022), Firms and inequalities, in IFS, ed., 'The IFS Deaton Review of Inequalities', IFS.
- Embrechts, P., Kluppelberg, C. & Mikosch, T. (1997), Time series analysis for heavy-tailed processes, in 'Modelling Extremal Events,', Springer, Berlin, Heidelberg, pp. 371–412.
- Fagiolo, G., Napoletano, M. & Roventini, A. (2008), 'Are output growth-rate distributions fat-tailed? some evidence from OECD countries', *Journal of Applied Econometrics* **23**(5), 639–669.
- Fama, E. F. (1965), 'Portfolio analysis in a stable Paretian market', *Management Science* **11**(3), 404–419.
- Foster, L., Grim, C., Haltiwanger, J. C. & Wolf, Z. (2021), Innovation, productivity dispersion, and productivity growth, in C. Corrado, J. Haskel, J. Miranda & D. Sichel, eds, 'Measuring and Accounting for Innovation in the Twenty-First Century', University of Chicago Press, p. 103.
- Gabaix, X. (2009), 'Power laws in economics and finance', *Annual Review of Economics* **1**(1), 255–294.
- Gabaix, X. (2011), 'The granular origins of aggregate fluctuations', *Econometrica* **79**(3), 733–772.

- Gaffeo, E. (2008), 'Lévy-stable productivity shocks', *Macroeconomic Dynamics* **12**(3), 425–443.
- Gaffeo, E. (2011), 'The distribution of sectoral TFP growth rates: International evidence', *Economics Letters* **113**(3), 252–255.
- Ghiglino, C. (2012), 'Random walk to innovation: Why productivity follows a power law', *Journal of Economic Theory* **147**(2), 713–737.
- Goldin, I., Koutroumpis, P., Lafond, F. & Winkler, J. (2021), 'Why is productivity slowing down?', *INET Oxford Working Paper* .
- Gopinath, G., Kalemli-Özcan, Ş., Karabarbounis, L. & Villegas-Sanchez, C. (2017), 'Capital allocation and productivity in South Europe', *The Quarterly Journal of Economics* **132**(4), 1915–1967.
- Gouin-Bonenfant, E. (2020), 'Productivity dispersion, between-firm competition and the labor share', *Draft* .
- Gu, W. (2019), 'Frontier firms, productivity dispersion and aggregate productivity growth in Canada', *International Productivity Monitor* (37), 96–119.
- Haldane, A. (2017), 'Productivity puzzles', *Speech given by Andy Haldane, Chief Economist, Bank of England at the London School of Economics* .
- Haltiwanger, J., Kulick, R. & Syverson, C. (2018), 'Misallocation measures: The distortion that ate the residual', *NBER Working Paper* **24199**.
- Heinrich, T. & Winkler, J. (2020), *finity: Test for Finiteness of Moments in a Distribution*. R package version 0.1.4.1.
URL: <https://CRAN.R-project.org/package=finity>
- Holly, S., Petrella, I. & Santoro, E. (2013), 'Aggregate fluctuations and the cross-sectional dynamics of firm growth', *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **176**(2), 459–479.
- Hsieh, C.-T. & Klenow, P. J. (2009), 'Misallocation and manufacturing TFP in China and India', *The Quarterly Journal of Economics* **124**(4), 1403–1448.
- Ijiri, Y. & Simon, H. A. (1977), *Skew distributions and the sizes of business firms*, Vol. 24, North Holland, Amsterdam.
- Ilzetzki, E. & Simonelli, S. (2017), 'Measuring productivity dispersion: Lessons from counting one-hundred million ballots', *CEPR Discussion Paper* **DP12273**.
- Jäger, K. & The Conference Board (2018), EU KLEMS growth and productivity accounts 2017 release, statistical module 1, Technical report, EU KLEMS Project. Revised 2018, Available online

http://www.euklems.net/TCB/2018/Metholology_EUKLEMS_2017_revised.pdf,
http://www.euklems.net/TCB/2018/EUKLEMS_2018_revision.pdf.

- Kalemli-Ozcan, S., Sorensen, B., Villegas-Sanchez, C., Volosovych, V. & Yesiltas, S. (2015), 'How to construct nationally representative firm level data from the ORBIS global database', *NBER Working Paper* **21558**.
- König, M. D., Lorenz, J. & Zilibotti, F. (2016), 'Innovation vs. imitation and the evolution of productivity distributions', *Theoretical Economics* **11**(3), 1053–1102.
- Lucas, R. E. & Moll, B. (2014), 'Knowledge growth and the allocation of time', *Journal of Political Economy* **122**(1), 1–51.
- Mandelbrot, B. (1960), 'The Pareto-Lévy law and the distribution of income', *International Economic Review* **1**(2), 79–106.
- Mandelbrot, B. (1963), 'New methods in statistical economics', *Journal of Political Economy* **71**(5), 421–440.
- McCulloch, J. H. (1986), 'Simple consistent estimators of stable distribution parameters', *Communications in Statistics-Simulation and Computation* **15**(4), 1109–1136.
- Mitchell, W. C. (1915), 'The making and using of index numbers', *US BLS Bulletin* **284**. Reprinted 1938, accessible at https://fraser.stlouisfed.org/files/docs/publications/bls/bls_0656_1938.pdf.
- Mizuno, T., Ishikawa, A., Fujimoto, S. & Watanabe, T. (2012), 'Power laws in firm productivity', *Progress of Theoretical Physics Supplement* **194**, 122–134.
- Moran, J., Secchi, A. & Bouchaud, J.-P. (2020), Statistical facts on growth rates, in J. Moran, ed., 'Statistical physics and anomalous macroeconomic fluctuations, PhD Thesis', EHESS press.
- Newman, M. E. (2005), 'Power laws, Pareto distributions and Zipf's law', *Contemporary Physics* **46**(5), 323–351.
- Nolan, J. P. (1998), 'Parameterizations and modes of stable distributions', *Statistics & Probability Letters* **38**(2), 187 – 195.
- Nolan, J. P. (2020), *Stable Distributions - Models for Heavy Tailed Data*, Springer.
- Office for National Statistics (2022), 'Firm-level labour productivity estimates from the annual business survey (abs): summary statistics'. <https://www.ons.gov.uk/economy/economicoutputandproductivity/productivitymeasures/datasets/firmlevellabourproductivityestimatesfromtheannualbusinesssurveyabssummarystatistics>.
- Oliveira-Cunha, J., Kozler, J., Shah, P., Thwaites, G. & Valero, A. (2021), Business time: How ready are UK firms for the decisive decade?, Technical report, The Resolution Foundation.

- Oulton, N. (2000), 'A tale of two cycles: Closure, downsizing and productivity growth in UK manufacturing, 1973-89', *National Institute Economic Review* **173**, 66–79.
- Perla, J. & Tonetti, C. (2014), 'Equilibrium imitation and growth', *Journal of Political Economy* **122**(1), 52–76.
- Resnick, S. I. (2007), *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*, Springer Series in Operations Research and Financial Engineering, Springer.
- Rotemberg, M. & White, T. K. (2021), 'Plant-to-table (s and figures): Processed manufacturing data and measured misallocation', *Working Paper* . https://wp.nyu.edu/mrotemberg/wp-content/uploads/sites/8049/2021/07/RW_plant_to_table_July_2021.pdf.
- Samuelson, P. A. (1967), 'Efficient portfolio selection for Pareto-Lévy investments', *Journal of Financial and Quantitative Analysis* **2**(2), 107–122.
- Schwarzkopf, Y., Axtell, R. & Farmer, J. D. (2010), 'An explanation of universality in growth fluctuations', *SSRN* **1597504**.
- Sornette, D. (2006), *Critical phenomena in natural sciences: chaos, fractals, selforganization and disorder: concepts and tools*, Springer Science & Business Media.
- Souma, W., Ikeda, Y., Iyetomi, H. & Fujiwara, Y. (2009), 'Distribution of labour productivity in Japan over the period 1996–2006', *Economics: E-journal* **3**(1).
- Syverson, C. (2011), 'What determines productivity?', *Journal of Economic Literature* **49**(2), 326–65.
- Trapani, L. (2016), 'Testing for (in)finite moments', *Journal of Econometrics* **191**(1), 57 – 68.
- Voitalov, I., van der Hoorn, P., van der Hofstad, R. & Krioukov, D. (2019), 'Scale-free networks well done', *Physical Review Research* **1**(3), 033034.
- Xia, Z. (2022), *skewtDist: Asymmetric Student t-distributions*. R package version 0.3.0.
URL: <https://github.com/dan9401/st>
- Yang, J., Heinrich, T., Winkler, J., Lafond, F., Koutroumpis, P. & Farmer, J. D. (2024), 'Measuring productivity dispersion: a parametric approach using the Lévy alpha-stable distribution', *Industrial and Corporate Change* p. dtae021.
- Zhu, D. & Galbraith, J. W. (2010), 'A generalized asymmetric Student-t distribution with application to financial econometrics', *Journal of Econometrics* **157**(2), 297–305.
- Zhu, D. & Galbraith, J. W. (2011), 'Modeling and forecasting expected shortfall with the generalized asymmetric Student-t and asymmetric exponential power distributions', *Journal of Empirical Finance* **18**(4), 765–778.

Appendix

4.A Data appendix

4.A.1 Raw data

We use the Orbis Europe firm level database, which is part of the Orbis data provided by Bureau van Dijk. The database encompasses more than 21 million firms of all sizes from 44 different European countries. Around 23 million firm-year observations for 7 million unique firms in 19 countries are used in the present analysis; see Table 4.A.2. The variables used are listed in Table 4.A.1.

Table 4.A.1: Variables from Orbis Europe used for the analysis

Orbis Code	Notation	Description
IDNR		Firm's identification number
EBITDA	π	Nominal Earnings Before Interest, Taxes, Depreciation & Amortization
STAF	ω	Nominal Wages (staff costs)
EMPL	L	Employment
NACE_PRIM_CODE		Industrial classification code (NACE Rev. 2)
CLOSDATE_year		Year

Notes: See Eq. 4.1 for the construction of VA and Labour productivity.

4.A.2 Data cleaning

The first step in processing the data is to ensure that no observations have missing values for their reporting IDNR or year. We only use unconsolidated data in order to avoid double counting. Since some accounts provide a default year for various reasons, observations where the reported year is not between 2006 and 2017 are removed. Negative sales, wages, and employment observations are regarded as missing. Furthermore, we remove firm-year observations where employment or value added are missing.

In addition, we discard country-year samples without at least 10,000 observations for five years or more. This reduces the list of countries considered from 44 to 19 countries (See Table 4.A.2).¹⁸

¹⁸For 25 countries, there was insufficient data: Albania, Austria, Belarus, Bosnia and Herzegovina, Cyprus, Greece, Iceland, Ireland, Kosovo, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Monaco, Montenegro, Netherlands, Norway, North Macedonia, Moldova, Russian Federation, Serbia, Switzerland, Turkey, Ukraine.

4.A.3 Deflation

The KLEMS database provides the most comprehensive data on deflation for the countries and industries covered by the Orbis Europe sample. In particular, tables `Statistical_Capital` and `Statistical_National-Accounts` provide value added and gross output for two-digit NACE Rev.2 industries for the countries in the Orbis Europe sample. Depending on its location, industry, and year of reporting, we deflate firms' value added using the value added deflator.

Table 4.A.2: Number of observations per country-year after cleaning

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Belgium	24,823	88,877	94,373	94,899	95,631	96,810	97,357	97,093	97,324	97,629	96,909	60,423
Bulgaria	0	12,231	38,205	38,136	39,190	81,215	93,304	94,817	95,577	99,594	99,453	102,942
Croatia	16,451	38,875	39,957	42,387	40,536	39,360	40,340	40,472	41,208	42,434	44,062	0
Czech Republic	24,328	52,697	46,049	70,294	73,284	77,457	76,924	78,795	75,475	77,139	60,192	40,554
Denmark	0	0	0	0	0	0	10,153	13,116	12,649	14,232	41,054	45,520
Estonia	0	14,505	16,095	13,939	15,932	16,909	18,088	19,017	20,185	20,511	20,926	19,876
Finland	14,585	31,271	30,235	32,541	32,344	38,403	38,749	39,827	40,825	37,529	35,797	36,760
France	188,682	248,892	235,860	249,724	259,352	233,105	199,456	207,906	207,671	155,439	128,023	82,804
Germany	19,537	36,256	37,810	38,440	39,498	42,164	50,222	73,045	40,261	37,871	33,155	0
Hungary	0	65,589	16,220	109,059	91,033	88,773	120,705	121,833	129,449	132,779	133,472	129,107
Italy	80,270	205,580	300,766	255,141	201,252	406,372	425,845	418,763	422,876	437,585	435,280	380,695
Poland	13,548	25,968	31,798	54,341	18,129	14,024	15,311	13,514	0	0	0	25,350
Portugal	97,015	202,418	208,151	201,864	197,862	192,167	187,614	184,607	187,208	191,268	193,563	193,080
Romania	0	167,761	175,666	161,536	162,342	173,394	179,953	182,637	178,535	183,662	191,385	0
Slovakia	0	16,158	22,700	33,137	32,910	35,012	33,821	39,242	43,622	34,241	34,956	0
Slovenia	0	0	0	9,659	29,250	29,453	27,599	27,442	28,932	29,746	30,959	31,813
Spain	270,505	424,877	468,998	461,257	436,631	424,314	404,536	390,900	392,478	403,412	405,003	345,054
Sweden	18,283	83,593	100,106	102,341	104,274	106,361	107,975	108,888	110,179	112,690	114,515	100,681
United Kingdom	16,128	49,873	53,686	51,906	50,259	49,457	49,832	51,090	52,126	52,793	52,818	0

4.A.4 Size proportion

Table 4.A.3 displays the number of observations and company size composition by country. Among the top three countries with the largest sample sizes (France, Italy, and Spain), France’s size composition closely aligns with the sample mean. Spain tends to have a relatively lower share of large and very large firms, while Italy has a lower share of small firms. For this reason, we showcase our main results regarding the distribution fit and model comparison (see Sec 4.3.2) using the French sample in the paper.

Table 4.A.3: Size proportion of country sample (%)

Country	Obs	Small	Medium	Large	Very Large
Belgium	1,042,148	72.18	20.78	5.85	1.19
Bulgaria	794,664	74.19	22.75	2.62	0.44
Croatia	426,082	81.30	16.07	2.22	0.41
Czech Republic	753,188	59.15	32.07	7.78	1.00
Denmark	136,724	62.38	27.21	7.29	3.12
Estonia	195,983	83.05	14.91	1.92	0.12
Finland	408,866	62.91	28.55	7.05	1.49
France	2,396,914	69.34	22.89	6.55	1.22
Germany	448,259	11.99	39.19	36.13	12.69
Hungary	1,138,019	82.23	15.43	2.02	0.32
Italy	3,970,425	56.54	35.40	7.04	1.03
Poland	211,983	35.22	43.29	18.48	3.01
Portugal	2,236,817	83.40	14.50	1.82	0.28
Romania	1,756,871	84.33	13.71	1.66	0.30
Slovakia	325,799	74.11	21.27	3.97	0.65
Slovenia	244,853	80.79	16.38	2.48	0.35
Spain	4,827,965	71.21	24.48	3.75	0.57
Sweden	1,169,886	70.62	24.59	4.11	0.68
United Kingdom	529,968	33.19	31.33	27.40	8.08
Mean	1,211,338	65.69	24.46	7.90	1.94

4.B The Lévy alpha-stable distribution

The Lévy alpha-stable distribution is a natural candidate for many distributions exhibiting heavy tails, as it emerges from the Generalized Central Limit Theorem (GCLT). We first describe the key properties of the Lévy alpha-stable distribution, and then discuss the GCLT.

4.B.1 Characteristics of the Lévy alpha-stable distribution

The Lévy alpha-stable distribution is a four-parameter distribution with parameters α (tail exponent), β (skewness), γ (scale), and δ (location). The density function exists in closed form only in a few special cases such as $\alpha = 2$ (Gaussian), $\{\alpha = 1, \beta = 0\}$ (Cauchy), and $\{\alpha = 1/2, \beta = 1\}$ (standard Lévy). Out of several alternatives, we use Nolan’s (1998) S_0

parametrisation, which has the characteristic function

$$\varphi(t) \equiv E[\exp(itX)] = \begin{cases} \exp\left(-|\gamma t|^\alpha \left[1 + i\beta \tan\left(\frac{\pi\alpha}{2}\right) \operatorname{sgn}(t)(|\gamma t|^{1-\alpha} - 1)\right] + i\delta t\right) & \alpha \neq 1, \\ \exp\left(-\gamma|t| \left[1 + i\beta \frac{2}{\pi} \operatorname{sgn}(t) \log(\gamma|t|)\right] + i\delta t\right) & \alpha = 1, \end{cases} \quad (4.7)$$

where i is the imaginary unit, $t \in \mathbb{R}$ is the argument of the characteristic function, and $\operatorname{sgn}()$ is the sign function. This particular parametrisation is useful for numerical work and statistical inference since the characteristic function is continuous in all four parameters (Nolan 2020, p.7).

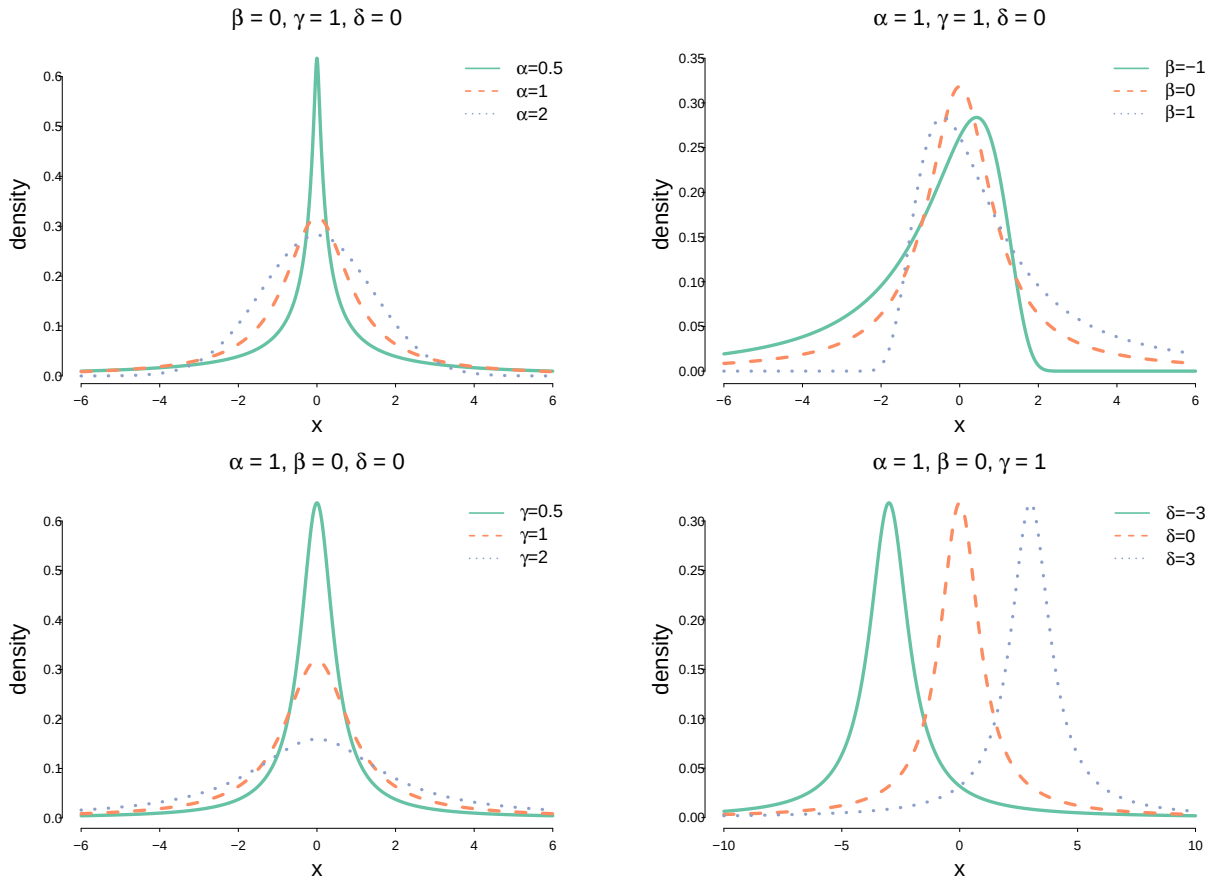


Figure 4.B.1: Probability densities of Lévy alpha-stable distributions for different parameter values: The top left plot varies the α parameter (tail exponent), which determines how heavy the tail is. The top right plot varies the β parameter (skew parameter). The bottom left plot varies the γ parameter, which determines the ‘scale’, or ‘width’, of the distribution. The last plot varies the δ parameter, which shifts the location of the modal value of the distribution.

Fig 4.B.1 shows the probability density function of the Lévy alpha-stable distribution, with three different values for each of the four parameters. Starting from the top left, the parameter $\alpha \in (0, 2]$ is also called the tail exponent because the tail of the distribution decays as a power law with exponent α , that is, $P(X > x) \approx Cx^{-\alpha}$. This implies that scaling the value of x by a factor h just scales the tail probability by $h^{-\alpha}$, since $h^{-\alpha}P(X > x) \approx C(hx)^{-\alpha}$.

For example, if $\alpha = 1$, values of x that are twice as large (i.e., $h = 2$) as some reference $\tilde{x}$ will be half as common, no less. As a result, with power law tails extreme values are very common and can dominate certain moments of the distribution. Thus, the tail parameter α is an important indicator of dispersion. The lower α , the thicker the tail (the more ‘tail dispersion’ there is). However, note that while a smaller α indicates more dispersion because of a heavier tail, a smaller α also leads to a higher concentration of values close to the mode, and thus, in this sense, to a lower dispersion in the body of the distribution.

The top right panel shows the effects of the ‘skew’ or ‘asymmetry’ parameter $\beta \in [-1, 1]$. For $\beta = 0$, the distribution is symmetric, for $\beta > 0$ right-skewed, and for $\beta < 0$ it is left-skewed (except when $\alpha = 2$, as the skew parameter β vanishes in that case). In practice, we will find that productivity distributions are right skewed.

The bottom-left panel shows the effect of the scale parameter $\gamma \in [0, +\infty]$. The higher the scale parameter γ , the wider the body of the distribution. The parameter γ is not dimensionless; it is expressed in the same units as the data. We can get more intuition on the role of γ by considering special cases (Nolan 2020, p.168). First, γ is highly related to the (25–75) Inter Quartile Range; in fact, when β is small and α is large (say, $\beta = 0.5$ and $\alpha = 1.3$), the interquartile range divided by 2 is a good estimator of γ . Second, for large enough α , the re-centred data (so that $\delta = 0$) can be used to estimate γ as the median of the absolute values.

Finally, the bottom right panel shows the effect of the location parameter $\delta \in (-\infty, +\infty)$, which is fairly intuitive even though it is not in general exactly related to key quantities such as the mode, mean, or median. For symmetric stable distributions, the median is equal to the location, and for asymmetric distributions the median remains close to the location as long as $|\beta|$ is not too large and α is not too small (Nolan 2020, p.168). When $\alpha > 1$, the sample mean is actually a consistent (but slowly converging) estimator of δ (Nolan 2020, p.174). Nolan (2020, p.92) defines an S_2 parametrisation specifically so that δ is the mode.

We can gain further intuition by considering special cases: For $\alpha = 2$, the distribution becomes a Gaussian with variance $2\gamma^2$ and mean δ . For $\alpha = 1, \beta = 0$ the distribution becomes Cauchy with scale parameter γ and location parameter δ . The role of δ and γ as location and scale parameters can be seen clearly by writing the Cauchy density as $f(x) \propto \left[1 + \left(\frac{x-\delta}{\gamma}\right)^2\right]^{-1}$. More generally, if $X \sim S_0(\alpha, \beta, \gamma, \delta)$, then $(X - \delta)/\gamma \sim S_0(\alpha, \beta, 1, 0)$.

One of the key characteristics of the Lévy alpha-stable distribution is that γ and α measure qualitatively different aspects of dispersion, providing a richer perspective on observed empirical patterns.

4.C Fitting procedure

4.C.1 Fitting methods

There are three primary estimation methods for the Lévy alpha-stable distribution in the context of asymmetric distributions: maximum likelihood (MLE), the quantile-based estimator of McCulloch (1986), and a method based on empirical characteristic function (Nolan

2020, Chap. 4). Among these, the quantile-based estimation is the fastest as it requires only a few quantiles to estimate the parameters. However, we noted that the quantile estimation can produce imprecise results, especially for countries where the empirical distribution is substantially right-skewed, such as Bulgaria, Estonia, Romania, Slovakia, and Slovenia. For instance, the quantile method occasionally returns a value of $\beta = 0.95$, the upper limit for β within this estimation method.

For this reason, we employ the MLE to obtain the most accurate estimates utilising the Limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS-B) algorithm (Byrd et al. 1995). However, even though the MLE demonstrates superior performance for the Lévy parameter estimation (Nolan 2020, Chap. 4), it is computationally highly prohibitive. This is because the density function of the Lévy alpha-stable distribution does not have a closed form, making its evaluation time-consuming. To get around this problem, we opt for using a sub-sample of the data. To ensure that this sub-sampling doesn't introduce bias into our MLE estimates, we conduct a Monte Carlo simulation. Specifically, we generate synthetic data of size 1,000,000 from the Lévy alpha-stable distribution using parameters $\alpha = 1.2, \beta = 0.55, \gamma = 1, \delta = 0$. From this data, we draw random samples of sizes $N = 5000, 10,000, 15,000, 20,000, 30,000$. For each sample size, we draw 100 samples and estimate the parameters using MLE. Subsequently, we compute the mean and both the 2.5th and 97.5th quantiles of these 100 observations.

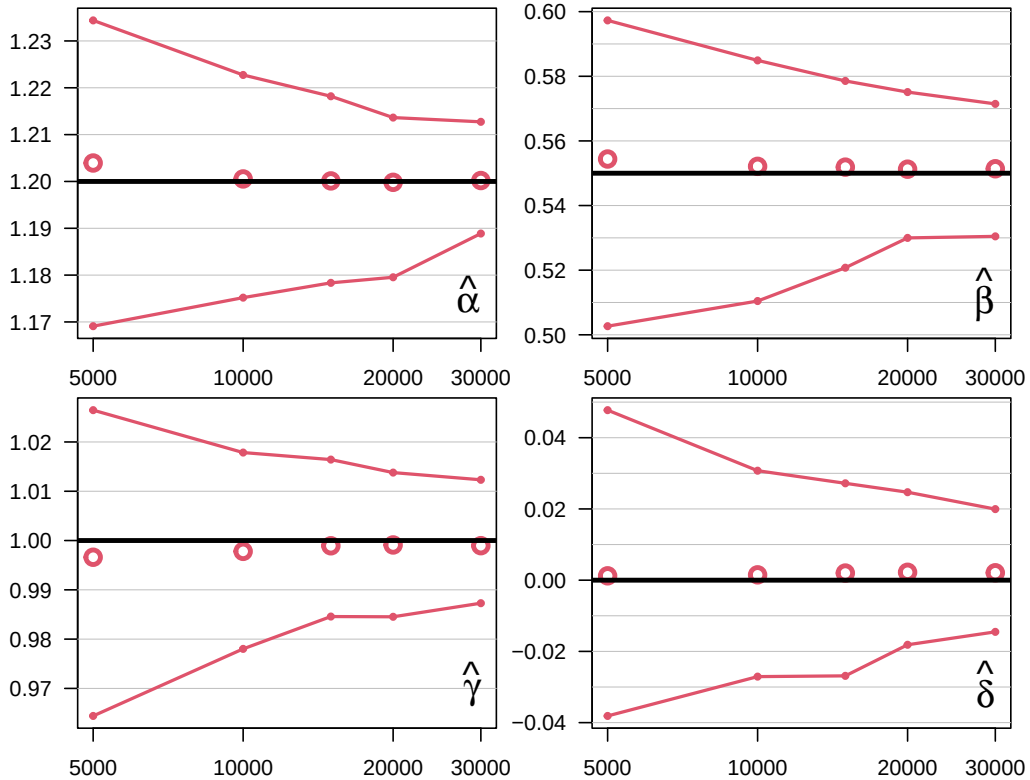


Figure 4.C.1: **Performance of the MLE in sub-sampling:** For each sample size, we sample 100 sets of i.i.d. values from the synthetic data of size 100,000 generated from the Lévy alpha-stable distribution using parameters $\alpha = 1.2, \beta = 0.55, \gamma = 1, \delta = 0$. The black lines show the true values of the parameters. The dots show the average of the estimated parameters over the 100 repetitions, and the line shows the 2.5th and 97.5th percentiles.

Fig. 4.C.1 displays the results. The black lines show the true values of the parameters while the red dots show the average of the estimated parameters over the 100 repetitions, and the red line shows the 2.5th and 97.5th percentiles. For all four parameters, the MLE accurately estimates the true value even with a very small sub-sample size. The error bar becomes smaller as the size of the sub-sample increases. When the size of the sub-sample reaches 30,000, the 95% error bar is only around 0.01 and 0.02 away from the true value for α and β , respectively. To conclude, we prefer to use the MLE on a single manageable sub-sample of the data than to use the quantile method on the full sample.

In using the MLE for Lévy parameter estimation in this paper, we employ a sub-sample of 20,000 and 100,000 observations for country-year sample and country sample estimates.

4.C.2 Estimated parameters

Table 4.C.1 presents the estimated parameters of the Lévy alpha-stable distribution using the MLE. When estimating country aggregate samples, we use sub-samples of 100,000 observations. For Croatia, the Czech Republic, Denmark, and Sweden, the built-in MLE function of the **StableEstim** package fails to operate properly. In these cases, we rely on the `optim()` function, which is the base optimisation function in R, to obtain the results.

Table 4.C.1: The estimated parameters of Lévy alpha-stable distribution by country

Country	$\hat{\alpha}$	$sd(\hat{\alpha})$	$\hat{\beta}$	$sd(\hat{\beta})$	$\hat{\gamma}$	$sd(\hat{\gamma})$	$\hat{\delta}$	$sd(\hat{\delta})$
Belgium	1.310	0.004	0.798	0.005	17.84	0.06	48.68	0.10
Bulgaria	1.027	0.003	0.769	0.004	4.95	0.02	7.45	0.03
Croatia	1.220	0.004	0.747	0.005	42.01	0.06	77.83	0.09
Czech Republic	1.195	0.004	0.791	0.005	190.96	0.80	306.84	1.00
Denmark	1.282	0.004	0.587	0.006	155.13	0.28	443.85	0.80
Estonia	1.164	0.004	0.855	0.004	5.86	0.02	10.13	0.03
Finland	1.415	0.004	0.769	0.006	16.50	0.05	41.63	0.09
France	1.341	0.004	0.830	0.005	15.29	0.05	40.33	0.08
Germany	1.156	0.004	0.787	0.004	21.11	0.08	49.03	0.11
Hungary	1.192	0.010	0.796	0.008	1312.80	36.55	2154.31	40.24
Italy	1.390	0.004	0.752	0.006	15.65	0.05	30.78	0.09
Poland	1.105	0.004	0.863	0.003	28.79	0.11	41.63	0.15
Portugal	1.346	0.004	0.688	0.006	6.84	0.02	12.12	0.04
Romania	1.148	0.004	0.768	0.004	12.69	0.05	15.50	0.07
Slovakia	1.176	0.004	0.795	0.004	8.03	0.03	12.25	0.04
Slovenia	1.291	0.004	0.801	0.005	7.56	0.03	21.50	0.04
Spain	1.289	0.004	0.674	0.006	11.10	0.04	25.78	0.06
Sweden	1.417	0.004	0.673	0.007	164.79	0.43	455.90	0.48
United Kingdom	1.093	0.004	0.545	0.005	18.71	0.08	31.01	0.09

Notes: The parameters are estimated using the MLE from subsamples of 100,000 observations. All MLE results are based on the Limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS-B) algorithm (Byrd et al. 1995).

4.D Reference model and comparison of the fit

As a reference model, we use the 5 parameter Asymmetric Student-t (AST) distribution by [Zhu & Galbraith \(2010\)](#) and the 5-parameter Asymmetric Exponential Power (AEP) distribution by [Bottazzi & Secchi \(2011\)](#). In the following, we give the functional forms, details of the fitting procedures, and a comparison of the goodness of the results.

4.D.1 AST and AEP: models, fitting procedure, and estimated parameters

5-parameter AST The probability density function of the 5-parameter AST distribution is ([Zhu & Galbraith 2010](#))

$$f(x) = \begin{cases} \frac{1}{\sigma} \left[1 + \frac{1}{\nu_1} \left(\frac{x-\mu}{2\alpha_{AST}\sigma K(\nu_1)} \right)^2 \right]^{-(\nu_1+1)/2} & \text{for } x \leq \mu \\ \frac{1}{\sigma} \left[1 + \frac{1}{\nu_2} \left(\frac{x-\mu}{2(1-\alpha_{AST})\sigma K(\nu_2)} \right)^2 \right]^{-(\nu_2+1)/2} & \text{for } x > \mu \end{cases}$$

with

$$K(\nu) = \Gamma((\nu+1)/2) / [\sqrt{\pi\nu}\Gamma(\nu/2)],$$

where μ is the location parameter, σ is the scale parameter, $\alpha_{AST} \in (0,1)$ is the skewness parameter, and $\nu_1 > 0$ and $\nu_2 > 0$ are the left and right tail parameters. $\Gamma()$ is the Gamma function. We fit the AST using the maximum likelihood estimation. For the computation, we use the R-package **SkewtDist** package ([Xia 2022](#)). Table 4.D.1 shows the estimated parameters of the 5-parameter AST distribution.

5-parameter AEP The probability density function of the 5-parameter AEP distribution is ([Bottazzi & Secchi 2011](#))

$$f(x) = \begin{cases} \frac{1}{A} \exp\left(-\frac{1}{b_l} \left| \frac{x-\theta}{a_l} \right|^{b_l}\right) & \text{for } x < \theta \\ \frac{1}{A} \exp\left(-\frac{1}{b_r} \left| \frac{x-\theta}{a_r} \right|^{b_r}\right) & \text{for } x > \theta \end{cases}$$

with

$$A = a_l b_l^{\frac{1}{b_l}} \Gamma\left(1 + \frac{1}{b_l}\right) + a_r b_r^{\frac{1}{b_r}} \Gamma\left(1 + \frac{1}{b_r}\right).$$

θ is the location parameter; the scale parameter is separate for the left side (a_l) and right side (a_r) of the distribution $x < \theta$ and $x > \theta$. b_l and b_r are a left-shape and a right-shape parameter. The distribution recovers the normal distribution $\mathcal{N}(\mu, \sigma)$ for parameter values $\theta = \mu, a_l = a_r = \sigma, b_l = b_r = 2$.

The distribution can be fitted with interval-constrained likelihood optimisation, which is, however, computationally expensive. We used the Subbotools package ([Bottazzi 2014](#)). To speed up the computation, we first perform a 3-parameter exponential power¹⁹ fit and

¹⁹This distribution is the standard Subbotin distribution and is obtained for $a_l = a_r$ and $b_l = b_r$ for the 5-parameter AEP.

Table 4.D.1: Estimated parameters of the 5-parameter AST distribution by country

Country	μ	σ	α_{AST}	ν_1	ν_2
Belgium	39.57	59.64	0.28	2.22	1.75
Bulgaria	3.91	16.58	0.20	1.03	1.34
Croatia	51.97	141.04	0.25	1.57	1.72
Czech Republic	182.78	648.84	0.24	1.61	1.61
Denmark	392.56	510.91	0.34	1.92	1.54
Estonia	4.82	20.22	0.16	1.22	1.87
Finland	37.89	54.09	0.36	3.18	1.65
France	32.24	51.21	0.28	2.45	1.84
Germany	37.69	69.18	0.26	1.58	1.33
Hungary	1365.14	4519.57	0.24	1.62	1.62
Italy	21.98	53.46	0.29	2.05	2.16
Poland	16.82	99.87	0.17	1.36	1.55
Portugal	10.03	22.11	0.34	2.11	1.63
Romania	5.67	43.53	0.20	1.23	1.68
Slovakia	6.49	27.39	0.22	1.41	1.66
Slovenia	16.27	25.72	0.23	1.89	1.99
Spain	21.11	36.73	0.31	1.83	1.62
Sweden	395.64	547.79	0.34	2.46	1.98
United Kingdom	21.93	61.51	0.30	1.14	1.29

Notes: The parameters are estimated using the MLE from the R-package **SkewtDist** package (Xia 2022).

then a 5-parameter AEP fit that is initialised with the estimates from the 3-parameter exponential power fit. Table 4.D.2 shows the estimated parameters of the 5-parameter AEP distribution.

4.E Derivation of the scaling of the sample standard deviation with sample size

Here we provide a highly stylised, heuristic derivation for the scaling of the sample standard deviation with sample size in the case of a Lévy alpha-stable distributed random variable. The key to this phenomenon is that because the theoretical moment is infinite, the larger the sample size, the higher is the chance that an extreme event is drawn. These extreme events are so extreme that they dominate the sum of squares from which the variance is computed. Thus, the larger the sample size, the larger the sample variance.

Sornette (2006) and Bouchaud & Potters (2003), for instance, provide more precise statements. Here we expose the argument in the simplest, albeit non-rigorous way. We discuss the maximum, but analogous arguments apply to the minimum.

First, we note that for N large enough, the sample maximum will be dictated by the tail. A key characteristic of the Lévy alpha-stable distribution is that it has power law tails, that is, for large x (Nolan 2020),

$$P(X > x) \sim x^{-\alpha}. \quad (4.8)$$

Now, in a sample of size N , we would hardly expect to see an extreme value that has

Table 4.D.2: Estimated parameters of the 5-parameter AEP distribution by country

Country	b_l	b_r	a_l	a_r	θ
Belgium	0.65	0.55	14.37	34.02	43.66
Bulgaria	0.88	0.36	12.12	15.58	17.44
Croatia	0.85	0.48	50.94	85.02	96.00
Czech Republic	1.00	0.40	355.56	460.68	575.00
Denmark	0.51	0.46	140.60	269.71	413.18
Estonia	0.46	0.60	3.45	13.39	4.65
Finland	1.13	0.43	27.73	31.15	61.00
France	0.83	0.49	16.88	27.96	45.86
Germany	0.50	0.41	16.62	43.97	44.50
Hungary	0.53	0.51	973.25	2711.20	1556.00
Italy	0.65	0.63	14.05	28.49	25.32
Poland	0.58	0.47	21.33	64.35	33.29
Portugal	0.51	0.48	5.74	11.48	10.39
Romania	0.46	0.58	9.12	28.26	5.35
Slovakia	0.50	0.55	5.73	17.24	6.93
Slovenia	0.65	0.67	5.01	15.82	16.37
Spain	0.51	0.50	9.38	20.28	21.77
Sweden	1.01	0.50	245.11	269.86	590.50
United Kingdom	0.41	0.41	18.26	36.68	25.01

Notes: The parameters are estimated using the interval-constrained likelihood optimisation from the Subbotools package (Bottazzi 2014)

chances of occurring less than $1/N$. Thus, we may define the ‘typical’ value of the maximum as the value $X_{\max}$ such that $1/N = P(X > X_{\max})$. Using Eq. 4.8, we have²⁰ $1/N \sim X_{\max}^{-\alpha}$, and solving for $X_{\max}$ gives

$$X_{\max} \sim N^{\frac{1}{\alpha}}, \quad (4.9)$$

For simplicity, let us assume a mean of zero²¹, so that the sample variance is just the average squared value,

$$\text{Var}(X) \sim \frac{1}{N} \sum_{i=1}^N X_i^2. \quad (4.10)$$

In a Lévy alpha-stable distribution, the square of the maximum (or minimum, if larger in absolute value) is so large that it dominates the entire sum of squares, such that we may approximate

$$\sum_{i=1}^N X_i^2 \approx X_{\max}^2 \sim N^{\frac{2}{\alpha}}, \quad (4.11)$$

where the last step uses Eq. 4.9. Now, inserting Eq. 4.11 into 4.10, and taking square root,

²⁰Sornette (2006) shows that this is the value of the maximum that is not exceeded with probability $1/e \approx 37\%$. Generalising to the value of the maximum that is not exceeded with probability p implies the condition $\frac{\ln(1/p)}{N} = P(X > X_{\max})$, which does not change the scaling of $X_{\max}$ with N . Newman (2005) shows that the expected value of the maximum also scales as $N^{\frac{1}{\alpha}}$.

²¹This is without loss of generality, see (2.8) in Cohen et al. (2020).

we find that the standard deviation depends on the sample size as

$$\sqrt{\text{Var}(X)} \sim N^{\frac{1}{\alpha}-\frac{1}{2}}.$$

Note that when $\alpha = 2$, so that the distribution is Gaussian, the sample standard deviation does *not* increase with sample size, as expected.

See also the derivations leading to (4.52) in [Sornette \(2006\)](#) and the heuristic proof based on the rank-size rule in [Gabaix \(2011, Proof of Prop. 2\)](#). A rigorous proof ([Gabaix 2011](#), [Cohen et al. 2020](#)) simply uses the fact that since X_i is (by assumption) regularly varying with index α , it follows that X_i^2 is regularly varying with index $\alpha/2$ ([Gabaix 2009](#), e.g. Eq.8), and we can apply the GCLT to the sum of X_i^2 to obtain the right scaling. The rigorous proof makes it clear that we expect the sample standard deviation to exhibit substantial variations, since the convergence is in distribution to a Lévy stable variable.

Chapter 5

Accounting for the slowdown in labour productivity

(With Ian Goldin, Pantelis Koutroumpis and François Lafond)

*This chapter is an excerpt from a paper titled ‘Why is productivity slowing down?’, published in *Journal of Economic Literature* ([Goldin et al. 2024](#)).*

The following excerpt was written for a literature review on the slowdown of labour productivity growth in advanced economies, following the financial crisis. The sections included are those I was mostly responsible for, and which best fit the narrative of the overall thesis. All further details are available in [Goldin et al. \(2024\)](#).

This study of labour productivity is a first step in understanding the methods used to study aggregates. Specifically, it seeks to evaluate whether labour productivity slowed down across all economic activities, or whether unproductive activities became a more important part of the economy. I re-visit some results in this paper in Chapter 6, where I take a closer look at the role of dispersion in the construction of economic indices.

5.1 Introduction

Labour productivity growth is widely seen as the main long-run determinant of per capita output growth and improving living standards¹.

The decline in measured labour productivity growth over recent decades is a matter of considerable concern and debate among academics, as it is in business and government. Three decades after Robert Solow’s famous quip that ‘you can see the computer age everywhere but in the productivity statistics’ (Solow 1987), this slowdown remains a puzzle, not least for those who believe that technological change is accelerating.

Table 5.1.1: Labour Productivity (LP) slowdown and per capita GDP gap

	LP growth		Slowdown	GDP per capita 2017	“Missing” GDP per capita
	1996–2005	2006–2017			
France	1.65	0.66	0.99	€30,512	€3,836
Germany	1.85	0.91	0.94	€35,217	€4,203
Japan	1.68	0.85	0.82	¥4,155,243	¥356,944
United Kingdom	2.21	0.45	1.75	£27,487	£6,443
United States	2.62	1	1.61	\$59,015	\$12,610

Notes: Growth of labour productivity is per hour worked, and GDP per capita is in 2017 national currency units, using data from EU-KLEMS 2019 (Stehrer et al. 2019) and the Conference Board. The periods for Japan (1995–2015) and the US (1998–2017) are slightly different due to data coverage, see Appendix 5.A for details.

The slowdown is indisputable. Table 5.1.1 demonstrates that labour productivity growth rates have at least halved since the 1996–2005 period, making GDP per capita in 2017 several thousand dollars lower than it would have been based on the previous trend (Syverson 2017). Why is productivity slowing down?

Key results and structure of the paper This paper synthesises a large literature that attempts to explain the slowdown. Before delving into explanations, in Section 5.2 we clarify the nature of the problem by using standard growth accounting for five large developed countries, France, Germany, Japan, the UK and the US, with data from the EU KLEMS 2019 (Stehrer et al. 2019). Comparing the period 2006–2017 against 1996–2005, we confirm the well-established result that most of the slowdown is driven by Total Factor Productivity (TFP) and capital deepening, with a smaller contribution from labour composition, and non-negligible variation across countries. The slowdown is pervasive across industries, with changing sectoral shares explaining very little of the slowdown (except in Germany). Manufacturing is the only industry that is a substantial contributor in all countries.

¹By definition, the growth in output per capita is equal to the growth in output per *worker* plus the growth of the ratio of number of workers over the total population. In this paper, we focus on the first term, labour productivity, but changes to labour participation and employment rates have been (and will continue to be) important, particularly due to aging (see e.g. Ramey et al. (2020) and Vollrath (2020)). We also mostly exclude from our discussion the debate over which measure of output is the best metric for welfare evaluation.

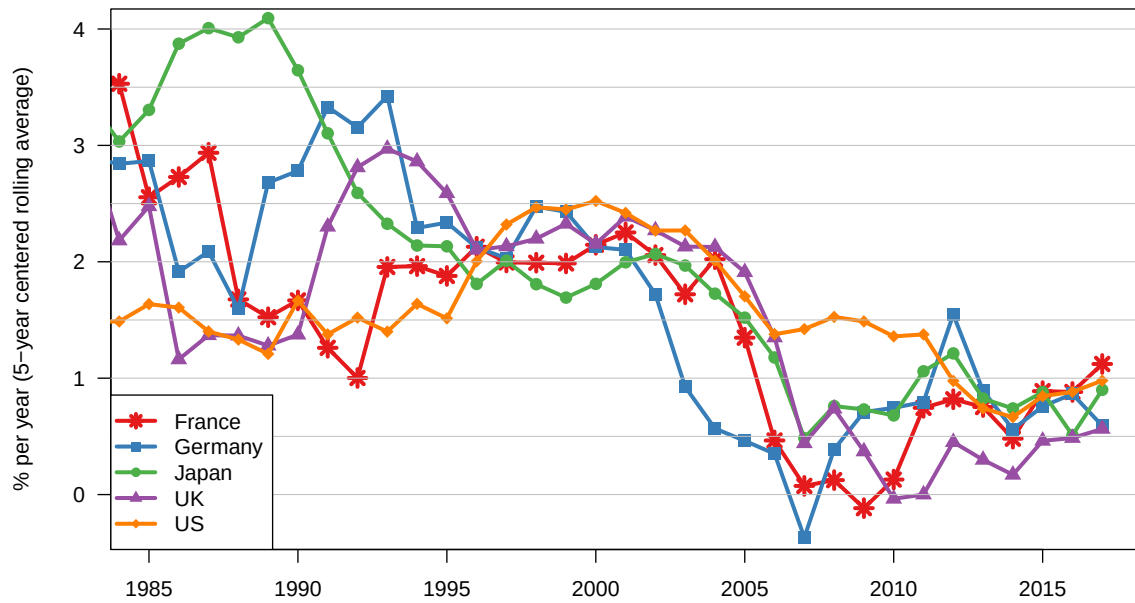


Figure 5.1.1: **Recent trends in labour productivity:** Data from the Long-Term Productivity Database (Bergeaud et al. 2016, 2021).

5.2 Accounting for the slowdown

In this section, we provide two standard decompositions of labour productivity growth, using the 2019 vintage of KLEMS (Stehrer et al. 2019). The first, ‘sources of growth’ decomposition, provides an organising framework for the rest of the paper. It separates labour productivity growth into an increase in inputs (labour composition and capital per worker), and a residual, corresponding to a pure increase in efficiency in the use of inputs, TFP. We find considerable heterogeneity between countries, and that the slowdown in TFP broadly represents the largest contributor to the labour productivity slowdown, with a smaller role played by the slowdown of capital deepening.

The second decomposition we present is a ‘within-between’ decomposition of labour productivity growth by industry. This allows us to show to which extent the slowdown is pervasive, or whether it is confined to specific industries or is due to a reallocation towards low productivity or low productivity growth industries. As with the sources-of-growth decomposition, there is some heterogeneity across countries, but patterns are discernible. The slowdown is broadly pervasive, but a large part of the overall slowdown can be traced back to key industries, in particular manufacturing. Structural change appears to play a limited role, which may reflect the relatively short time periods considered.

5.2.1 Contributions of inputs growth and TFP

Growth Accounting The premise of growth accounting is that aggregate output grows either because more inputs are used, or because they are used more efficiently. Solow (1957) introduced a straightforward method to produce this decomposition: the contribution of each input is computed as its growth rate times its share in income. The contribution of ef-

efficiency is then the part of output growth that is left after the contribution of all inputs has been accounted for. This decomposition is rooted in clear economic assumptions: a stable and smooth functional relationship between inputs and outputs at the economy-wide aggregation level, inputs paid at their marginal product, constant returns to scale, and Hicks-neutral technical change. Solow (1957) originally found that most of US growth over 1909–1949 was not due to the growth of inputs, but to inputs being used more efficiently. This was dubbed the ‘Solow residual’ and came to be described as ‘a measure of our ignorance’ (Abramovitz 1956), prompting a significant strand of research into improving measurement of inputs to reduce this unexplained growth of output. In particular, Jorgenson & Griliches (1967) showed the importance of improvements in human capital. Much research today still concerns better measurements of inputs, such as intangible capital.

While major efforts of data collection and harmonisation have taken place, modern growth accounting still often starts from a relatively simple decomposition, which we report here. Throughout the paper, we denote real output by Y , the number of workers by L (in practice, KLEMS and OECD STAN use hours worked), the capital stock by K , and we define labour productivity $y = Y/L$ and capital per worker $k = K/L$. The growth rate in real output per unit of labour, $\Delta \log y_t$, can be decomposed as

$$\Delta \log y_t = \Delta \log A_t + (1 - \alpha_t) \Delta \log k_t + \alpha_t \Delta \log h_t, \quad (5.1)$$

where A_t denotes TFP, h_t is an index of the composition of the labour force, and Δ is the first difference operator, i.e. $\Delta x_t \equiv x_t - x_{t-1}$. The labour compensation share of income, α_t , is computed as a Divisia index $\alpha_t \equiv (w_t L_t / Q_t + w_{t-1} L_{t-1} / Q_{t-1}) / 2$, where $Q_t = P_t Y_t$ is nominal output, and w_t is the wage rate per unit of labour.

This decomposition makes it possible to trace the sources of growth, and thus the sources of the slowdown – efficiency (TFP), physical capital, or human capital.

Results Table 5.2.1 reports the decomposition from Eq. 5.1. TFP and capital deepening are the largest contributors. Labour composition appears to contribute only modestly, and only in Germany and the UK². Labour composition is an index of labour services computed assuming that workers in specific gender, age and educational attainment groups are paid their marginal productivity. Because its changes are driven in large part by changes in the relative size of each group, it is unlikely to change very quickly. It is therefore not surprising that it contributes only marginally to the productivity slowdown, compared to capital deepening and TFP.

While the relative contributions of TFP and capital deepening are balanced in the US, we cannot simply assume that this is the case everywhere. In particular, capital deepening is almost the only source of decline in Japan³, and TFP is almost the only source of decline

²For Germany, Elstner et al. (2018) points to labour market reforms of the early 2000’s, which increased the employment of low-skill workers. However, our results for the UK differ from other studies that rely on data from the ONS, the UK’s statistical agency (Goodridge et al. 2018, Riley et al. 2018). KLEMS notes its labour composition index also differs from previous vintages in the case of the UK, pointing to discrepancies in labour survey data managed by the ONS and Eurostat (Stehrer et al. 2019).

³Japan’s TFP growth from KLEMS data shows an improvement post-2005. The literature in this area, cov-

in France. However, in Appendix 5.B, we repeat the exercise using the OECD Productivity database, and find that capital deepening contributed more than 20% in France, and TFP almost 30% in Japan (the OECD Productivity database does not include labour composition). Throughout the paper, we will discuss TFP and capital deepening as the main causes of the slowdown.

Table 5.2.1: Sources of growth decomposition for the slowdown in labour productivity growth pre- and post-2005

		$\Delta \log y_t$	$\Delta \log A_t$	$(1 - \alpha_t) \Delta \log k_t$	$\alpha_t \Delta \log h_t$
<i>France</i>	1996-2005	1.65	1.18	0.16	0.30
	2006-2017	0.66	0.17	0.09	0.40
	Slowdown	0.99	1.01	0.07	-0.09
	Share	1.00	1.02	0.07	-0.10
<i>Germany</i>	1996-2005	1.85	1.10	0.61	0.15
	2006-2017	0.91	0.87	0.07	-0.03
	Slowdown	0.94	0.23	0.54	0.17
	Share	1.00	0.24	0.57	0.18
<i>Japan</i>	1995-2005	1.68	0.29	1.07	0.33
	2006-2015	0.85	0.31	0.26	0.28
	Slowdown	0.82	-0.02	0.80	0.04
	Share	1.00	-0.03	0.98	0.05
<i>United Kingdom</i>	1996-2005	2.21	1.14	0.70	0.37
	2006-2017	0.45	0.30	0.18	-0.02
	Slowdown	1.75	0.84	0.53	0.39
	Share	1.00	0.48	0.30	0.22
<i>United States</i>	1998-2005	2.62	1.37	1.09	0.16
	2006-2017	1.00	0.46	0.38	0.17
	Slowdown	1.61	0.91	0.71	-0.01
	Share	1.00	0.57	0.44	-0.00

Notes: Data from EU-KLEMS 2019.

These results are largely in agreement with previous studies, which use different datasets and slightly different periods. Table 5.2.2 reports the qualitative conclusions from papers relying on growth accounting to explain the productivity slowdown.

Table 5.2.2 synthesises the results of existing growth accounting studies on the recent productivity slowdown. Not all studies use comparable breakdowns in years; many, for example, will compare productivity growth pre- and post-2007, instead of 2005. Not all studies use comparable data on inputs, either: notable differences emerge when calculating the contributions of labour composition or Information and Communication Technology (ICT) capital in isolation. We make an arbitrary judgement on the contribution to the slowdown based on the result of a given paper and a given input, from high (++), modest (+), negligible (0), to worsening (–) the slowdown. When a given input does not feature in the study, we

ered in Jorgenson et al. (2018) and revisited by Baily et al. (2020), emphasises Japan's recovery from its lost decade of growth in the 1990's, followed by TFP levels catching up to the US in the late 2000's.

Table 5.2.2: Proposed sources for the labour productivity growth slowdown from 13 growth accounting studies with diverse data sources

	Data	Labour Composition	Non-ICT Capital	ICT Capital	TFP
<i>France</i>					
van Ark (2016)	TED	+	+	+	++
Cette et al. (2016)	LTPD			0	++
Fernald & Inklaar (2020)	KLEMS 17,12 ^b , PWT ^c	0/–	+/ ^d		++
Gordon & Sayed (2019)	KLEMS 17,12 ^a	–	++	0	++
Oulton (2019)	KLEMS 17		++		++
Inklaar et al. (2019)	KLEMS 17		+		++
<i>Germany</i>					
van Ark (2016)	TED	–	+	–	++
Cette et al. (2016)	LTPD			0	++
Baily et al. (2020)	OECD		++		+
Fernald & Inklaar (2020)	KLEMS 17,12 ^b , PWT ^c	0/–	+/ ^d		++
Gordon & Sayed (2019)	KLEMS 17,12 ^a	–	++	0	++
Oulton (2019)	KLEMS 17		++		++
Inklaar et al. (2019)	KLEMS 17		+		0
<i>Japan</i>					
Baily et al. (2020)	OECD		++		+
Jorgenson et al. (2018)	KLEMS, NSA		++		–
<i>UK</i>					
Riley et al. (2018)	NSA	–	+		++
Goodridge et al. (2018)	NSA	–	+ ^e		++
Tenreyro (2018)	NSA	0		++	++
van Ark (2016)	TED	+	+	+	++
Cette et al. (2016)	LTPD			+	++
Fernald & Inklaar (2020)	KLEMS 17,12 ^b , PWT ^c	0/–	+/ ^d		++
Gordon & Sayed (2019)	KLEMS 17,12 ^a	–	++	0	++
Oulton (2019)	KLEMS 17		++		++
Inklaar et al. (2019)	KLEMS 17		++		++
<i>US</i>					
Baily & Montalbano (2016)	NSA	0	++		++
Murray (2018)	NSA	0	+		++
van Ark (2016)	TED	+	+	+	++
Cette et al. (2016)	LTPD			+	++
Baily et al. (2020)	OECD		+		++
Gordon & Sayed (2019)	KLEMS 17,12	0	++	+	++
Oulton (2019)	KLEMS 17		++		++
Inklaar et al. (2019)	KLEMS 17		++		++

Notes: ^a Aggregated as EU-10 ^b Aggregated as EU-8 ^c Aggregated as EU-15 ^d Calculated as the capital-output ratio ^e A separate intangible capital term yielded a negligible (0) contribution.

Data sources, and their shorthand, are: one or more country-specific national statistical agencies (NSA), Total Economy Database (TED), Long Term Productivity Database (LTPD), Penn World Tables (PWT), OECD Statistics (OECD), and various vintages of EU KLEMS (KLEMS 1X). The contribution of proposed sources to the slowdown are denoted by a symbol; large (++), modest (+), negligible (0), worsening (–). A missing component within a paper is reflected by a blank entry.

leave the entry blank; this means that a study which only considers non-ICT capital growth will have the corresponding entry filled, even though their aggregate capital measure may well include ICT capital, which we can only leave blank. We also record the data used by these various studies, as well as other idiosyncrasies, such as country aggregates that often appear for European countries.

Broadly speaking, this confirms our results in Section 5.2: TFP is the main source of the slowdown, except in Japan, while capital deepening is also important, but labour composition is not found to explain much.

5.2.2 Industry-specific contributions and structural change

A hypothesis for the productivity slowdown is Baumol’s disease. [Baumol \(1967\)](#) theorised that service industries have a lower intrinsic capacity to increase their labour productivity, perhaps because they require in-person services, in contrast to manufacturing. Because similar wage growth would apply to all industries, and because the demand for services such as health and entertainment tends not to decline despite increasing relative cost, these low-productivity growth industries represent an increasing aggregate share of spending, leading to declining aggregate productivity growth.

To examine this hypothesis we ask: do all industries suffer from a productivity slowdown, or is it worse in some industries than in others? Is the slowdown of the aggregate an artifact of the changing relative sizes of industries, with industries with low levels of productivity growth becoming larger?

Our preferred method in measuring structural change is from [Tang & Wang \(2004\)](#) (see also [Nordhaus \(2002\)](#) and [Riley et al. \(2018\)](#)). It decomposes aggregate labour productivity growth into a ‘within’ and a ‘between’ contribution from N industries as

$$\frac{\Delta y_t}{y_{t-1}} = \underbrace{\sum_{i=1}^N \left[\frac{Q_{i,t-1}}{Q_{t-1}} \times \frac{\Delta y_{i,t}}{y_{i,t-1}} \right]}_{\text{Within}} + \underbrace{\sum_{i=1}^N \left[\Delta \left(\frac{P_{i,t} L_{i,t}}{P_t L_t} \right) \times \frac{y_{i,t-1}}{y_{t-1}} \times \left(1 + \frac{\Delta y_{i,t}}{y_{i,t-1}} \right) \right]}_{\text{Between}}. \quad (5.2)$$

The ‘within’ term keeps the relative shares of the sector constant; it is simply the weighted average of industry-level growth rates, with the weights given by the sectors’ nominal shares of total output $Q_{i,t-1}/Q_{t-1}$. The between term, however, contains a factor that reflects structural change, notably the changes in the relative shares of industries, where industry size is in terms of the share of labour resources $L_{i,t}/L_t$ and relative price $P_{i,t}/P_t$. The changes in shares are multiplied by both industry *growth rates* and industry relative productivity *levels*. As noted by [Nordhaus \(2002\)](#) and [Tang & Wang \(2004\)](#), aggregate productivity can still go up even if the growth rates of each industry are zero (i.e. the last factor is equal to 1), because resources are reallocated toward sectors with a high productivity *level*.

Eq. 5.2 measures growth in percentages, rather than log points as in Eq. 5.1, so there will be slight differences in the calculated growth rates and slowdowns in Tables 5.2.1 and 5.2.3.

Results Table 5.2.3 shows the results. We report only key industries, chosen because they are often mentioned in the literature and/or were found to be important in our results. The main takeaway is the strong contribution to the slowdown of manufacturing in all countries. Other substantial contributors are more country-specific: ICT service industries in France and the UK, wholesale and retail trade in Germany, Japan and the US. Financial and Insurance Activities also played a role in the UK and US. The reallocation term is large in Germany, and significant in France, but not elsewhere. Together, these components explain much of the slowdown in each country. The large contribution of ‘Other industries’ in the UK is driven by Real Estate (which is notoriously difficult to measure and dropped in most other studies) and the Oil and Gas industry (Goodridge et al. 2018).

Table 5.2.3: Industry decomposition for the slowdown in labour productivity growth pre- and post-2005

	Total	Manufacturing	Wholesale, Retail and Repair	Financial and Insurance Activities	Information and Communication	Other	Reallocation
<i>France</i>							
1996-2005	1.67	0.65	0.17	0.09	0.23	0.52	0.02
2006-2017	0.67	0.28	0.12	0.03	0.11	0.23	-0.10
Slowdown	1.00	0.37	0.05	0.05	0.11	0.29	0.12
Share	1.00	0.37	0.05	0.05	0.11	0.29	0.12
<i>Germany</i>							
1996-2005	1.87	0.69	0.31	-0.08	0.17	0.45	0.33
2006-2017	0.93	0.46	0.17	0.06	0.18	0.15	-0.08
Slowdown	0.95	0.23	0.15	-0.14	-0.01	0.31	0.41
Share	1.00	0.25	0.16	-0.14	-0.01	0.32	0.43
<i>Japan</i>							
1996-2005	1.74	0.86	0.33	0.15	0.16	0.00	0.24
2006-2015	0.87	0.48	0.07	0.05	0.07	-0.05	0.24
Slowdown	0.87	0.37	0.26	0.09	0.09	0.05	0.01
Share	1.00	0.43	0.30	0.11	0.10	0.06	0.01
<i>United Kingdom</i>							
1996-2005	2.24	0.51	0.16	0.37	0.32	0.64	0.25
2006-2016	0.42	0.12	0.20	0.01	0.07	-0.28	0.29
Slowdown	1.82	0.38	-0.04	0.35	0.25	0.92	-0.05
Share	1.00	0.21	-0.02	0.19	0.14	0.51	-0.02
<i>United States</i>							
1998-2005	2.54	0.96	0.55	0.29	0.50	0.40	-0.16
2006-2017	0.92	0.20	0.11	0.04	0.45	0.31	-0.19
Slowdown	1.61	0.76	0.43	0.26	0.05	0.09	0.02
Share	1.00	0.47	0.27	0.16	0.03	0.06	0.01

Notes: Data from the EU-KLEMS 2019.

Table 5.2.3 is broadly in line with previous work, although methodological and aggregation differences make a systematic comparison more difficult than for the decomposition by factors of production.

The US experience is defined by strong TFP growth pre-2005 in ICT *using* industries, highlighting a point often emphasised by [Gordon \(2016\)](#): productivity growth can be thought of as an adjustment of the levels, with an innovation leading to a new normal level of productivity. [Baily & Montalbano \(2016\)](#), [Cette et al. \(2016\)](#), [Murray \(2018\)](#) and [Baily et al. \(2020\)](#), among others, thus demonstrate that the industries responsible for most of the slowdown in TFP post-2004/05 are those which experienced an acceleration pre-2004/05, namely manufacturing, wholesale and retail trade services, and, to some degree, agriculture. Some studies, such as [Cette et al. \(2016\)](#), [Inklaar et al. \(2019\)](#) and [Baily et al. \(2020\)](#), highlight a strong TFP slowdown in ICT *producing* industries.

[Cette et al. \(2016\)](#), [van Ark \(2016\)](#) and [Gordon & Sayed \(2019\)](#) directly contrast the European experience with that of the US; ICT using industries did not experience much growth pre-2005, and the slowdown in manufacturing is not due to ICT producers specifically. [Inklaar et al. \(2019\)](#) specifically searched for the best industry taxonomy for the productivity slowdown, but, except for a pattern of slowdown in offshoring industries, are largely inconclusive for Europe. In addition to manufacturing, studies of the UK place greater emphasis on financial industries, and also some combination of information and communication services ([Riley et al. 2018](#), [Tenreyro 2018](#)), wholesale and retail trade ([Goodridge et al. 2018](#)), oil and gas ([Goodridge et al. 2018](#), [Riley et al. 2018](#)) and professional, scientific and technical services ([Tenreyro 2018](#)). In all, the slowdown for Europe is more widespread across industries.

In line with our results, the reallocation between industries in France, Japan, UK and US is rarely seen as an important factor ([Byrne et al. 2016](#), [Murray 2018](#), [Tenreyro 2018](#), [Cantner et al. 2018](#), [Nishi 2019](#)), and actually improved labour productivity in the UK ([Goodridge et al. 2018](#), [Riley et al. 2018](#)). However, the strong effect of reallocation that we find for Germany appears missing from the literature, and warrants further research. In line with our results, the literature offers little evidence that Baumol's cost disease is strong enough to explain the productivity slowdown over the fairly short time scales we are considering, although [Nishi \(2019\)](#) and [Duernecker et al. \(2019\)](#) highlight long-term, secular patterns in Japan and the US.

In summary, reallocation fails to explain the pervasive productivity slowdown, which is therefore due to a decline in at least some industries. Indeed, some industries are more affected than others, with manufacturing being a strong contributor to the slowdown due to both its decline in productivity and its relatively large size. High contributions from other industries appear more country-specific, although the evidence suggests that the current slowdown may reflect a pause in the adjustment of productivity towards higher levels initiated by the ICT revolution.

Bibliography

Abramovitz, M. (1956), 'Resource and output trends in the United States since 1870', *The American Economic Review* 46(2), 5–23.

- Baily, M. N., Bosworth, B. P. & Doshi, S. (2020), 'Lessons from productivity comparisons of Germany, Japan, and the United States', *International Productivity Monitor* **38**, 81–103.
- Baily, M. N. & Montalbano, N. (2016), 'Why is US productivity growth so slow? Possible explanations and policy responses', *Brookings Institution, Hutchins Center Working Paper* 22.
- Baumol, W. J. (1967), 'Macroeconomics of unbalanced growth: The anatomy of urban crisis', *American Economic Review* **57**(3), 415–426.
- Bergeaud, A., Cette, G. & Lecat, R. (2016), 'Productivity trends in advanced countries between 1890 and 2012', *Review of Income and Wealth* **62**(3), 420–444.
- Bergeaud, A., Cette, G. & Lecat, R. (2021), Long term productivity database, Dataset v. 2.4, Banque de France. http://www.longtermproductivity.com/download/BCLDatabase_online_v2.4.xlsx (accessed May 2021).
- Byrne, D. M., Fernald, J. G. & Reinsdorf, M. B. (2016), 'Does the United States have a productivity slowdown or a measurement problem?', *Brookings Papers on Economic Activity* **2016**(1), 109–182.
- Cantner, U., Graf, H., Prytkova, E. & Vannuccini, S. (2018), 'The compositional nature of productivity and innovation slowdown', *Jena Economic Research Papers* **2018-006**.
- Cette, G., Fernald, J. & Mojon, B. (2016), 'The pre-Great Recession slowdown in productivity', *European Economic Review* **88**, 3 – 20.
- Duernecker, G., Herrendorf, B. & Valentinyi, A. (2019), 'Structural Change within the Services Sector and the Future of Cost Disease', *CEPR Discussion Paper* **DP12467**.
- Elstner, S., Feld, L. P. & Schmidt, C. M. (2018), The German Productivity Paradox - Facts and Explanations, CESifo Working Paper Series 7231, CESifo.
- Fernald, J. & Inklaar, R. (2020), 'Does disappointing European productivity growth reflect a slowing trend? Weighing the evidence and assessing the future', *International Productivity Monitor* **2020**(38), 104–135.
- Goldin, I., Koutroumpis, P., Lafond, F. & Winkler, J. (2024), 'Why is productivity slowing down?', *Journal of Economic Literature* **62**(1), 196–268.
- Goodridge, P., Haskel, J. & Wallis, G. (2018), 'Accounting for the UK productivity puzzle: A decomposition and predictions', *Economica* **85**(339), 581–605.
- Gordon, R. J. (2016), *The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War*, Princeton University Press.
- Gordon, R. J. & Sayed, H. (2019), The industry anatomy of the transatlantic productivity growth slowdown, Working Paper 25703, National Bureau of Economic Research.

- Inklaar, R., Jäger, K., O'Mahony, M. & van Ark, B. (2019), European productivity in the digital age: evidence from EU KLEMS, *in* B. M. Fraumeni, ed., 'Measuring Economic Growth and Productivity: Foundations, KLEMS Production Models, and Extensions', 1st edn, Elsevier, p. 75–94.
- Jorgenson, D. W. & Griliches, Z. (1967), 'The explanation of productivity change', *The Review of Economic Studies* **34**(3), 249–283.
- Jorgenson, D. W., Nomura, K. & Samuels, J. D. (2018), Progress on measuring the industry origins of the Japan-U.S. productivity gap, Working paper, Harvard University.
- Murray, A. (2018), 'What explains the post-2004 U.S. productivity slowdown?', *International Productivity Monitor* **34**, 81–109.
- Nishi, H. (2019), 'Sources of productivity growth dynamics: Is Japan suffering from Baumol's growth disease?', *Review of Income and Wealth* **65**(3), 592–616.
- Nordhaus, W. D. (2002), 'Productivity growth and the new economy', *Brookings Papers on Economic Activity* **2002**(2), 211–244.
- OECD (2021a), Indicators on trends in business dynamism, Dataset OECD STAN database, February 2021, OECD. <https://stats.oecd.org>, Industries and Services > STAN.
- OECD (2021b), Productivity and ULC by main economic activity (ISIC rev.4), Dataset April 2021, OECD. <https://stats.oecd.org>, Productivity and ULC - Annual, Total Economy.
- Oulton, N. (2019), 'The UK and western productivity puzzle: Does Arthur Lewis hold the key?', *International Productivity Monitor* **36**, 110–141.
- Ramey, V. A. et al. (2020), 'Secular stagnation or technological lull?', *Journal of Policy Modeling* **42**(4), 767–777.
- Riley, R., Rincon-Aznar, A. & Samek, L. (2018), 'Below the Aggregate: A Sectoral Account of the UK Productivity Puzzle', *ESCoE Discussion Paper* **2018-06**.
- Solow, R. M. (1957), 'Technical change and the aggregate production function', *The Review of Economics and Statistics* **39**(3), 312–320.
- Solow, R. M. (1987), 'We'd better watch out', *New York Times Book Review* **36**.
- Stehrer, R., Bykova, A., Jäger, K., Reiter, O. & Schwarzhappel, M. (2019), Industry Level Growth and Productivity Data with Special Focus on Intangible Assets, Report, The Vienna Institute for International Economic Studies.
- Syverson, C. (2017), 'Challenges to mismeasurement explanations for the US productivity slowdown', *Journal of Economic Perspectives* **31**(2), 165–86.

- Tang, J. & Wang, W. (2004), 'Sources of aggregate labour productivity growth in Canada and the United States', *The Canadian Journal of Economics / Revue Canadienne d'Economie* 37(2), 421–444.
- Tenreyro, S. (2018), 'The fall in productivity growth: causes and implications'.
URL: <https://www.bankofengland.co.uk/speech/2018/silvana-tenreyro-2018-peston-lecture>
- The Conference Board (2020), Total Economy DatabaseTM, Dataset July 2020, The Conference Board. <https://www.conference-board.org/data/economydatabase/total-economy-database-productivity>.
- van Ark, B. (2016), Europe's productivity slowdown revisited: A comparative perspective to the United States, in P. Askenazy, L. Bellmann, A. Bryson & E. M. Galbis, eds, 'Productivity Puzzles Across Europe', Oxford Scholarship Online.
- Vollrath, D. (2020), *Fully Grown: why a stagnant economy is a sign of success*, University of Chicago Press.

Appendix

5.A Slowdown and missing GDP

In the spirit of [Syverson \(2017\)](#), to highlight the consequences of slowing productivity, we compute what GDP would have been, should the productivity slowdown not have occurred. Formally, labour productivity growth between two periods t and $t - 1$ is

$$g_t \equiv \Delta \log y_t = \log \frac{Y_t}{L_t} - \log \frac{Y_{t-1}}{L_{t-1}},$$

where Y denotes real value added (labelled VA_Q by KLEMS), L units of labour, in our case hours worked (H_EMP), and Δ is the first difference operator. KLEMS 2019 provides the labour productivity growth rates (LP1_G) directly for our five countries, although certain years may be missing in specific instances. Since the data for (Y/L) generally starts in 1995, the first observed growth rate in France, Germany and the United Kingdom is in 1996. However, data on productivity growth rates for the United States only starts from 1998, while in Japan there is an observation for 1995. The data for labour productivity growth extend to 2017, with another exception for Japan, for which data end in 2015.

With these caveats in mind, we compute labour productivity growth for our five countries by averaging the growth rates across all years in the base period (0), 1996–2005, using

$$g_{(0)} \equiv \frac{1}{T_{(0)}} \sum_{t=1996}^{2005} g_t, \quad (5.3)$$

where $T_{(0)} = 2005 - 1996 + 1 = 10$. These values are listed in the first column of Table 5.1.1, multiplied by 100 to denote percentage points (pp).

For the period 2006–2017 (or 2015 in the case of Japan), denoted (1), this *realised* average rate of labour productivity growth is defined as

$$g_{(1)} \equiv \frac{1}{T_{(1)}} \sum_{t=2006}^{2017} g_t, \quad (5.4)$$

where $T_{(1)} = 2017 - 2006 + 1 = 12$. In the third column, we compute the slowdown in labour productivity growth as the difference between the two average growth rates,

$$\text{Slowdown} \equiv g_{(0)} - g_{(1)}.$$

In column four, we calculate GDP per capita in 2017 (2015 for Japan) as

$$\text{GDP per capita}_{2017} \equiv \frac{Y_{2017}}{N_{2017}}, \quad (5.5)$$

where N_{2017} is the mid-year population count taken from the Conference Board's Total Economy DatabaseTM ([The Conference Board 2020](#)) for each country⁴.

To calculate 'missing GDP', column five, we write a counterfactual GDP per capita in 2017 in terms of a counterfactual level of labour productivity,

$$\frac{\tilde{Y}_{2017}}{N_{2017}} = \frac{Y_{2017}}{N_{2017}} \frac{\tilde{Y}_{2017}}{Y_{2017}} = \frac{Y_{2017}}{N_{2017}} \frac{\tilde{y}_{2017}}{y_{2017}}, \quad (5.6)$$

using a tilde to denote a variable's 'counterfactual' and because $\tilde{L} = L$. The labour productivity levels were the same until 2005, so

$$\frac{\tilde{y}_{2017}}{y_{2017}} = \frac{y_{2005} \exp(g_{(0)} T_{(1)})}{y_{2005} \exp(g_{(1)} T_{(1)})} = \exp((g_{(0)} - g_{(1)}) T_{(1)}),$$

since $\tilde{g}_{(1)} = g_{(0)}$ is the growth rate that would have happened had labour productivity growth continued on its 1996–2005 trend after 2005. Substituting this back into Eq. 5.6, we have

$$\frac{\tilde{Y}_{2017}}{N_{2017}} = \frac{Y_{2017}}{N_{2017}} \exp((g_{(0)} - g_{(1)}) T_{(1)}), \quad (5.7)$$

Taking stock, we compute the missing GDP in 2017 (again, 2015 for Japan) as

$$\text{Missing GDP per capita}_{2017} \equiv \frac{\tilde{Y}_{2017}}{N_{2017}} - \frac{Y_{2017}}{N_{2017}} = \frac{Y_{2017}}{N_{2017}} \left[\exp((g_{(0)} - g_{(1)}) T_{(1)}) - 1 \right],$$

using Eq 5.7, and the values defined in Eqs 5.3, 5.4 and 5.5. Finally, note that volume indices for real GDP are indexed to 2010 prices in KLEMS. We convert GDP and missing GDP per capita to 2017 prices by multiplying by the price index (VA.PI) for 2017, divided by 100, or equivalently by taking directly nominal values (VA), instead of volumes in 2010 prices. The results are in the fifth column of Table 5.1.1.

In principle, for the purpose of constructing Table 5.1.1, using data from the Total Economy DatabaseTM would have been better due to its better time coverage. We have constructed such a table and it does not change our narrative very much. The slowdown (2006–2017 compared to 1996–2005, for all countries) ranges from 0.68 (Germany) to 1.77 (UK). We prefer to use EU KLEMS 2019 throughout the paper for consistency.

Table 5.B.1: Growth accounting results using OECD Productivity data

		$\Delta \log y_t$	$\Delta \log A_t$	$(1 - \alpha_t) \Delta \log k_t$
<i>France</i>	1995-2005	1.74	0.96	0.77
	2006-2017	0.71	0.15	0.55
	Slowdown	1.03	0.80	0.22
	Share	1.00	0.78	0.21
<i>Germany</i>	1995-2005	1.54	0.79	0.74
	2006-2017	0.87	0.63	0.23
	Slowdown	0.68	0.17	0.51
	Share	1.00	0.24	0.76
<i>Japan</i>	1995-2005	2.11	0.84	1.25
	2006-2017	0.75	0.48	0.27
	Slowdown	1.36	0.36	0.98
	Share	1.00	0.27	0.72
<i>United Kingdom</i>	1995-2005	2.22	1.73	0.45
	2006-2017	0.47	0.09	0.37
	Slowdown	1.75	1.64	0.08
	Share	1.00	0.94	0.05
<i>United States</i>	1995-2005	2.27	1.36	0.88
	2006-2017	1.06	0.46	0.59
	Slowdown	1.21	0.90	0.29
	Share	1.00	0.75	0.24

5.B Contributions of TFP and capital deepening using OECD Productivity database

Table 5.B.1 shows the results. The most noticeable difference with EU KLEMS is the substantially smaller slowdown of capital deepening for the US. The OECD data (OECD 2021b) slightly mitigates the result from KLEMS that the source of the slowdown is only TFP in France and only capital deepening in Japan.

5.C Contributions of industries and reallocation using OECD's STAN

As a robustness check for the industry level decomposition in labour productivity growth, we reproduce the decomposition using data from the OECD's STAN database (OECD 2021a), Table 5.C.1. The downside is that hours worked data for Japan are missing, and the productivity series generally do not extend beyond 2015. Despite these shortcomings, results from the industry-level decomposition, using the same method of Tang & Wang (2004), are almost identical to those derived from the KLEMS 2019 data, visible in Table 5.2.3.

⁴<https://conference-board.org/data/economydatabase>, version July 2020.

Table 5.C.1: Industry decomposition for the slowdown in labour productivity growth with OECD STAN data

	Total	Manufacturing	Wholesale, Retail and Repair	Financial and Insurance Activities	Information and Communication	Other	Reallocation
<i>France</i>							
1996-2005	1.61	0.62	0.16	0.09	0.23	0.49	0.02
2006-2015	0.67	0.27	0.09	0.06	0.11	0.21	-0.08
Slowdown	0.94	0.35	0.07	0.03	0.11	0.28	0.10
Share	1.00	0.37	0.07	0.03	0.12	0.30	0.11
<i>Germany</i>							
1996-2005	1.87	0.69	0.31	-0.08	0.17	0.45	0.33
2006-2015	0.87	0.39	0.16	0.06	0.21	0.14	-0.09
Slowdown	1.01	0.30	0.15	-0.13	-0.04	0.31	0.42
Share	1.00	0.30	0.15	-0.13	-0.04	0.30	0.42
<i>United Kingdom</i>							
1996-2005	2.18	0.50	0.17	0.27	0.31	0.66	0.27
2006-2015	0.48	0.14	0.18	0.00	0.08	-0.18	0.26
Slowdown	1.70	0.36	-0.01	0.27	0.23	0.84	0.01
Share	1.00	0.21	-0.01	0.16	0.14	0.50	0.01
<i>United States</i>							
1996-2005	2.36	0.91	0.56	0.32	0.27	0.40	-0.10
2006-2015	0.97	0.25	0.08	0.09	0.22	0.45	-0.11
Slowdown	1.39	0.66	0.49	0.24	0.05	-0.05	0.01
Share	1.00	0.47	0.35	0.17	0.03	-0.03	0.01

Chapter 6

Accounting for variety

Abstract

I introduce a general method to account for the distribution of underlying components (variety) in the growth of an aggregate quantity, using the notion of entropy. The cross-entropy of GDP with respect to a benchmark captures the change in its distribution, and thus how well this benchmark matches data for price and volume indices across time. This ‘error’ changes demonstrably over time. Accounting of variety also lends itself to a decomposition of labour productivity growth by a technology component, and the allocation of labour plus demand.

6.1 Introduction

Is Gross Domestic Product (GDP) a good measure for welfare? Real output of goods and services is the accepted measure of historical living standards. To track its evolution, statisticians count growth in expenditures, also ‘nominal’ output, hoping to separate growth in price (inflation) versus real output:

$$\text{Real output growth} = \text{Nominal output growth} - \text{Price Inflation.}$$

Nominal output is observed in currency units, but *what* exactly should we measure the price, and real output, of? Choosing a measure for price requires aggregating prices from hundreds of industries, millions of firms, and many more product types. In abstract terms, ‘what’ we pay for at any given point in time is in the distribution of nominal output among a number of ‘types’ of output: industries, products, firms, etc.

This paper notes that even if the question of ‘what’ we want to measure is settled, ‘what’ we pay for changes over time. For example, we know the distribution of US nominal output in 2000 to rely more on service and IT industries compared to 1950. If we are interested in measuring historical real output growth from the perspective of the average firm in 2000, more weight should be placed on inflation for service and IT industries in the 1950’s, even if they constituted a smaller share of nominal output. The result is a worse approximation of real output growth for the average firm in 1950.

It follows that the importance placed on certain goods should change with the distribution of nominal output as preferences evolve. However, in a recent study, [Baqaee & Burstein \(2021\)](#) insist on fixing weights, to the extent that they change due to income effects or taste shocks, when measuring economic welfare. The important contribution is their challenge to what we should interpret as a historically consistent, and intuitive, measure for living standards. The composition of expenditures they seek to deflate may not agree with the weights assigned to price components, which this paper shows to hold significant practical implications.

One illustrative thought experiment demonstrating the importance of a changing output distribution at any level of aggregation is Hulten’s Paradox. The ICT revolution in the 1990’s sparked a discussion on measuring welfare improvements from the changing quality of products, in addition to their real output. Besides practical difficulties in implementing such ‘hedonic’ price indices, [Hulten \(1997\)](#), in response to [Nordhaus’s \(1997\)](#) claim that the hedonic price of light dropped faster than its traditional counterpart, famously cautions the use of quality adjustments since they likely overstate actual welfare gains. A passage from [Hulten’s \(1997\)](#) comment, evaluating [Nordhaus’s \(1997\)](#) revised price deflator, provides the key insight:

[...]a person possessing the average disposable income in America today should be willing to accept a massive reduction in spending power – from \$17,200 to the \$90–430 range – in order to avoid being sent back in time to an equivalent status in colonial America. Alternatively, it suggests that the average colonial should

prefer living in the America of today, with as little as \$90 per year, to staying put in the late eighteenth century.

The contentious point in Hulten's quote is that a basket of goods for the colonial American evolved very differently from a representative basket of goods consumed today. The \$90 today prices a different 'representative' unit than \$90 in colonial America. An individual today would have indeed turned out quite poor in colonial America, if the representative basket of goods was substantially more scarce compared to the basket of a colonial American. Sending him back in time requires centuries of devolving preferences and substitutions. This substitution also entails a scarcer basket of goods for a colonial American, today.

Section 6.2 formalises this issue using index number theory, and introduces an extra entropy term, alongside price and real output, which encapsulates the shift in consumption patterns. Entropy is used in many fields as a measure of 'observational variety', and therefore lends an elegant application in the present economic context. The result emerges naturally from Jensen's Inequality, but also benefits from deep intuition; indices aggregate across a variety, so a change in variety hurts the 'representativeness' of an index.

The entropy term captures a bias in real output when extrapolated from the difference in nominal output and price inflation. Furthermore, this bias travels in one direction with time, suggesting that the representativeness of a basket of goods from the past should worsen, whereas it improves for today's basket. In practice, this means any pair of price and real output indices that favour recent consumption bundles will overestimate the price and real output growth rates of older bundles. This finding can be used to evaluate the bias in popular indexation methods. For example, Divisia indices update product weights in each period, which achieves a good result in terms of minimising the influence of the cross-entropy bias.

This result applies to any aggregate index. One application is in measuring the reallocation effect studied in the growth accounting field, whereby resources move between high and low productivity industries. In this context, the entropy of employment by industry contributes to a labour productivity decomposition, to the degree that the allocation of employment diverges from sources of production.

Section 6.3 measures the bias from cross-entropy in two applications. I test the effect in a transparent manner, using industry price and output data from US KLEMS for the period 1947–2014. Forming production profiles from 65 industries, the real output growth rate experienced by the 1947 bundle is over 1pp smaller than that experienced by the 2014 bundle. Without correcting for cross-entropy, this means that we would overstate today's living standards by around 70% compared to those of 50 years ago when using consumption patterns from today, as would be the case if data on industry-level real output was unavailable. Overall, the industries that made up larger shares of nominal output in 1947 experienced both lower inflation and real output growth rates. It follows that nominal output grew more uniform between industries, as reflected by an increasing rate of entropy. A popular view is that the economic structure of the US shifted out of agriculture and into professional or health services, due to some type of 'cost disease'.

Applying the notion of entropy to labour productivity is instructive for the wider applicability of entropy to economic questions. I extend the output exercise by including a labour index, which yields a labour productivity decomposition between i) the productivity growth rate experienced by the average industry, ii) the divergence of expenditures from ‘important’ industries, and iii) the convergence of the distribution, by industry, of labour inputs to the importance of any given industry assigned by the weighting scheme. Using EU KLEMS data, I find that a significant part of the post-2005 slowdown in labour productivity growth in Germany can be attributed to a slowdown in the allocation of labour inputs.

Related literature A number of key arguments characterise the persistent debate on what real output should measure. Complications arise when new products enter, or old products leave, the market: [Aghion et al. \(2019\)](#) estimate that imputations for inflation of products subject to creative destruction leads to an overestimation of the true inflation rate by an average of 0.5%. Most of this stems from hotel, restaurant and retail trade industries. There exists an active literature that seeks to interpret changes in product variety in terms of welfare ([Redding & Weinstein 2019](#), [Baqaee & Burstein 2021](#)). [Deaton \(2010\)](#) famously re-draws the international poverty line to address differences in product variety across countries. The accounting framework I present is complementary to those efforts, but extends beyond the scope of output and welfare measurements.

Revisiting Hulten’s Paradox for a moment, [Gordon \(2009\)](#) forces a point on Consumer Price Index (CPI) measurement, demonstrating that US women’s apparel products experienced higher rates of inflation under a hedonic price model. To be clear: the entropy problem is distinct, in the sense that real output growth rates may appear unreasonably high regardless of downward biases in aggregate deflators. The cross-entropy term behaves in opposing directions for aggregate prices weighted to the start of the period, from which consumption baskets diverge, versus the end, towards which consumption baskets converge. In this way, changing variety *solves* Hulten’s paradox, instead of correcting it.

Information Theory concepts like entropy feature prominently in econometrics. [Maa-soumi \(1993\)](#) offers a broader review, and explains their usefulness in studying the role of aggregation. This was an area of concern to [Theil \(1967\)](#) in his study of income inequality across population groups. In contrast, this paper focuses on distributional divergence over time, instead of a cross-sectional dimension.

Roadmap Section 6.2 presents the decomposition, and explains the role of entropy in the present context. Section 6.3 applies this decomposition to US KLEMS data to demonstrate the scale of the cross-entropy problem, and extends it to study the aggregate productivity growth slowdown. Section 6.4 concludes.

6.2 Indexing variety

In this section, I develop a statistical accounting framework to aggregate nominal output – GDP – although it applies to any index that seeks to aggregate across a variety of types:

employment, exchange rates, monetary aggregates, etc. The role of variety emerges statistically as the *entropy* of the index, which theoretically evolves alongside incomes as long as preferences are not specifically Cobb-Douglas.

6.2.1 Entropy to measure variety

National statistical agencies measure GDP and its composition between industries. The relationship between aggregate and industry nominal outputs, themselves the products of industry prices times real outputs, adheres to Definition 6.1.

Definition 6.1 (GDP). *GDP, Y , is the sum of N industry nominal outputs Y_i , which themselves are equal to the product of the industry-specific price level P_i and real output level Q_i :*

$$Y = \sum_{i=1}^N Y_i = \sum_{i=1}^N P_i Q_i. \quad (6.1)$$

Note that Definition 6.1 implies the existence of industry-level deflators P_i , which simplifies the exposition. When dealing with long-term economic data, statisticians typically observe prices from consumer surveys, then extrapolate an aggregate real output index from aggregate GDP. I make two assumptions throughout this paper. First, the number of types (industries) N is large. Second, GDP and industry prices P_i are always known.¹ Industry-specific nominal or real outputs are not necessarily observed.²

To track historical living standards, statisticians look to retrieve an index that aggregates real outputs across industries over time. Since these are unobserved, the next best thing is to construct an aggregate price index with which to deflate GDP. Definition 6.2 states the type of index studied throughout this paper.

Definition 6.2 (Price and real output indices). *The price(real output) index equals a geometric weighted average of industry prices(real outputs),*

$$\tilde{P}_t = \Delta \sum_{i=1}^N \omega_{i,t} \log P_{i,t}, \quad \tilde{Q}_t = \Delta \sum_{i=1}^N \omega_{i,t} \log Q_{i,t}, \quad (6.2)$$

where t denotes time, ω_i denotes a positive weight assigned to industry i such that $\sum_i \omega_i = 1$, and Δ is the difference operator, $\Delta x \equiv x_t - x_{t-1}$.

Industry weights ω_i belong to a $1 \times N$ vector $\mathbf{\Omega} = \{\omega_1, \omega_2, \dots, \omega_N\}$, which I refer to as an *indexation scheme*. This is distinct from the distribution of industry nominal output shares

¹One complication, especially for longer time series, is that surveyed prices are often consumer prices, not producer prices. One symptom of this problem is that a price index used to deflate *domestic* output may include import prices. While this challenge is beyond the scope of the present paper, the assumption offers a best-case scenario in measuring historical living standards.

²Historically, the breakdown of GDP by industry is not always available, for two main reasons; i) tracking expenditures for granular product classifications, which can include hundreds of product types, is a significant practical challenge, and ii) data attrition increases the longer the time series, even for coarser industry aggregates.

$y_i = Y_i/Y$, populating a $1 \times N$ vector $\mathbf{Y} = \{y_1, y_2, \dots, y_N\}$, which I refer to as *consumption patterns*.

An important remark is that Definition 6.2 restricts the scope of the paper to indices constructed from geometric averages. Examples are the Törnqvist, Sato-Vartia and Divisia indices, but alternatives exist, notably the Fisher, Laspeyres and Paasche indices (Törnqvist 1936, Sato 1976, Vartia 1976, Fisher 1922). A sprawling literature, duly explained by Eichhorn (1978), proposes sets of tests to evaluate which index is ‘best’. The main contender to the type of index introduced by Definition 6.2 is Fisher’s ideal index.³

Indices of the type in Definition 6.2 produce a residual, because

$$\tilde{P}_t + \tilde{Q}_t = \Delta \log Y_t \iff \Delta \log P_{i,t} = \Delta \log P_{j,t} \text{ and } \Delta \log Q_{i,t} = \Delta \log Q_{j,t}, \quad \forall i, j. \quad (6.3)$$

Unless price and quantity growth is exactly equal for each category i , the resulting indices will not sum to GDP growth. The key contribution of the present paper is to give this residual a name. This is desirable, because it yields an exact decomposition of GDP in an economically meaningful way, without resorting to assumptions required by complicated demand systems. Proposition 6.1 introduces this decomposition, applied to GDP growth.

Proposition 6.1 (Accounting for variety). *GDP can be decomposed exactly between a price index, real output index, and the cross-entropy of consumption patterns relative to the indexation scheme:*

$$\Delta \log Y_t = \underbrace{\Delta \sum_{i=1}^N \omega_{i,t} \log P_{i,t}}_{\text{Price inflation: } \tilde{P}_t} + \underbrace{\Delta \sum_{i=1}^N \omega_{i,t} \log Q_{i,t}}_{\text{Real output growth: } \tilde{Q}_t} + \underbrace{\Delta \sum_{i=1}^N \omega_{i,t} \log \left(\frac{Y_t}{Y_{i,t}} \right)}_{\text{Entropy change: } \Delta H(\mathbf{\Omega}_t, \mathbf{Y}_t)}. \quad (6.4)$$

Extrapolated output growth: $\tilde{X}_t$

Proof. The difference between GDP and an index of industry nominal outputs, following indexation scheme $\mathbf{\Omega}_t$, is

$$\log Y_t - \sum_{i=1}^N \omega_{i,t} \log Y_{i,t} = \sum_{i=1}^N \omega_{i,t} \log \left(\frac{Y_t}{Y_{i,t}} \right), \quad (6.5)$$

which is the cross-entropy of the indexation scheme $\mathbf{\Omega}_t$ and the consumption pattern $\mathbf{Y}_t$. Following convention, this term is written as $H(\mathbf{\Omega}_t, \mathbf{Y}_t)$. Substituting $\log Y_{i,t} = \log P_{i,t} + \log Q_{i,t}$, which follows from Definition 6.1, and re-arranging yields the desired result. $\square$

There are two immediate observations on the residual, termed ‘cross-entropy’, in Proposition 6.1. First, cross-entropy is identical to the negative log-likelihood of the nominal out-

³I briefly summarise Eichhorn (1978) on the trade-off between Fisher-type indices and Törnqvist-type indices, and its relevance to the present context. Fisher-type indices fail a *Base Test*, in that the total inflation rate between two periods is not equal to the sum of inflation rates in enclosed sub-periods, whereas Törnqvist-type indices fail a *Product Test*, in that the aggregate price and real output indices do not sum to the GDP index.

put data, $\mathbf{Y}_t$, relative to $\mathbf{\Omega}_t$. Therefore, minimising cross-entropy is similar to maximising the log-likelihood of all $Y_{i,t}$, given that we assume weights $\omega_{i,t}$. The second interpretation is that of entropy itself: cross-entropy quantifies the information required to encode Y_t , given that the encoding scheme is optimised for $\mathbf{\Omega}_t$. It is, loosely, a measure of the variation in $\mathbf{Y}_t$ explained by the variation in $\mathbf{\Omega}_t$.

In practice, Proposition 6.1 suggests that statisticians have some freedom in choosing an indexation scheme that is economically meaningful, in that they produce a model $\mathbf{\Omega}_t$ to explain variation in expenditures across industries $\mathbf{Y}_t$. This comes at the price of some information loss, to the extent that the model does not reproduce true consumption patterns.

The failure in indices of the type in Definition 6.2 for the product test in Eq. 6.3 follows from Proposition 6.1. Corollary 6.1.1 finds that no such index can perfectly satisfy the product test: there always remains a positive residual from cross-entropy, which is bounded by the entropy of aggregate GDP.

Corollary 6.1.1. *The choice of indexation scheme $\mathbf{\Omega}_t$ which minimises $H(\mathbf{\Omega}_t, \mathbf{Y}_t)$ is*

$$\omega_{i,t} = \frac{Y_{i,t}}{Y_t} \quad \forall i. \quad (6.6)$$

Under this indexation scheme, minimal cross-entropy can be written as

$$H(\mathbf{\Omega}_t, \mathbf{Y}_t) = \sum_{i=1}^N \frac{Y_{i,t}}{Y_t} \log \left(\frac{Y_t}{Y_{i,t}} \right) = H(\mathbf{Y}_t), \quad (6.7)$$

which is Shannon's (1948) definition of entropy, applied to aggregate nominal output Y_t .

Proof. Follows from the definition of cross-entropy (Cover & Thomas 2005). □

From Corollary 6.1.1, the cross-entropy residual can be minimised by setting the indexation scheme equal to the distribution of GDP. The existence of any residual is immediate from Eq. 6.5, which is a standard representation of Jensen's Inequality. When the indexation scheme is chosen to be equal to the distribution of GDP among industries, the cross-entropy term reduces to the entropy of GDP, $H(\mathbf{Y}_t)$, as defined by Shannon (1948). This index is typically referred to as the Divisia index when applied to continuous time.

The surprising implication of Corollary 6.1.1 is that the fidelity of any pair of price and real output indices to the product test has a hard boundary equal to the entropy of GDP. This has particular consequences for long time series, where the entropy of GDP itself might change over time. This point is developed further below.

6.2.2 Intermezzo: what is entropy?

Here, I illustrate what the entropy of GDP measures. Formally, recall that the market economy is accounted for by the transactions across N goods in a given period of time, for which the sum is GDP. The number of possible arrangements for those transactions is given by the

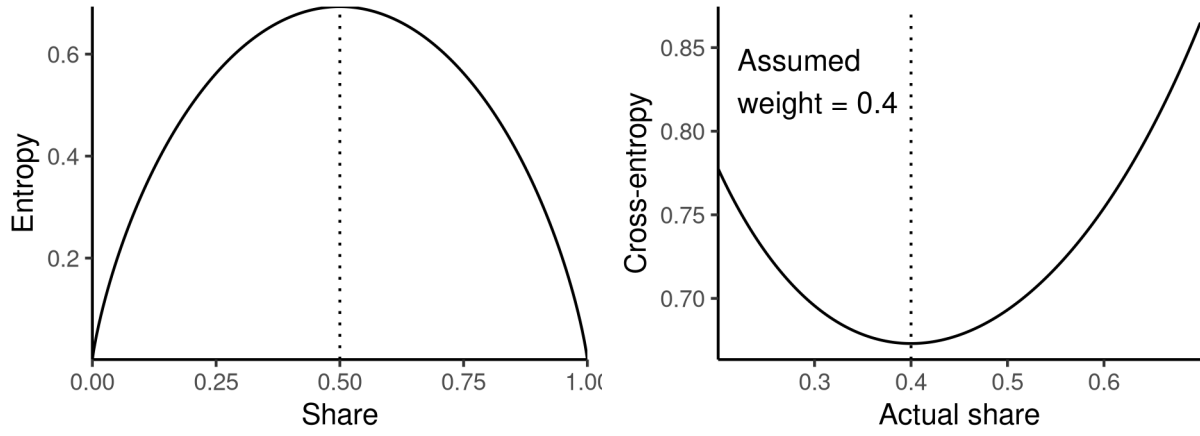


Figure 6.2.1: **Plotting entropy for two industries:** Total entropy (Left) is maximal at the dotted line, where the share of each industry is equal at 1/2. By fixing a weight (Right) at the dotted line, a change in the actual share increases cross-entropy.

multinomial coefficient:

$$\binom{Y!}{Y_1!, Y_2!, \dots, Y_N!} = \frac{Y!}{Y_1! Y_2! \dots Y_N!} = \frac{Y!}{(Y y_1)! (Y y_2)! \dots (Y y_N)!},$$

where $y_i = Y_i/Y$. This is simplified, under large N , using the Stirling approximation for factorials:

$$\log \left[\frac{Y!}{(Y y_1)! (Y y_2)! \dots (Y y_N)!} \right] \approx Y \sum_{i=1}^N y_i \log \left(\frac{1}{y_i} \right), \quad (6.8)$$

which corresponds to [Shannon's \(1948\)](#) entropy. To better interpret this term, assume that, on a given day, €10 of fruit are traded on a market consisting of two types, apples and coconuts. If spending is uniform, the number of possible allocations is

$$\frac{10!}{5!5!} = 252.$$

On the other extreme, if all but one Euro is spent on one fruit alone, the number of possible arrangements is

$$\frac{10!}{9!1!} = 10.$$

This is why entropy is often described as a measure of uncertainty, or *observational variety*: from the consumer's perspective, there are many more ways to spend an endowment equally among options, rather than prioritising some ahead of others.⁴ An additional Euro has many possible ways of being spent in a uniform economy, but fewer in a concentrated economy.

Returning to the problem at hand, a weighting scheme $\mathbf{\Omega} \stackrel{d}{=} \mathbf{Y}$ yields an indexation error

⁴Two technical remarks: i) observational zeros do not appear ($0 \log 0 = 0$), and ii) each Euro is treated equally.

equal to the entropy of GDP. Therefore, the indexation error *worsens* as variety *increases*. In the example, a real output index will be more inaccurate when expenditures are similar between apples and coconuts, rather than concentrated in apples. This is visualised on the left of Figure 6.2.1: if all expenditures are on apples, then the growth rate of apples produced is a perfect measure for aggregate growth, and entropy is zero. Alternatively, if expenditures are allocated uniformly, the average growth rate is a worse approximation of the growth rates in either apples or coconuts. When national statisticians extrapolate an aggregate real output index from a time series of industry GDP and prices, it will invariably be progressively *more* biased *upwards* if variety in GDP increases. This is particularly interesting for fixed weights, visualised on the right of Figure 6.2.1; the cross-entropy between the assumed weight versus the actual share can change over time, even if entropy itself is already low.

I write ‘economic’ variety because this result naturally extends beyond industry-level aggregation, right down to the product level. It is not hard to imagine that variety today is orders of magnitude larger than a century ago. In fact, much of this is reflected in the constant revisions and expansions of industry classifiers: what products are similar enough that we can assume a common price level? The answer inevitably evolves with innovation and creative destruction, as per [Aghion et al. \(2019\)](#).

Comparing distributions Entropy is a metric for variety in one variable, but cross-entropy combines this measure for two distributions. The notion of ‘accounting’ for variety I propose is no more than a log-likelihood function for the distribution of GDP by industry against an assumed weighting scheme. In economics, these weights typically constitute a ‘model’. Proposition 6.1 suggests that an accurate measure for aggregate real output requires a ‘good’ model when quantities are not directly observed by industry.

The right panel of Figure 6.2.1 demonstrates the extent to which variation in the true share of an industry, plotted along the horizontal axis, affects cross-entropy as measured by Eq. 6.4 when the assumed weight is fixed to 0.4. One issue with cross-entropy is that it is difficult to compare across different scenarios, because it is jointly sensitive to the performance of the model in addition to variety in the original data. Therefore, the *Kullback-Leibler* (KL) divergence is used as a common metric that standardises the divergence between two distributions ([Kullback & Leibler 1951](#)):

$$\begin{aligned} D_{KL}(\mathbf{\Omega}||\mathbf{Y}) &= H(\mathbf{\Omega}, \mathbf{Y}) - H(\mathbf{\Omega}) \\ &= \sum_i^N \omega_i \log \frac{\omega_i}{y_i}, \end{aligned}$$

where $D_{KL}(x||y)$ denotes the KL divergence of y from x . The desirable property of the KL divergence between two distributions is that it is zero when both distributions are identical. This is also seen from the minimised cross-entropy in Eq. 6.7. Further properties of the KL divergence measure are discussed by [Cover & Thomas \(2005\)](#).

6.2.3 Theoretical motivation

So far, I motivated the integration of entropy into macroeconomic indices by an intriguing accounting relationship. This comes with its own downside, since the result in Proposition 6.1 might strike the reader as tautological. Uneven growth rates must cause relative expenditure shares to diverge. However, theoretical assumptions also yield a variable level of entropy, except under Cobb-Douglas aggregation. Entropy may therefore offer new insights into the nature of heterogeneous consumption in a demand system. I demonstrate this using a general framework.

Modelling consumption shares A parametric model can rationalise the share of consumption allocated to a given industry. Using such models, economists build ‘cost of living’ indices that aggregate industry prices to meet their benchmark ‘standard of living’, the corresponding aggregate real output index. Various assumptions can yield different parameterisations, and thus different weighting schemes by which to aggregate across industries. The main parameter that governs aggregation in these models is

$$\varepsilon_i = \frac{dY}{dQ_i} \frac{Q_i}{Y} \quad (6.9)$$

$$= \frac{Y_i}{Y} = y_i. \quad (6.10)$$

Parameter ε_i in Eq. 6.9 is defined as the *elasticity* of output Y with respect to quantity Q_i . Eq. 6.10 follows from assuming that Q_i is paid price P_i equivalent to its marginal output (the derivative in Eq. 6.9). This relationship is expected to hold at the system’s equilibrium under perfect competition. Further extensions consider markups charged under imperfect competition, and other sources for ‘wedges’ (Baqee & Farhi 2020).

Parameterisation using a utility function This elasticity features in a functional relationship between aggregate welfare and quantity inputs: a ‘utility function’ (or its dual with respect to prices, the ‘cost function’). Writing this parameter explicitly is useful because it clearly delineates the realm of variability that these models offer under perfect competition. This is crucial for understanding the theoretical underpinning of entropy, which emerges from re-arranging Eq. 6.10:

$$\log Y = \log P_i + \log Q_i + \log \frac{1}{\varepsilon_i} \quad (6.11)$$

$$= \sum_{i=1}^N \omega_i \log P_i + \sum_{i=1}^N \omega_i \log Q_i + \sum_{i=1}^N \omega_i \log \frac{1}{\varepsilon_i}, \quad (6.12)$$

where $\sum_i \omega_i = 1$ as before. Eq. 6.11 demonstrates the earlier tautology: without measurement error, an index can be built from a single industry alone only if quantity growth is equal among all industries. In other words, aggregated output elasticities are residuals that only change when quantity growth is uneven among industries. The aggregation in Eq. 6.12 is desirable because it tracks ‘aggregate’ prices and quantities according to a scheme Ω ,

whose purpose can either be statistical – attenuating measurement error – or economic – tracing some aggregate ‘welfare’.

When is entropy constant? It is important to note that different utility aggregators yield various predictions for what the elasticity should be. Therefore, changes in the output distribution can be interpreted in a variety of ways, depending on the assumptions on the utility function. Proposition 6.2 states a null result: what utility functions predict constant elasticities, and, consequently, a fixed output distribution?

Proposition 6.2 (Constant elasticity). *Aggregator $Y = f(Q_i)$ exhibits a constant elasticity ε_i with respect to quantity Q_i if, and only if, it holds functional form*

$$f(Q_i) = Q_i^{\varepsilon_i} c, \quad (6.13)$$

where c is constant with respect to Q_i .

Proof. See Appendix 6.A.1. □

Utility functions that follow the form of Eq. 6.13 are commonly referred to as ‘Cobb-Douglas’. According to Proposition 6.2, Cobb-Douglas aggregators have constant elasticities that are dependent on a variety of goods, resulting in a constant entropy for the corresponding quantity index as income levels change. Preferences that generate constant expenditure shares as a function of income levels are known as *homothetic* preferences. This has significant implications, as it highlights that GDP alone is not a sufficient welfare metric unless preferences are Cobb-Douglas. Other utility functions require real output indices that are adapted to incorporate changes in preferences.

A prime example of such an adaptation is Baqaee & Burstein (2021). Their analysis considers the extent to which particular industries become more prevalent in consumption as incomes increase. Practical scenarios include healthcare and professional services, which grew as a share of GDP in advanced economies, in contrast to agricultural and certain industrial services, which generally shrank. Redding & Weinstein (2019) also introduce a demand system to assign an entropy-like residual, specifically for the Sato-Vartia index, some economic insight.

Does entropy increase over time? One reflection follows an established literature on the evolution of human wants as they progressively satisfy their needs (Maslow 1943). As spending on relatively niche outlets increases, and allocations turn more uniform, entropy invariably increases. More concretely, certain utility forms predict output elasticities as a function of input levels, allowing for some flexibility in assessing changes in output variety. Table 6.2.1 tabulates these elasticities for three popular functions used in microeconomic theory, with two inputs Q_1, Q_2 . Section 6.3 offers empirical evidence for a broad trend in the historical entropy rate of the US economy.

Table 6.2.1: Demand elasticities for select utility functions

Name	$f(Q_1, Q_2)$	ε_1
Cobb-Douglas	$AQ_1^{a_1}Q_2^{a_2}$	a_1
CES	$A[a_1Q_1^\sigma + a_2Q_2^\sigma]^\frac{1}{\sigma}$	$\frac{a_1Q_1^\sigma}{a_1Q_1^\sigma + a_2Q_2^\sigma}$
Translog	$A + \sum_i a_i \log Q_i + \sum_{i,j} a_{i,j} \log Q_i \log Q_j$	$a_1 + \sum_i a_{1,i} \log Q_i$

Notes: Three candidate utility functions are tabulated, along the output elasticity ε_1 for one good Q_1 . Constant coefficients are A , a generic multiplier, a_i , a technical coefficient, and σ , which yields an elasticity of substitution.

One notable feature is that these utility functions, namely the Constant Elasticity of Substitution (CES) and the Translog, preserve the Cobb-Douglas as a special case. In both instances, output elasticities, and thus entropy, change if growth in utility components is uneven. This is readily seen in the third column of Table 6.2.1. In a CES scenario where Q_1 grows faster than Q_2 , the expenditure share on Q_1 is predicted to decline if both components are complements ($\sigma < 0$), or increase if both are substitutes ($\sigma > 0$). This is due to the indirect effect on prices, thus breaking the tautology pointed out at the beginning of this section.

The Translog function obtains a similar result, where the signs of the second-order coefficients determine the direction for the expenditure share. Starting input levels will make a difference, since entropy will specifically increase (decrease) if expenditure shares grow more uniform (concentrated). It is possible to derive similar elasticities for other forms of utility.

6.2.4 Choosing weights

As per Corollary 6.1.1, changing indexation scheme Ω_t can ameliorate, or worsen, the bias in real output in Proposition 6.1. However, one detail is that this corollary considers a static indexation scheme, built with information learnt in period t alone. Since quantity and price indices are defined by two-period changes, can we get around the bias by finding weights that minimise the change in cross-entropy, instead of its level?

Generally, yes. Table 6.2.2 categorises certain well-known indexation schemes, and demonstrates how cross-entropy enters the indexation formula. This variety in methods is due to the fact that some ‘optimal’ index should fulfill multiple criteria, which Eichhorn (1978) proves are impossible to satisfy simultaneously. Additionally, the desirability of certain indices depends on their interpretation within economic theory, or simple intuition.

Examples of weighting schemes A prominent choice for weighting scheme Ω_t defined in continuous time is Divisia weights (Divisia 1925). Hulten (1973) demonstrates that, under the parameterisation in Eq. 6.10 plus an assumption on linear homogeneity, so that $\sum_i \varepsilon_{i,t} = 1$, these indices are path-independent. The Divisia index is conventionally approximated with two-period averages in expenditure shares for discrete time observations. Table 6.2.2 suggests that the bias due to cross-entropy under the Divisia approximation is

Table 6.2.2: Cross-entropy under different indexation schemes

Name	$\omega_{i,t}$	$\Delta H(\mathbf{\Omega}_t, \mathbf{Y}_t)$
Divisia	$\frac{1}{2}(y_{i,t} + y_{i,t-1})$	$\frac{1}{2}[D_{KL}(\mathbf{Y}_{t-1} \parallel \mathbf{Y}_t) - D_{KL}(\mathbf{Y}_t \parallel \mathbf{Y}_{t-1})]$
Törnqvist	$\frac{1}{2}(y_{i,t1} + y_{i,t0})$	$\frac{1}{2}[D_{KL}(\mathbf{Y}_{t0} \parallel \mathbf{Y}_{t1}) - D_{KL}(\mathbf{Y}_{t1} \parallel \mathbf{Y}_{t0})]$
Sato-Vartia 2	$\frac{y_{i,t} - y_{i,t-1}}{\log y_{i,t} - \log y_{i,t-1}} / \sum_{i=1}^N \frac{y_{i,t} - y_{i,t-1}}{\log y_{i,t} - \log y_{i,t-1}}$	0
Chained	$y_{i,t-1}$	$D_{KL}(\mathbf{Y}_{t-1} \parallel \mathbf{Y}_t)$

Notes: The bias in real output due to cross-entropy in output shares, listed in the final column, varies by indexation scheme, named and expressed in the first two columns. $D_{KL}(x \parallel y)$ is the Kullback-Leibler (KL) divergence of y from x (Kullback & Leibler 1951, Cover & Thomas 2005).

small, specifically one half of the difference in the KL divergence of output shares in periods t and $t - 1$. This term is large if fluctuations in output shares $y_{i,t}$ are large, but close to zero otherwise.

Since Diewert’s (1976) seminal contribution, statistical agencies adopted the use of standard techniques in forming aggregate price indices, motivated by economic theory. One recommended method is the Törnqvist index, which adopts a similar form to Eq. 6.4. Specifically, it assigns weights equal to the average output shares between two arbitrary periods – t_0 and t_1 in Table 6.2.2. Törnqvist weights coincide with the Divisia approximation when computed over rolling observations. Diewert motivates the properties of this index, which he defines to be ‘superlative’, by demonstrating how it benchmarks the price to an assumed average utility between the two periods of choice. Practically, it is useful because it may cancel the cross-entropy term in Eq. 6.4, depending on how linearly the true weights change over time.

Whereas the Törnqvist index corresponds to Translog utility, the Sato-Vartia index corresponds to CES preferences (Sato 1976, Vartia 1976).⁵ The Sato-Vartia index’s own desirable property is that it adjusts for substitution between inputs via some constant elasticity, although at the price of violating the monotonicity axiom of Eichhorn & Voeller (1976) (Reinsdorf & Dorfman 1999). Redding & Weinstein (2019) further critique this index, demonstrating that it is problematic when preferences change over time. As far as accounting for variety is concerned, one feature that stands out in Table 6.2.2 is that the cross-entropy residual under Sato-Vartia indexation is precisely zero. The weights cancel the change in logarithmic output shares, and the remaining expression sums to zero.

Finally, I mention the chain index. The chain index, which updates weights to the previous period’s output shares, is widely used in statistical agencies’ CPI computations. The weights for consumption are updated annually and used to form monthly CPI estimates. This index provides a straightforward interpretation of price and quantity changes, as it is tied to a single representative basket from the recent past. This feature is particularly useful in high-frequency settings where there are persistent fluctuations, such as seasonal goods. In terms of accounting for variety, the cross-entropy term for a chained index reflects the

⁵The specific version I consider is its second variant, Sato-Vartia 2.

simplest and most intuitive form for a change in entropy, which is the KL divergence of the current consumption basket from the previous basket.

6.2.5 Summary

This section demonstrates that real output indices of the type defined in Definition 6.2 are biased if extrapolated from GDP using a counterpart price index, to the degree that the indexation scheme differs from consumption patterns. In addition, this bias from cross-entropy in the indexation scheme and consumption patterns enters real output growth indices by at least the amount that the entropy of GDP changes. This bias is larger if the indexation scheme does not match consumption patterns, as is the case for common, economically-motivated real output indices.

Importantly, this bias remains present as long as preferences are not Cobb-Douglas. Therefore, this result suggests that adjusted real output indices are needed to account for varying preferences in order to obtain an accurate representation of economic activity. The main limitation in this argument is that it is derived assuming perfect competition. Recent studies investigate the role of markups and taxes that further distort the observed output distribution [Baqae & Farhi \(2020\)](#). However, further adjustments for such wedges should manifest *in addition to* the cross-entropy term that relies on modelling output shares under perfect competition.

6.3 Indices, revisited

6.3.1 Real output: Hulten's Paradox in KLEMS data

What does Proposition 6.1 imply for the measurement of living standards in the long run? To demonstrate the role of changing consumption patterns, I revisit Hulten's Paradox. [Hulten \(1997\)](#) established the paradox in response to [Nordhaus's \(1997\)](#) claim that prices grew slower than commonly accepted. He noted that the revision implies faster volume growth: when extrapolated back in time, this correction in prices leaves our ancestors with very little income in real terms. A paradox manifests in the observation that these ancestors were demonstrably capable of procuring sustainable lifestyles.

Solving the paradox The exchange centred on pricing attributes for products (dis)appearing for the first time, but this paper takes issue with one particular step in Hulten's logic.⁶ He asserts that if true price growth is overestimated, and volume growth underestimated, the average colonial and modern Americans are indifferent between the average living standard of colonial America, and that afforded with \$90–\$430 in 1997. However, not only is the budget of 90–430 1997 dollars fixed, but so is the 1997 assumed consumption basket. The decisions from both modern and colonial Americans to travel across time should not be

⁶Quality improvements, the like [Nordhaus \(1997\)](#) investigates, are indeed important in biasing inflation indices upwards in matched-model methods; the availability of characteristics is increasing with new products, but on paper their prices may appear the same. [Gordon \(2009\)](#) proposes that, just as matched models bias inflation upwards, so they may bias inflation downwards, with apparel prices as a case in point.

equivalent, because they adhere to different consumption patterns. In absolute terms, goods consumed in 1997 were likely scarce in colonial America – think of the entertainment a new Game Boy offers – while goods consumed by colonial Americans are abundant in 1997. The modern CPI Hulten uses to deduce living standard improvements one hundred years ago is off the mark.

Table 6.3.1: Yearly GDP growth with different industry indices: US, 1947–2014

	$\Delta \log Y_t$	$\tilde{P}_t$	$\tilde{Q}_t$	$\Delta H(\Omega_t, Y_t)$
Basket from 1947	6.33	3.17	2.40	0.76
Basket from 2014	6.33	3.53	3.48	-0.68
<i>Difference</i>	0.00	-0.36	-1.08	1.44
Törnqvist 1947&2014	6.33	3.35	2.94	0.04
Chained	6.33	3.36	3.15	-0.18
Divisia	6.33	3.32	3.01	0.00

Notes: This table decomposes average nominal GDP growth from 1947 to 2014 between the average growth in price, volumes, and change in cross-entropy. Weights Ω_t are labelled in the first column, and computed as per Table 6.2.2. All values denote log points. These observations aggregate across 65 industries, for which data are made available by World KLEMS.

Is the bias in real output due to entropy significant in practice? Returning to Eq. 6.4, the cross-entropy term increases over time as real output growth diverges between industries from the perspective of a colonial American, even when keeping inflation rates equal across industries. From the perspective of the modern American, aggregate GDP appears to have grown substantially, but this is only due to a decline in cross-entropy: they happen to consume more goods that experienced higher real output growth rates.

Data In order to capture the order of magnitude of this difference, I reproduce an aggregate GDP deflator from US World KLEMS data.⁷ These data cover 65 industries, from 1947 to 2014. The files provide yearly GDP by industry in current and constant prices. From this, I back out industry-specific price indices by dividing reported industry-level nominal gross output by its respective output in real quantities. Using these gross output deflators as P_i 's, I can estimate each term in Eq. 6.4 using different weighting schemes. Specifically, in Table 6.3.1 I compare the composition of average, yearly GDP growth for a representative basket in 1947 – fixing weights to industry shares of GDP in 1947 – to a representative basket in 2014 – with weights fixed to industry shares in 2014. In addition, I include a Törnqvist index averaging 1947 and 2014, a chained index, and a rolling Törnqvist average to approximate a Divisia index, as outlined by Table 6.2.2.

Results The first column in Table 6.3.1 reports the average yearly nominal GDP growth rate, which is always constant. The second column reports the annual average price index, and the third the quantity index, constructed with the weighting scheme labelling each

⁷<http://www.worldklems.net/data.htm>

row. Using the annual industry-level nominal output shares, I retrieve the average cross-entropy residuals corresponding to those schemes in the fourth column. Given that the decomposition in Proposition 6.1 is exact, the final three columns always add up to the nominal output growth rate reported in the first column.

The first result is that, as expected, real output growth rates are higher when indexing industry contributions to the distribution of nominal output from 2014 in the second row, relative to 1947 in the first row. This difference, reported in the third row of the third column, is significant, in the order of 1 percentage point (pp) per year. This translates to overstated real welfare improvements of about 70% after seven decades, if consumption patterns in 2014 are used as the benchmark.

The second, perhaps surprising outcome of this exercise is that aggregate deflators contribute *positively* to this gap. In the first and second rows of the second column, the average product in 1947 turned out *less* expensive than the average product in 2014. Therefore, a resolution to Hulten’s Paradox may not lie at all in revising aggregate price indices upward by finding negative CPI biases, but rather in accounting for the change in industrial variety of GDP.

How does this compare to measures of welfare motivated by economic theory? The fourth row of Table 6.3.1 reports the contributions to GDP growth by adopting average weights between 1947 and 2014 as the benchmark. As expected, the loss of information by this index is substantially lower, although slightly positive. This would suggest that industry shares evolved relatively smoothly during that period.

The fifth row reports the indices and cross-entropy for chained weights. As stated previously, this weighting scheme is perhaps most intuitive, since it benchmarks price and quantity growth to the most recent reporting period. The average rate of information loss between two subsequent years, which coincides with the KL divergence reported in Table 6.2.2, is relatively small at 0.18pp, almost 3% of average GDP growth.

Finally, the sixth row demonstrates that approximating a Divisia index yields a tiny cross-entropy term, although not exactly equal to zero. Therefore, this index is desirable for those who want to reduce the cross-entropy term to the greatest extent possible, short of using the Sato-Vartia 2 indexation scheme.

Appendix 6.B.1 demonstrates that CPI weights from statistical agencies produce similar results, even though they include in the order of 200 product groups instead of coarse industry classes.

Sources of structural change It remains unclear whether these aggregate statistics conceal meaningful trends in the structure of the US economy over the last seventy years. Figure 6.3.1 plots the entropy rate of US GDP, $H(Y_t)$, for years reported by World KLEMS. The underlying unit plotted on the vertical axis is the base of the natural logarithm, Euler’s number, conventionally known as a ‘nat’ (Cover & Thomas 2005).⁸ These industry data demonstrate a stunning reversal in the entropy rate, which has increased up to the early

⁸Base two measures of entropy are known as ‘bits’, which are used to measure information in binary computing systems.

1980s, but since declined. This suggests that US industrial output has been most *uniform* in the early 1980s, but turned relatively more concentrated in recent years.

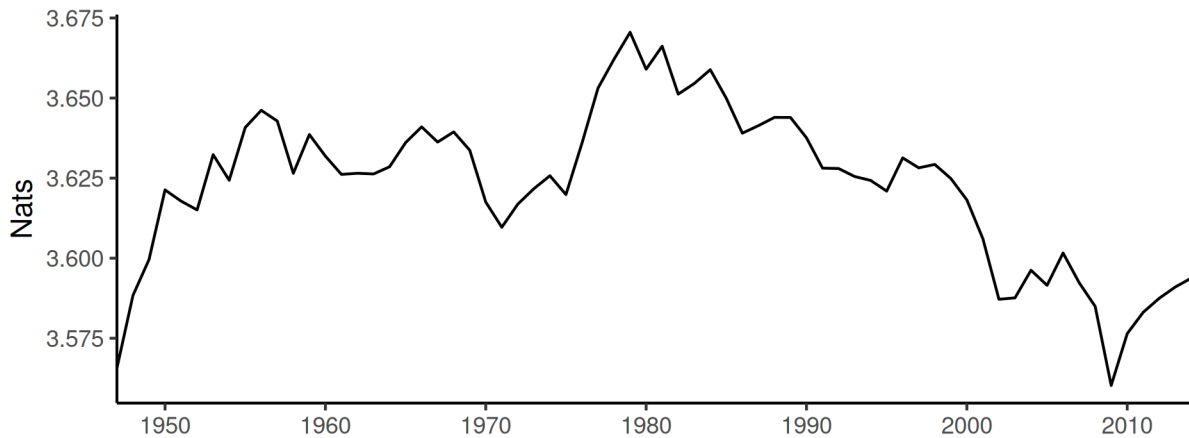


Figure 6.3.1: **Entropy of US GDP increases until 1980, but has declined since:** This figure plots the entropy rate of US GDP, measured in nats, from 1947 to 2014. Data are made available by World KLEMS.

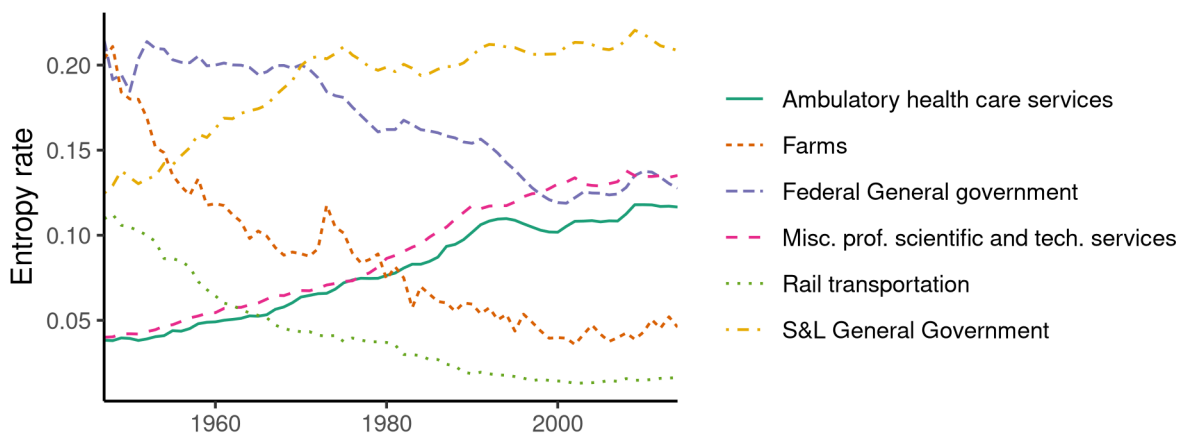


Figure 6.3.2: **Largest contributions to US entropy, 1947–2014:** This figure plots six industries that have the three largest positive and negative contributions to the entropy of US GDP. Data are made available by World KLEMS.

Figure 6.3.2 helps understand where most of the change in US entropy originated, by plotting the entropy rate contributions

$$h_{i,t} = -y_{i,t} \log y_{i,t}$$

for some industries i . I plot these rates for the three largest positive and negative contributors to the change in the total entropy rate observed, over the whole sample, by comparing the sums of entropy rates $\sum_{t=1947}^{2014} h_{i,t}$ between industries. The industries that emerge are well-documented in studies of structural change for the US economy. The largest declining industries are agriculture, rail transport, but also output from federal government activities. On the other hand, the rise of health care, professional services, and S&L general government contributed most to an increasing entropy rate.

A promising avenue for further research constitutes aligning changes in the distribution and concentration of US industrial output with known drivers of structural change. One important phenomenon is the emergence of IT industries, and service industries generally. A second driver is the rise of global trade, which significantly reduced the importance of certain manufacturing industries in US domestic output. A third is cost disease (Baumol 1967), which likely affected agriculture.

6.3.2 Labour productivity: technology or allocation?

Cross-entropy is a measure of how closely two distributions overlap. Recall, from Figure 6.2.1, that cross-entropy is minimised when the two distributions are identical. This observation lends itself to a useful application to the decomposition of labour productivity. To see this, I first define a labour productivity index in Definition 6.3.

Definition 6.3 (Real productivity index). *Real aggregate productivity growth $\Delta \log q_t$ equals GDP growth, minus labour input growth and an aggregate inflation index:*

$$\Delta \log q_t = \Delta \log Y_t - \Delta \sum_{i=1}^N \omega_{i,t} \log P_{i,t} - \Delta \log L_t. \quad (6.14)$$

A common choice for weights $\omega_{i,t}$ in the productivity literature is the rolling Törnqvist average, which coincides with the Divisia approximation expressed in Table 6.2.2. Unlike before, productivity entails the use of two aggregate statistics; nominal output and labour, both of which are distributed among N industries. This means that the behaviour of the aggregate index $\Delta \log q_t$ hinges on the entropy of both nominal output and labour input. This leads to an elegant decomposition of aggregate labour productivity between: i) ‘technology’ (how many more units do industries produce per unit of labour, on average) and the allocation of ii) ‘demand’ (are expenditures for industries, that are weighted relatively more, relatively lower) and iii) labour (are labour input shares equal to industry weights). Proposition 6.3 outlines the components of this decomposition.

Proposition 6.3 (Aggregate real productivity decomposition). *Aggregate real productivity growth can be decomposed exactly between technology, plus the allocation of demand and labour:*

$$\Delta \log q_t = \underbrace{\Delta \sum_{i=1}^N \omega_{i,t} \log \left(\frac{Q_{i,t}}{L_{i,t}} \right)}_{\text{Technology}} + \underbrace{\Delta D_{KL}(\mathbf{\Omega}_t \| \mathbf{Y}_t)}_{\text{Demand}} + \underbrace{\Delta D_{KL}(\mathbf{\Omega}_t \| \mathbf{L}_t)}_{\text{Labour}}. \quad (6.15)$$

Proof. See Appendix 6.A.2.

□

Cost disease Proposition 6.3 shares some similarities with the method presented by Tang & Wang (2004), but offers an improved interpretation and additivity of log growth rates.

The decomposition predicts that demand allocation can contribute positively to the labour productivity growth index. This is reasonable in the sense that important industries, with larger weights, should cost less, and thus capture a smaller expenditure share. The demand allocation term constitutes an important link to [Baumol's \(1967\)](#) cost disease hypothesis, by which industries with low productivity growth capture a larger share of output. The precise version of the cost disease hypothesis that the demand allocation term captures will specifically depend on the choice of weights.

Table 6.3.2: Allocation in the slowdown of labour productivity growth

		$\Delta \log q_t$	Technology	Demand	Labour
<i>France</i>	1995–2005	1.64	1.58	0.14	-0.08
	2006–2017	0.64	0.73	-0.02	-0.08
	Slowdown	1.01	0.85	0.16	0.00
	Share	1.00	0.84	0.15	0.00
<i>Germany</i>	1995–2005	1.84	1.49	0.16	0.18
	2006–2017	0.87	0.96	0.02	-0.10
	Slowdown	0.97	0.54	0.15	0.28
	Share	1.00	0.55	0.15	0.29
<i>Japan</i>	1995–2005	1.75	1.48	0.12	0.15
	2006–2015	0.80	0.56	-0.00	0.25
	Slowdown	0.95	0.92	0.13	-0.11
	Share	1.00	0.98	0.13	-0.11
<i>United Kingdom</i>	1995–2005	2.18	1.90	0.18	0.11
	2006–2017	0.38	0.06	-0.02	0.34
	Slowdown	1.80	1.83	0.20	-0.23
	Share	1.00	1.02	0.11	-0.13
<i>United States</i>	1997–2005	2.45	2.58	0.22	-0.35
	2006–2017	0.88	1.02	-0.01	-0.13
	Slowdown	1.57	1.56	0.23	-0.22
	Share	1.00	0.99	0.15	-0.14

Notes: This table reports the sources of the labour productivity slowdown in five advanced economies, using the decomposition in Eq. 6.15. Weights are two-period Törnqvist averages of nominal output shares. Data from EU-KLEMS 2019.

Labour misallocation While demand allocation contributes positively to the labour productivity growth index, labour allocation has a negative effect. This is because, if an industry receives a greater amount of labour inputs than its weight in the weighting scheme, it creates a reallocation opportunity by which relatively more labour input is sent to relatively important industries. In Proposition 6.3, this re-allocation can reduce the KL divergence of labour from the scheme towards zero. In contrast to [Tang & Wang \(2004\)](#), the decomposition does not focus exclusively on the productivity levels of industries. Rather, it measures re-allocation according to the weighting scheme that is chosen, which does not have to reflect real output levels.

Explaining the slowdown in labour productivity growth Are the allocation terms significant for the infamous labour productivity slowdown? I use the conventional approach of two-period averages for the industries' nominal value added shares as the indexation scheme. Labour inputs are defined as number of hours worked. To summarise the problem, the first column reports real aggregate labour productivity growth (from a price index derived using appropriate weights) as an average for years pre- and post-2005, for France, Germany, Japan, the UK and the US. The slowdown is then defined as the difference between those two average growth rates; these range from around 1pp for the first three countries, to more than 1.5pp in the UK and US. As Table 6.3.2 demonstrates, these allocation terms explain the labour productivity slowdown in varying proportions.

One of the key findings of the study is that the reallocation of demand appears to have worsened in all countries, explaining between 10 and 15% of the slowdown across the board. However, the experience with labour reallocation is more varied. In Germany, labour reallocation worsened after 2005, and explains almost 30% of the labour productivity slowdown. On the other hand, labour reallocation does not appear to have changed significantly in France, and has actually contributed positively to labour productivity in Japan, the UK, and the US. Taken together, these results indicate that demand and labour reallocation account for almost half of the slowdown in labour productivity for Germany. However, for the other countries, the pure technology component, which measures the average industry's ability to produce more per hour worked, is the single explanation.

6.4 Conclusion

In this paper, I highlight the importance of accounting for product variety in the analysis of aggregate quantities. I demonstrate that failing to consider the changing representation of products over time can lead to biased estimates of real output. For instance, the real output growth experienced by a modern product may appear higher simply because the product was rarely consumed in the past, irrespective of its actual inflation rate. Similarly, the current real output growth of a representative product from the past may be underestimated due to its diminishing importance over time. This bias can be quantified by the cross-entropy of GDP between the benchmark year and the remaining period. I interpret this measure as the loss of information inherent in an index that accumulates over time. These findings highlight the need for careful consideration of product and industry variety in the study of aggregate quantities.

The loss of information due to accounting for variety has a significant impact on measuring GDP growth, even in fine-grained industry breakdowns. In the US World KLEMS data, for instance, it results in a doubling of the real output index level when using current industry compositions of GDP, compared to those from seven decades earlier. This finding is consistent with a more detailed analysis of CPI weights from 1998 to 2018, as well as with consumer price data from Eurostat, which suggests that extrapolated real output growth rates from aggregated inflation data are about 1–2pp higher than actual growth rates from observed real output data. Additionally, by incorporating the allocation of demand and

labour inputs, a decomposition of labour productivity provides insight into the origins of the labour productivity slowdown in Germany since 2005.

The central argument of the paper is that variety plays a crucial role in analysing heterogeneous economic data, and this warrants careful consideration. Although the study focuses on a single application in the context of long-term economic growth, it suggests that accounting for variety could be relevant in other areas such as monetary aggregates and employment. There are several measures of variety that extend beyond the one examined in this paper, which could offer fresh perspectives. As researchers have access to increasingly granular datasets, this subject gains even greater significance, expanding the horizon beyond the analysis of singular aggregate quantities. The implications of this approach are broad, and further exploration could lead to new insights and innovative applications.

Bibliography

- Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J. & Li, H. (2019), 'Missing Growth from Creative Destruction', *American Economic Review* **109**(8), 2795–2822.
- Baqaei, D. & Burstein, A. (2021), Welfare and output with income effects and taste shocks, Working Paper 28754, National Bureau of Economic Research.
- Baqaei, D. R. & Farhi, E. (2020), 'Productivity and misallocation in general equilibrium', *Quarterly Journal of Economics* **135**(1), 105–163.
- Baumol, W. J. (1967), 'Macroeconomics of unbalanced growth: The anatomy of urban crisis', *American Economic Review* **57**(3), 415–426.
- Bresnahan, T. F. & Gordon, R. J., eds (1997), *The Economics of New Goods*, University of Chicago Press.
- Cover, T. M. & Thomas, J. A. (2005), *Entropy, Relative Entropy, and Mutual Information*, John Wiley & Sons, Ltd, chapter 2, pp. 13–55.
- Deaton, A. (2010), 'Price indexes, inequality, and the measurement of world poverty', *The American Economic Review* **100**(1), 5–34.
- Diewert, W. (1976), 'Exact and superlative index numbers', *Journal of Econometrics* **4**(2), 115 – 145.
- Divisia, F. (1925), 'L'indice monétaire et la théorie de la monnaie', *Revue d'économie politique* **39**, 1121.
- Eichhorn, W. (1978), *Systems of Functional Equations in the Theory of the Price Index*, Addison-Wesley Publishing Company, chapter 8, pp. 151–172.
- Eichhorn, W. & Voeller, J. (1976), *Theory of the Price Index: Fisher's Test Approach and Generalizations*, Vol. 140 of *Lecture Notes in Economics and Mathematical Systems*, Springer, Berlin.

- Fisher, I. (1922), *The Making of Index Numbers*, Houghton-Mifflin.
- Gordon, R. J. (2009), Apparel prices 1914-93 and the Hulten/Bruegel paradox, in W. E. Diewert, J. S. Greenlees & C. R. Hulten, eds, 'Price Index Concepts and Measurement', University of Chicago Press, pp. 85–128.
- Hulten, C. R. (1973), 'Divisia index numbers', *Econometrica* **41**(6), 1017–1025.
- Hulten, C. R. (1997), Comment on "Do Real-Output and Real-Wage Measures Capture Reality? The History of Lighting Suggests Not", in [Bresnahan & Gordon \(1997\)](#), pp. 66–70.
- Kullback, S. & Leibler, R. A. (1951), 'On Information and Sufficiency', *The Annals of Mathematical Statistics* **22**(1), 79 – 86.
- Maasoumi, E. (1993), 'A compendium to information theory in economics and econometrics', *Econometric Reviews* **12**(2), 137–181.
- Maslow, A. H. (1943), 'A theory of human motivation', *Psychological Review* **50**(4), 370–396.
- Nordhaus, W. D. (1997), Do Real-Output and Real-Wage Measures Capture Reality? The History of Lighting Suggests Not, in [Bresnahan & Gordon \(1997\)](#), pp. 27–70.
- Redding, S. J. & Weinstein, D. E. (2019), 'Measuring Aggregate Price Indices with Taste Shocks: Theory and Evidence for CES Preferences', *The Quarterly Journal of Economics* **135**(1), 503–560.
- Reinsdorf, M. B. & Dorfman, A. H. (1999), 'The Sato–Vartia index and the monotonicity axiom', *Journal of Econometrics* **90**(1), 45–61.
- Sato, K. (1976), 'The ideal log-change index number', *The Review of Economics and Statistics* **58**(2), 223–228.
- Shannon, C. E. (1948), 'A Mathematical Theory of Communication', *The Bell System Technical Journal* **27**, 379 – 423.
- Tang, J. & Wang, W. (2004), 'Sources of aggregate labour productivity growth in Canada and the United States', *The Canadian Journal of Economics / Revue canadienne d'Economie* **37**(2), 421–444.
- Theil, H. (1967), *Economics and information theory*, Studies in mathematical and managerial economics ; v. 7, North-Holland ; Rand McNally, Amsterdam : Chicago.
- Törnqvist, L. (1936), 'The Bank of Finland's consumption price index', *Bank of Finland Monthly Bulletin* **10**, 1–8.
- Vartia, Y. O. (1976), 'Ideal log-change index numbers', *Scandinavian Journal of Statistics* **3**(3), 121–126.

Appendix

6.A Proofs

6.A.1 Proposition 6.2

Dropping subscript i from Eq. 6.9, elasticity

$$\varepsilon = \frac{dY}{dQ} \frac{Q}{Y} \quad (6.16)$$

is assumed constant. To derive the functional form of $Y = f(Q)$ that satisfies constant ε , re-arrange the expression, then integrate both sides:

$$\int f'(Q) dQ = \int \varepsilon \frac{f(Q)}{Q} dQ \quad (6.17)$$

$$= f(Q) + C = \varepsilon \int \frac{f(Q)}{Q} dQ \quad (6.18)$$

$$= \varepsilon f(Q) \log Q - \frac{\varepsilon^2}{2} f(Q) \log^2 Q + \frac{\varepsilon^3}{6} f(Q) \log^3 Q - \mathcal{O}(\log^4 Q) \quad (6.19)$$

$$= f(Q) \left[1 - \sum_{n=0}^{\infty} \frac{\varepsilon^n (-1)^n}{n!} (\log Q)^n \right] \quad (6.20)$$

$$= f(Q) \left[1 - \sum_{n=0}^{\infty} \frac{1}{n!} (\log Q^{-\varepsilon})^n \right] \quad (6.21)$$

$$= f(Q) - f(Q) \frac{1}{Q^\varepsilon}, \quad (6.22)$$

$$\Rightarrow f(Q) = Q^\varepsilon c, \quad (6.23)$$

where $C = -c$ is the constant of integration. Step 6.18 is possible because ε is assumed not to vary with Q . Step 6.19 follows from recursively integrating by parts and substituting $\varepsilon = Qf'(Q)/f(Q)$. The terms are collected in the sum of Step 6.20 after adding and subtracting one. This summation term is a Taylor expansion of $\exp(x)$ about zero, and thus simplifies to Step 6.22. Re-arranging yields the final result.

6.A.2 Proposition 6.3

Starting from productivity defined in Definition 6.3,

$$\Delta \log q_t = \Delta \log Y_t - \Delta \sum_{i=1}^N \omega_{i,t} \log P_{i,t} - \Delta \log L_t \quad (6.24)$$

$$= \Delta \sum_{i=1}^N \omega_{i,t} \log Q_{i,t} - \Delta \sum_{i=1}^N \omega_{i,t} \log L_{i,t} + \Delta \sum_{i=1}^N \omega_{i,t} \log \frac{Y_t}{Y_{i,t}} - \Delta \sum_{i=1}^N \omega_{i,t} \log \frac{L_t}{L_{i,t}} \quad (6.25)$$

$$= \Delta \sum_{i=1}^N \omega_{i,t} \log q_{i,t} + \Delta \sum_{i=1}^N \omega_{i,t} \log \frac{\omega_{i,t} Y_t}{Y_{i,t}} - \Delta \sum_{i=1}^N \omega_{i,t} \log \frac{\omega_{i,t} L_t}{L_{i,t}} \quad (6.26)$$

$$= \Delta \sum_{i=1}^N \omega_{i,t} \log q_{i,t} + \Delta D_{KL}(\mathbf{\Omega}_t \| \mathbf{Y}_t) - \Delta D_{KL}(\mathbf{\Omega}_t \| \mathbf{L}_t), \quad (6.27)$$

where $q_{i,t} = Q_{i,t}/L_{i,t}$, and $D_{KL}(x||y)$ is the Kullback-Leibler (KL) divergence of y from x (Kullback & Leibler 1951).

6.B Additional results

6.B.1 Does the level of aggregation matter?

Since the entropy of GDP clearly matters in measuring historical living standards, one wonders whether the 65 industries offered by World KLEMS paint an accurate picture. As mentioned in Section 6.2, data on disaggregated output, and thus weights, are hard to come by. Privacy protection rules also prevent publication of the most granular decomposition in industry output.

BLS data However, through remarkable efforts of statistical agencies in recent decades, data on components of consumption are more extensive, reliable and granular. On the product level, the BLS is tasked with assigning weights to products consumed by households for the purpose of estimating its CPI, which entails the use of surveys. The BLS publishes price indices for 182 products from 1998 to 2018, of which 10 have data for only a subset of years.⁹ These prices are accompanied by official weights used to form aggregates, from surveys of what consumers buy for day-to-day living. After building a price index from these data, one can deflate any consumption series. I supplement the results from BLS data using the Harmonised Index of Consumer Prices (HICP) published by Eurostat, showing that the cross-entropy bias typically varies between 1–4pp per year across 33 European countries.

Section 6.2 establishes that this real output series is biased by the cross-entropy of CPI weights and the actual distribution of consumption. This bias is minimal when the weights are equal to the distribution of output, at which point it is equal to the entropy of consumption. Therefore, even though expenditure data is unavailable for the 182 products, the weights alone determine a ‘best-case’ outcome for this bias.

⁹Omitting the incomplete series does not significantly alter results.

Table 6.B.1: Yearly growth in consumer price and cross-entropy: US, 1998–2018

	$\widetilde{P}_t$	$\Delta H(\Omega_t)$
Basket from 1998	2.07	0.57
Basket from 2018	2.12	-0.48
<i>Difference</i>	-0.05	1.05
Törnqvist 1998&2018	2.08	0.06

Notes: BLS data on prices and weights for 182 products are used to form price indices in the first column, fixing weights to 1998 and 2018. The second column computes the average cross-entropy of those weights and the weights in all other years, denoted ω . All values denote log points. Data are made available by the BLS.

Results: US Table 6.B.1 computes this bias with CPI weights provided by the BLS. There are persistent differences between the CPI and the PPI, so these results do not compare directly to those of Table 6.3.1. However, the magnitude of the bias is surprisingly similar, at about 1pp per year. The bias for the Törnqvist index is also close, however small, at 0.06pp. At a rate of 1pp, deflating consumption using current expenditure patterns overstates improvements in living standards from fifty years ago by *at least* 65%. Another finding reproduced in Table 6.B.1 is the smaller inflation rate for the index based on 1998 consumption patterns.

Eurostat data The European HICP, maintained by Eurostat, mimics the construction of the BLS' CPI, with the advantage of offering granular data for 33 European countries. Published data for the HICP include five levels of aggregation, with the 5-digit level counting up to 264 product categories. Decent coverage is available at that level for France, Lithuania and Slovenia. For the remaining countries, data is available at the 4-digit level, which includes up to 72 product categories. Generally, the price data span from 1996 to 2020, with some exceptions. Using those observations, I repeat the exercise from BLS CPI data for each European country. This will illustrate how varied the change in cross-entropy can be across different economies.

Results: HICP By using weights from the first and final years available, Figure 6.B.1 reproduces the difference in average yearly changes in cross-entropy on the left, the difference in average yearly price growth rates in the middle, and the average yearly change in cross-entropy when using a Törnqvist average weight. These are comparable to row three, and the third column in row four, of Table 6.B.1.

The middle chart in Figure 6.B.1 demonstrates that the average rate of information loss from cross-entropy for an aggregate real output index is between 1–2pp per year. These observations align closely to the 1.05pp found for US CPI data, but exhibit some variation, with 6.31pp an extreme case in Lithuania. It is not clear whether the level of aggregation has an obvious impact on cross-entropy estimates, since the number of countries publishing 5-digit level data is too low for a robust hypothesis test.

The second chart suggests that yearly price growth rates are higher when assigning weights from the first year of middle relative to the final year. This evidences some sort

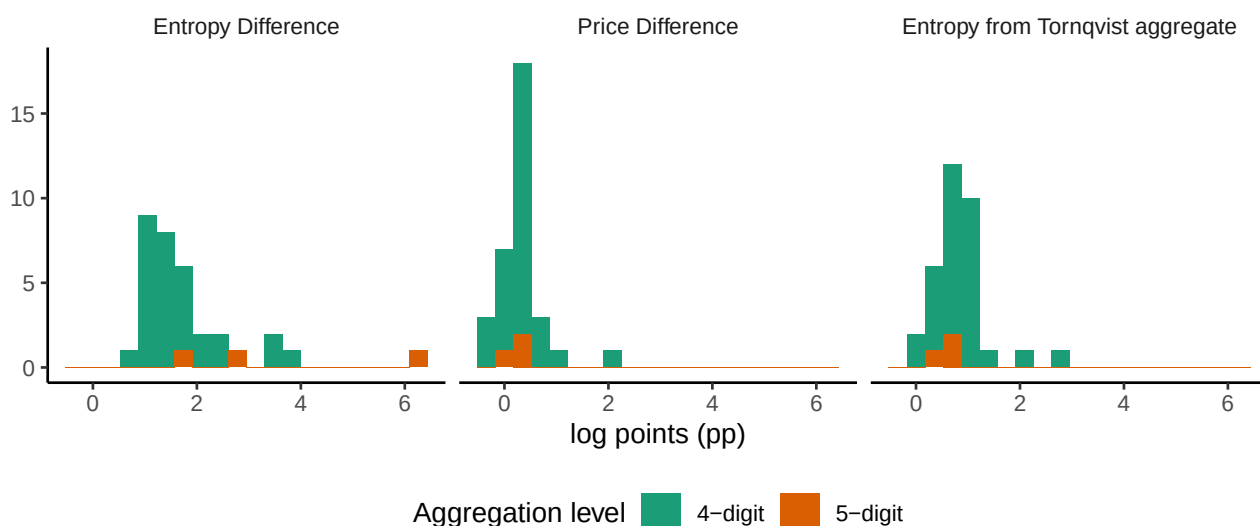


Figure 6.B.1: Comparing price and entropy from first and last available weights in HICP data: 33 European countries, 1997–2020: Indexing aggregate price inflation to the start of the observation period yields an estimate for entropy (Left) that is around 1–2pp higher relative to an index composed of final-year weights. Overall inflation rates (Middle) are generally similar, if not slightly larger when indexing to final-year weights. A Törnqvist index (Right) averaging first- and final-year baskets lowers the average amount of cross-entropy in the index, but can still be in the order of 1pp.

of substitution bias, by which consumption in later years favours products which grew relatively cheaper.

Finally, the right chart indicates that even Törnqvist aggregation, by which an average weight is derived from weights in the first and final years, can be subject to substantial bias from a change in cross-entropy. This suggests that the small result in US data, seen for industry deflators in Table 6.3.1 and in CPI weights in Table 6.B.1, may be special cases. In contrast, consumption patterns in other countries are subject to substantial shifts, even within an observation period of 20 years.

Chapter 7

Volatility in forward rates

Abstract

This paper proposes a model for the Forward Rate Curve, and its fluctuations, founded on the operations of a banking sector. The returns of a representative bank are maximised given constraints on the duration of its assets. The shape, fluctuations and extremes of the curve are predicted to vary with linear duration, as well as geometric duration. This reflects a representative bank's tendency to discount short-term cash flows hyperbolically, but long-term cash flows exponentially. I validate these predictions using data on Eurodollar future prices from 2013 to 2020.

7.1 Introduction

A few undesirable features plague textbook models of the yield curve. The idea that long-dated forward rates of interest reflect expected short rates, known as the *Expectations Hypothesis* (EH), fails to explain the persistent overshooting of long-term rates relative to their eventual spot rates (Fama & Bliss 1987, Campbell & Shiller 1991). Statistical models that rely on non-arbitrage conditions fail to explain why rates at the one to five year horizon are more volatile (Vasicek 1977, Nelson & Siegel 1987).

This paper proposes a model for the Forward Rate Curve (FRC) based on the operations of a banking sector. The key role of a bank driving the model is its ability to hold a short position on low-maturity deposits while maintaining a long position on high-maturity bonds. I find that the model fits the FRC, and its fluctuations, well under most circumstances. An interesting feature is that the data on interest rate futures matches the discount factor for a bank that uses hyperbolic discounting in the short term – a time horizon of about three years – but exponential discounting in the long term – beyond three years. The corresponding discount rates determine the maturity at which the magnitude of FRC fluctuations peaks, which is at the point where hyperbolic and exponential discounting meet.

I develop the model using assumptions based on the operations of a hypothetical, representative bank. A bank that seeks to maximise the prospective return on its bond portfolio would buy bonds that offer the highest yield irrespective of maturity. In a void, this implies that a competitive equilibrium yields uniform forward rates. The difficulty from the bank's perspective is that the maturity of its deposits is uncertain – households may withdraw them at any moment. I express this uncertainty as a constraint on the average maturity – or *duration* – of the cash flow it derives from its bond portfolio. The bank is tasked with maximising the return from its bonds subject to that constraint, which I show to resemble an entropy maximisation programme for the distribution of assets as a function of maturity.

The solution to the programme expresses the FRC as a linear function of maturity. This has the convenient interpretation that future cash flows are discounted by a representative bank, with a discount factor that is inversely related to the duration of deposits and regulates the slope of the FRC. An additional constraint on the geometric mean of maturity introduces a curvature-like term to the FRC, and can be interpreted as the marginal utility of cash flows discounted by a bank using hyperbolic discounting. Hyperbolic discounting weighs forward rates at the front end of the curve more favourably, whereas exponential discounting dominates the back end. I refer to this mix as 'gamma' discounting, since the discount function resembles a Gamma distribution.

The relative importance of exponential versus hyperbolic discounting, evaluated by their respective discount rates, determines the location of a turning point in the FRC. Importantly, it locates the maturity of the extreme forward rate fluctuation when these durations change, namely in the belly of the FRC. That is where hyperbolic and exponential discounting meet, and jointly determine the aggregate discount factor. The belly's exposure to two sources of variation implies that it exhibits larger fluctuations.

I fit the model's parameters to data on Eurodollar future prices from 2013 to 2020. For

the most part, the model matches observed patterns in FRC fluctuations at a quarterly frequency. Variation in linear duration generates volatility in spreads when the slope changes, whereas variation in geometric duration generates volatility when the body of the FRC moves. The model explains most of the FRC's shapes over time, except in exceptional circumstances such as the 2019 'taper tantrum'. There is limited evidence that the trade-off between linear and geometric duration explains peaks in FRC volatility, largely due to noisy parameter estimates at a daily frequency, and insufficient observations at the quarterly frequency.

Related literature Economic models of the FRC follow two perspectives. The first is the EH, and considers forward rates to reflect some expectations of the future spot rates of interest. [Campbell & Shiller \(1991\)](#) famously show that forward rates at the back end of the FRC are persistently higher than their realised spot rate. [Fama & Bliss \(1987\)](#) report a similar finding, and specifically note that the one year rate today predicts changes in the one year rate in over two years' time. [Fama \(2006\)](#) assigns this pattern to the reversion of interest rates to some time-varying, long run expected value. [Gabaix \(2012\)](#) proposes that the back end of the FRC is susceptible to rare, but large, negative shocks to consumption. This risk can explain higher yields, and may explain the observed mean reversion if such disaster risks fade over time.

One downside to the EH approach is that it does not use all the information available on the entire shape of the curve. The empirical exercises rest on studying the spreads on forward rates at arbitrary time horizons. One exception is the study by [Cochrane & Piazzesi \(2005\)](#), which evaluates the joint predictability of fluctuations along the entire yield curve. They find that the belly of the curve serves as an important predictor of bond returns, which may not be surprising given that its higher volatility ought to be compensated by higher expected returns. And yet none of these studies offer a compelling explanation for the excessive volatility of forward rates in the belly of the FRC specifically, beyond some notion that it is the average time horizon at which money markets 'revert' to a 'mean'. This paper uses the maturity-swapping mechanism to determine the shape of the entire FRC, in addition to the size of its fluctuations over time.

The second perspective, more prevalent in the recent literature, views the FRC as a stochastic process for the underlying Stochastic Discount Factor (SDF). [Vasicek's \(1977\)](#) canonical model, for example, assumes that the SDF follows an AR(1) process with standard normal shocks. This model has been modified to include time-varying volatility and risk premia ([Dai & Singleton 2000](#), [Duffee 2002](#)), but still does not explain the non-linearity in the magnitude of return fluctuations without introducing arbitrary factors.

Instead of relying on a stochastic shock to interest rates, this paper views the FRC as the distribution of cash flows over time. Specifically, it reflects the price of earning a constant rate of interest at various maturities. Borrowing from a literature in statistics that studies the scaling of random variables ([Frank 2018](#)), I find that the FRC is well modelled by the shape of a Gamma distribution. This means that rates at the front end scale geometrically with maturity, but linearly at the back end. This can be easily interpreted economically as a

representative bank that discounts both hyperbolically, as well as exponentially.

Structure Section 7.2 presents key patterns in the data for Eurodollar futures, namely the shape of the FRC, its evolution and the volatility of its fluctuations conditional on the future's maturity. Section 7.3 develops the model, for which Section 7.4 offers parameter estimates using the data from Section 7.2. Section 7.5 concludes.

7.2 Data

This paper uses daily data on Eurodollar (ED) futures downloaded from the Chicago Mercantile Exchange (CME). Additional data on three month USD LIBOR rates are downloaded from IBORate. The historical panel begins on the 19th of March, 2013, and ends on the 28th of February, 2020.

I give a brief description of interest rate futures. Readers familiar with ED futures may skip ahead to the features of the data relevant to this paper.

7.2.1 Overview of forward rates and futures

Price P quoted for an ED future contract (which I refer to as 'future' for the rest of the paper) is a function of maturity m ,

$$P_t(m) = 1 - f_t(m),$$

where $f_t(m)$ denotes the International Money Market (IMM) index, m is the maturity of the future, and t indexes time. For simplicity, I assume that the size of each contract is \$1, although the actual price quoted by CME is 100 times $P_t(m)$. The IMM index, henceforth referred to as the *forward rate*, is meant to follow the three month USD London Interbank Offered Rate (LIBOR). Rate $f_t(0)$, whose maturity is at time t , must equal the LIBOR, and is referred to as the *spot rate*.

After m periods, the future expires, and is settled between the receiver (buyer) and payer (seller) of the future. This usually involves a cash transaction between the two counterparties, with the payment depending on the realised spot rate $f_{t+m}(0)$. If the spot rate is lower than the contract rate, the payer pays the receiver the difference $f_t(m) - f_{t+m}(0)$, assuming a contract size of \$1. If, instead, the realised rate is higher than the contract rate, the receiver pays the payer $f_{t+m}(0) - f_t(m)$.

Who needs interest rate futures? The following examples will help understand the mechanisms and motivations that govern these futures. Suppose that a bank wants to lock in the interest rate on a loan it plans to give at time $t + m$. The purpose of the future is to let the bank fix the interest rate to $f_t(m)$ at time t . In any case, the loan it issues at $t + m$ earns the realised spot rate $f_{t+m}(0)$. If the bank decided to purchase the future, it will earn, or owe, additional money to the payer when settling the future contract. The bank earns $f_t(m) - f_{t+m}(0)$ if this realised spot rate is lower than the contract rate. If the spot rate is higher, the bank

instead pays $f_{t+m}(0) - f_t(m)$. In both instances, the profit earned by the bank equals $f_t(m) -$ the rate it decided to lock in at time t .

The payer in this example may have been motivated to expose themselves to interest rate risk, for example because they have knowledge at time t of what the spot rate $f_{t+m}(0)$ will be. Alternatively, they may have decided to lock in a rate for an amount they seek to borrow. Suppose, now, that the bank plans to take such a loan at $t + m$, instead of issuing one, so it decides to sell the same future, making it the payer. At time $t + m$, the bank will borrow money on which it will owe the realised rate $f_{t+m}(0)$. This time, the bank earns $f_{t+m}(0) - f_t(m)$ from the receiver if the realised rate is higher. On the other hand, the bank pays $f_t(m) - f_{t+m}(0)$ to the receiver if the realised rate is lower. In both cases, the total cost of borrowing at time $t + m$ will be $f_t(m)$, instead of an uncertain rate $f_{t+m}(0)$.

For banks, pension funds and national governments that rely on trillions of dollars in financing, the availability of these instruments is invaluable.

7.2.2 Forward Rate Curve

For notational convenience, I re-scale the raw data for future prices to reflect the continuously compounded rate earned from a loan that charges interest equal to IMM index $f_t(m)$:

$$p_t(m) = e^{-f_t(m)}.$$

This way, the earnings to the receiver issuing a loan at time $t + m$ is always \$1, and the price of that dollar at time t is $p_t(m)$. As before, $f_t(m)$ is the forward rate quoted by the exchange.

The set of forward rates available on the market at time t constitutes the *Forward Rate Curve* (FRC). In practice, 44 or 45 unexpired Eurodollar futures are available on a given day. The first six contracts expire on a monthly basis, and the remaining expire on a quarterly basis. The longest dated contract expires in roughly ten years' time. Further details are provided in Appendix 7.A.

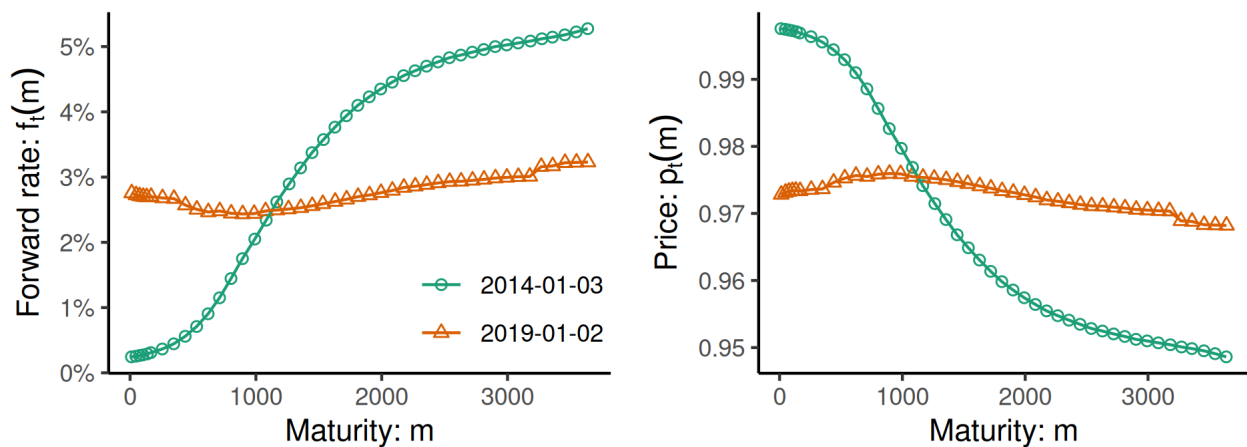


Figure 7.2.1: **FRC generally slopes upwards, and prices downwards:** This figure plots the observed FRC and corresponding prices for two arbitrary dates. The plots demonstrate the inverse relationship between the forward rate and the corresponding contract price.

The key benefit of these products is their liquidity, which allows for daily observations on the entire FRC. Two examples are illustrated in Figure 7.2.1, for the third of January, 2014, and the second of January, 2019. Each dot represents a future, with its maturity, in days, graphed on the horizontal axis, and its closing forward rate (Left) and price (Right) on the vertical axis.

Two features stand out in Figure 7.2.1. First, the FRC is generally increasing in slope. The overall range of forward rates evolved substantially between 2013 and 2020, as illustrated by the declining forward rate for the longest-maturity contract in Figure 7.2.2. Besides a brief moment where the spot rate exceeded the longest-maturity contract's forward rate, the FRC does not slope downwards for extended periods of time.

Second, the relationship between forward rates and maturities is not necessarily linear, and demonstrates a pronounced deviation near the mid-point of the curve. This pattern is reflected in the hump-shaped relationship between future prices and maturities in Figure 7.2.1 (Right).

7.2.3 Fluctuations along the FRC

While the shape of the FRC can change dramatically over the course of years, daily fluctuations also reveal some important features. I give a brief description of daily FRC fluctuations by decomposing its variance by the variance of the spot rate, and the contribution from the variance in the spread between forward rates and the spot rate.

The log return to holding a future with maturity m is

$$\begin{aligned}\Delta \log p_t(m) &= -\Delta f_t(m) \\ &= -[\Delta f_t(0) + \Delta(f_t(m) - f_t(0))] \\ &= -[\Delta f_t(0) + \Delta s_t(m)],\end{aligned}\tag{7.1}$$

where $s_t(m)$ is the *spread* between the forward rate at maturity m and spot rate $f_t(0)$, and Δ is the difference operator, $\Delta x_t \equiv x_t - x_{t-1}$.

Figure 7.2.2 plots the evolution of the LIBOR and the spread on the highest maturity ED future. The first striking element is the degree to which the spread fluctuates more than the LIBOR. Even though the Federal Reserve fixes the spot rate, there is still a significant amount of volatility along the rest of the FRC. Second, it demonstrates how the FRC flattened considerably over the course of the observed sample, which is mostly accounted for by the rise in the LIBOR starting 2016. In the meantime, the ten year rate remained relatively stable, to the point that spreads dipped below zero in 2019.

However, the LIBOR and ten year rates only constitute two points on the entire FRC. To get a better picture, I decompose the variance in daily fluctuations of each point on the FRC using Eq. 7.1:

$$\sigma^2[\Delta f_t(m)] = \sigma^2[\Delta f_t(0)] + \sigma^2[\Delta s_t(m)] + 2\sigma[\Delta f_t(0), \Delta s_t(m)],\tag{7.2}$$

where $\sigma^2[x]$ is the variance of x , and $\sigma[x, y]$ the covariance between x and y . The total

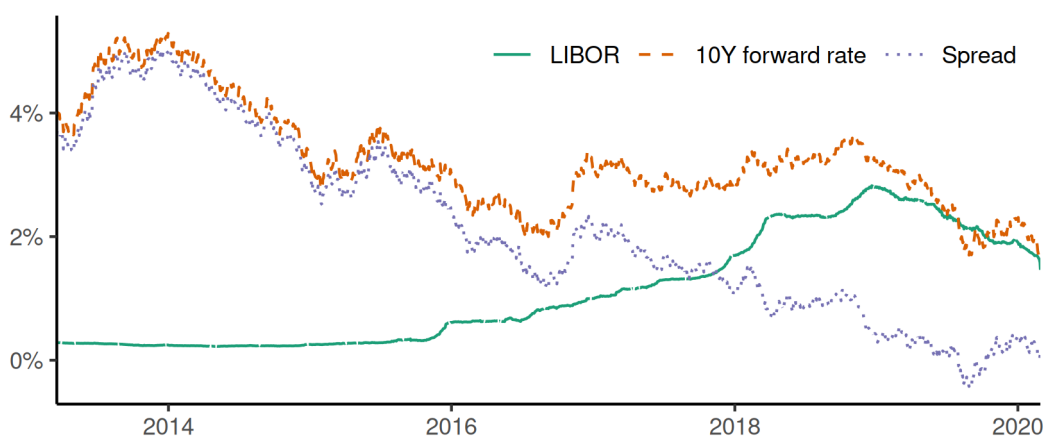


Figure 7.2.2: **Longest maturity (~10 years) forward rate fluctuates more than the spot rate:** This figure plots the time series for the LIBOR rate (solid green), the forward rate for the highest maturity ED future (dashed orange), and the difference – *spread* – between the two (dotted purple), on a daily frequency. The highest maturity future usually expires at a maturity of ten years, and is therefore labelled as the ‘10Y future’. LIBOR data made available by IBORate.

variation in daily forward rate fluctuations is the sum of variation in spot rate fluctuations, variation in spreads, and two times the covariance of spot rate and spread fluctuations.

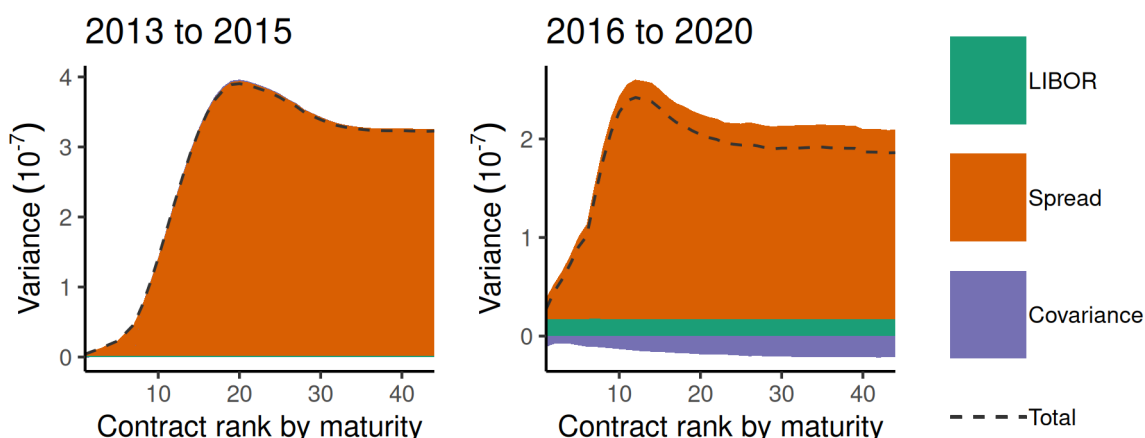


Figure 7.2.3: **Daily changes in forward rates are mostly explained by fluctuations in spreads:** This figure plots the volatility of forward rates (vertical axis) for future contracts, ranked by maturity (horizontal axis). The rank of a future is determined by its location in the list of futures ordered by maturity, from smallest to largest, on each day. Volatilities, denoted by the dashed line, are computed for the period 20/03/2013 to 31/12/2015 on the left, and 01/01/2016 to 28/02/2020 on the right. The contributions to forward rate volatility are attributed to the LIBOR (green) versus spreads (orange), and their covariance (purple), according to Eq 7.2.

I use Eq. 7.2 to decompose the variance in daily fluctuations of the FRC for two periods that characterise the sample period: 2013 to 2015, where spot rates were limited by the Zero Lower Bound (ZLB), and 2016 to 2020, where the Federal Reserve gradually lifted its target rates. I use daily observations of the LIBOR to reflect changes in the spot rate.

Figure 7.2.3 plots the results of the variance decomposition. The first striking feature is

that movements in the FRC were almost entirely driven by fluctuations in spreads. The low contribution from fluctuations in the LIBOR is not too surprising, given i) the ZLB and ii) a strict forward guidance policy by the Federal Reserve.

A second interesting feature is the hump-shaped pattern in the forward rate variance by future maturity. Roughly speaking, futures expiring in two to three years fluctuate more than at other points along the curve.

7.2.4 Level, slope and curvature

Ultimately, these results point to the three, well-known components to the FRC: its level, determined by the spot rate, its slope, determined by spreads on long-maturity futures, and its curvature, determined by spreads for futures expiring in roughly two to three years.

These patterns are simultaneously present in daily snapshots of the FRC, as well as the magnitude of its daily fluctuations. Many statistical models exist which describe those three components. The main contribution of this paper is to offer more economic insight into the behaviour of the FRC, as well as possible explanations for the extreme fluctuations observed in the belly of the curve.

7.3 Model

The purpose of this section is to predict the shape and evolution of the FRC. The two desirable features I reproduce from the data are i) the slope in forward rates and ii) the peak in the magnitude of forward rate fluctuations. Section 7.4 evaluates the fit of the model to the data.

I consider a market for futures, where a representative bank is tasked with maximising its interest income earned on household deposits. The key inputs to the model are the duration of household deposits, and the rate of return earned by the bank. Given a set of initial conditions, the bank is tasked with re-arranging its portfolio of futures to match changes in duration, subject to transaction costs.

7.3.1 Notation and setup

The model includes an initial period and final period, $t - 1$ and t , to describe the FRC composed of M futures traded at market prices $p_t(m)$. In practice, the future is assigned to a margin account, and a cash transaction occurs on the settlement date between the payer and the receiver. To simplify the exposition, I assume that the bank purchases a future with maturity m for the full price of $p_t(m)$, which translates to the bank providing full margin on their trade. Furthermore, I assume that prices are continuous and differentiable, but develop the model using discrete maturities. Discrete maturities are arguably more appropriate, since futures expire on a monthly or quarterly basis, and they specifically price LIBOR rates on three-month deposits. Much of the literature on fixed income derivatives relies on continuous time methods.

Key variable: future prices Future prices $p_t(m)$ are the key quantities the bank seeks to control. These represent the prices of earning one dollar for sure, at times $t + m$. On aggregate, future prices determine both the interest accrued on a loan issued at a future date, but also measure the average holding of the representative bank. This means that the sum of prices

$$\sum_{m=1}^{M-1} p_t(m) = \mathcal{P}_t$$

is both the price of a contract that pays one dollar in every period, as well as the net asset value of the bank. I develop this intuition further in Appendix 7.B.1. Quantity $\mathcal{P}_t$ is also referred to as the *money market account*.

Before proceeding, I assume that the spot rate $p_t(0)$ is fixed exogenously. In practice, this rate is set by the central bank, and it is therefore unreasonable to endogenise it here. This assumption implies that the system effectively governs the interest rate spreads at each maturity $t + m$.

Between time $t-1$ and t , the aggregate money market account changes by a fixed amount

$$\Delta \mathcal{P}_t = \sum_{m=1}^{M-1} \Delta p_t(m), \quad (7.3)$$

where $\Delta x_t \equiv x_t - x_{t-1}$ is the difference operator. In order to measure fluctuations along the FRC, I constrain the second moment of future price changes $\mathcal{V}_t^2$ to

$$\mathcal{V}_t^2 = \sum_{m=1}^{M-1} \frac{[\Delta p_t(m)]^2}{p_{t-1}(m)}, \quad (7.4)$$

where initial prices $p_{t-1}(m)$ serve as weights.

This setup is a general method to decompose temporal fluctuations in a cross-section of heterogeneous variables, as demonstrated in Frank (2018). Eq. 7.3 represents the total amount of liquidity availability in the system, and can be interpreted as the net asset value of a representative bank, as per Appendix 7.B.1. Eq. 7.4 is similar to a constraint on transaction costs in portfolio management problems. In asset pricing terms, it regulates the amount of cross-sectional variation in future prices, which will be important to decompose the different sources of risk to future prices.

Key targets: returns versus duration Given the constraints on future prices, I express three further quantities that govern the evolution of the FRC, founded in behavioural assumptions. The first behavioural assumption is that the system seeks to maximise its prospec-

tive Return on Assets (RoA) $\mathcal{R}_t$, which equals

$$\begin{aligned}\mathcal{R}_t &= \sum_{m=1}^{M-1} p_t(m) \log \frac{B_t(m)}{B_t(m+1)} \\ &= - \sum_{m=1}^{M-1} p_t(m) \log p_t(m),\end{aligned}\tag{7.5}$$

where $B_t(m)$ is the price of a bond that pays \$1 at time $t + m$ (see Appendix 7.B.1 for details). Eq. 7.5 outlines the impact of a hypothetical bank re-organising its bond portfolio on its prospective RoA. If bond prices are exogenous, it would park its funds at the particular spot in the FRC – defined by the series of future prices $-\log p_t(m) = f_t(m)$ – that earns the highest return. Future prices also reflect the net asset value of the entire banking system, which offers an aggregate ‘annuity’ paying one dollar every period $t + m$.

Extending this logic leads to the second behavioural constraint, which regulates the duration $\mathcal{D}_t$ of the cash flows:

$$\mathcal{D}_t = \sum_{m=1}^{M-1} p_t(m)m.\tag{7.6}$$

Duration $\mathcal{D}_t$ is high when the front end of the FRC is cheap, and the back end is expensive. Conversely, duration is low when the front end is expensive, but the back end cheap. Duration effectively measures the timing with which the aggregate annuity returns the ‘principal’ amount of hypothetical deposits, as opposed to returning ‘interest’, because prices $p_t(m)$ reflect the net asset value of loans made at time t .

In the context of the hypothetical bank, this can be interpreted as the timing with which deposits are repaid. Recall that future prices reflect the net asset value of the bank, which must ultimately equal the value of deposits. An interesting insight, which I don’t develop here, is that duration is relevant in setting the FRC if the bank faces uncertainty in the time it takes for depositors to withdraw their funds. If the bank only knows the *average* time – the duration – of its deposits, it will seek to allocate its bond portfolio so that it can meet withdrawals *on average*. To better develop this framework would require additional details on capital requirements, but detracts from the main point of the paper: when investing in long-term projects, the banking sector is constrained in the profits it can generate, determined in Eq. 7.5, because it must meet potential withdrawals in its deposits. The *rate* at which these withdrawals take place, determined by duration in Eq. 7.6, will regulate the distribution of assets as a function of their maturity, and, thus, the shape of the FRC.

For the third behavioural assumption, I note that the data presented in Section 7.2 suggest that the FRC does not only scale linearly with maturity, but also geometrically. As a result, I propose a third behavioural assumption that constrains the geometric duration of

the FRC:

$$\mathcal{G}_t = \sum_{m=1}^{M-1} p_t(m) \log m. \quad (7.7)$$

The economic meaning of this constraint will become more clear towards the end of this section. It effectively translates into a hyperbolic discount factor, in contrast to duration in Eq. 7.6 which is shown to reflect an exponential discount factor. This geometric duration will introduce curvature into the FRC, which is an important component when estimating the model empirically.

7.3.2 Dynamics of the FRC

This subsection studies the volatility of future prices as a function of the key aggregates defined in Section 7.3.1. Since volatility depends on variation in price changes $\Delta p_t(m)$, it is necessary to relate these price dynamics to the evolution of those aggregates: the money market account, the RoA and durations (linear and geometric).

Updates to future prices $\Delta p_t(m)$ impact the prospective RoA $\mathcal{R}_t$. The partial change to the RoA is

$$\begin{aligned} \partial \mathcal{R}_t &= \sum_{m=1}^{M-1} \Delta p_t(m) \log \frac{p_t(m)}{p_{t-1}(m)} \\ &= - \sum_{m=1}^{M-1} \Delta p_t(m) \log p_{t-1}(m) + \sum_{m=1}^{M-1} \Delta p_t(m) \log p_t(m). \end{aligned} \quad (7.8)$$

I focus on this partial change in RoA because it only considers the additional returns that stem from re-pricing the entire FRC by $\Delta p_t(m)$. In principle, total returns would also include an additional level effect, but it is negligible when changes in returns $p_t(m)/p_{t-1}(m)$ are small.

The first term in Eq. 7.8 expresses the gain in the RoA thanks to updating prices by $\Delta p_t(m)$, but the second term reflects a quantity change that stems from a change in average prices between periods t and $t-1$. To quantify the impact of this second term on the change in the RoA, I study this quantity for the FRC at equilibrium. Definition 7.1 proposes the equilibrium condition for the FRC, for which I derive an expression in Lemma 7.1.

Definition 7.1 (FRC equilibrium). *The FRC is in equilibrium when it maximises RoA $\mathcal{R}_t$, whilst satisfying constraints on duration $\mathcal{D}_t$ and $\mathcal{G}_t$ and the money market account $\mathcal{P}_t$.*

Lemma 7.1. *Prices $p_t(m)$, which constitute the FRC, are in equilibrium when*

$$\log p_t(m) = -1 - \mu - \beta m - \gamma \log m, \quad (7.9)$$

where μ , β and γ are constants.

Proof. See Appendix 7.B.2. □

Lemma 7.1 gives the eventual functional form for the FRC, which will be useful later. On its own, it does not help study fluctuations in the FRC. This is because changes in prices $p_t(m)$ must stem from changes in parameters μ , β and γ , as is evident in Eq. 7.9. However, Lemma 7.1 leads to the following expression for the change in RoA:

$$\partial \mathcal{R}_t = - \sum_{m=1}^{M-1} \Delta p_t(m) \log p_{t-1}(m) - (1 + \mu) \Delta \mathcal{P}_t - \beta \Delta \mathcal{D}_t - \gamma \Delta \mathcal{G}_t, \quad (7.10)$$

where changes in the money market account and duration, both linear and geometric, act as constraints on the degree to which the RoA can increase. Importantly, any changes in these aggregates must stem from changes in the FRC. Taking duration $\mathcal{D}_t$ as an example,

$$\Delta \mathcal{D}_t = \Delta \sum_{m=1}^{M-1} p_t(m) m = \sum_{m=1}^{M-1} \Delta p_t(m) m,$$

so that duration can only change with the re-pricing of futures.

The second term of Eq. 7.8 thus embodies a quantity change in the RoA triggered by changes in either the aggregate money market account, or the durations. In contrast, the second term resembles *entropy production*: the degree to which interest income is increasingly uniform over maturities m . Chapter 6 demonstrates how changes in distribution and entropy can affect aggregates. Chapter 8 applies this concept to the behaviour of stock prices.

Eq. 7.10 states the impact of important aggregates that feature in the money market on fluctuations in the FRC. The degree to which the FRC fluctuates, and where, can be solved for by introducing the measure for volatility from Eq. 7.4. Taking the constraints in Eq. 7.10 as given, I can now write the Lagrangian function for volatility $\mathcal{V}_t^2$ in terms of constraints on the RoA, duration and the money market account:

$$\begin{aligned} \mathcal{L} = & \sum_{m=1}^{M-1} \frac{[\Delta p_t(m)]^2}{p_{t-1}(m)} - 2\kappa \left(- \sum_{m=1}^{M-1} \Delta p_t(m) \log p_{t-1}(m) - \partial \mathcal{R}_t \right) \\ & - 2\beta \left(\sum_{m=1}^{M-1} \Delta p_t(m) m - \Delta \mathcal{D}_t \right) \\ & - 2\gamma \left(\sum_{m=1}^{M-1} \Delta p_t(m) \log m - \Delta \mathcal{G}_t \right) \\ & - 2\mu \left(\sum_{m=1}^{M-1} \Delta p_t(m) - \Delta \mathcal{P}_t \right), \end{aligned} \quad (7.11)$$

where I re-scale multipliers β , γ , and μ for convenience. This expression effectively assigns cross-sectional volatility in the FRC to changes in various aggregate quantities, namely i) the RoA $\partial \mathcal{R}_t$, ii) linear duration $\Delta \mathcal{D}_t$, iii) geometric duration $\Delta \mathcal{G}_t$, and iv) the money market account $\Delta \mathcal{P}_t$. Since the second derivative of Eq. 7.11 is positive, the Lagrangian serves to minimise volatility. Effectively, the task for the representative bank in this system is to min-

imise the re-allocation of its assets, for example to avoid transaction costs. It is compelled to re-allocate some amount due to changes in certain aggregates, which double as benchmarks it needs to meet: its ability to generate returns, to repay deposits and manage liquidity.

Proposition 7.1 offers a solution to the programme with an expression for the returns on futures, denoted

$$r_t(m) \equiv \frac{\Delta p_t(m)}{p_{t-1}(m)}.$$

Proposition 7.1 (FRC dynamics). *FRC volatility $\mathcal{V}_t^2$ is at its minimum when*

$$r_t(m) = \mu - \kappa \log p_{t-1}(m) + \beta m + \gamma \log m, \quad (7.12)$$

where the vector of parameters

$$\begin{bmatrix} \mu & \kappa & \beta & \gamma \end{bmatrix}' = (\mathbf{X}'\mathbf{\Omega}\mathbf{X})^{-1} \mathbf{X}'\mathbf{\Omega}\mathbf{R}, \quad (7.13)$$

an apostrophe denotes a matrix transpose, $\mathbf{X}$ is the $(M-1) \times 4$ matrix of values for $-\log p_{t-1}(m)$, m and $\log m$, and ones in the first column, and m incrementing by one in every row of $\mathbf{X}$,

$$\mathbf{X} = \begin{bmatrix} 1 & -\log p_{t-1}(1) & 1 & 0 \\ 1 & -\log p_{t-1}(2) & 2 & \log(2) \\ 1 & -\log p_{t-1}(3) & 3 & \log(3) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & -\log p_{t-1}(M-1) & M-1 & \log(M-1) \end{bmatrix}, \quad (7.14)$$

matrix $\mathbf{\Omega}$ is a diagonal $(M-1) \times (M-1)$ matrix of weights $\omega_{t-1}(m)$,

$$\mathbf{\Omega} = \begin{bmatrix} \omega_{t-1}(1) & 0 & 0 & \dots & 0 \\ 0 & \omega_{t-1}(2) & 0 & \dots & 0 \\ 0 & 0 & \omega_{t-1}(3) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \omega_{t-1}(M-1) \end{bmatrix}, \quad (7.15)$$

and $\mathbf{R}$ is the rank $(M-1)$ vector of future price returns $r_t(m)$,

$$\mathbf{R}' = \begin{bmatrix} r_t(1) & r_t(2) & r_t(3) & \dots & r_t(M-1) \end{bmatrix}'. \quad (7.16)$$

Proof. See Appendix 7.B.3. □

Proposition 7.1 demonstrates how the FRC fluctuates as a function of three aggregates: the money market account, the RoA on fixed income, and durations. The solution is a linear model, where returns on futures are a function of the future's characteristics – its previous price, maturity and log maturity – and returns on those characteristics. This functional form is strongly reminiscent of typical asset pricing models. I use this methodology to reconstruct factor models for stock price returns, where market capitalisation stands out as

an important characteristic, in Chapter 8.

One difference to typical models is that Eq. 7.35 includes parameters weighted by future prices in Ω . In particular, movements in key aggregates governing the FRC rely more heavily on changes at the front end of the curve. Given the motivation in Section 7.3.1, this is because the front end is the more valuable part of the curve, in terms of net assets available to banks. It is natural that matters concerning profitability, and commitments to repay deposits, should therefore influence benchmarks in the near term to a greater degree.

Under classical assumptions, the parameters in Eq. 7.35 are straightforward to estimate using historical ED data. The FRC dynamics suggested by Proposition 7.1 are also instructive regarding the behaviours outlined in Section 7.2, namely i) the shape of the FRC and ii) the maximal fluctuations in the FRC.

7.3.3 FRC equilibrium

The equilibrium condition for the FRC, stated in Lemma 7.1, can be seen in Eq. 7.12 of Proposition 7.1 by setting all returns to zero, $r_t(m) = 0$. It is the set of points where the system does not seek to re-price futures, since it already satisfies all constraints.

It is difficult to solve for the parameters in Lemma 7.1 using the discrete system for future prices. However, finding some solution would help to better understand the role different constraints play on the shape of the FRC.

Continuous maturity I proceed by introducing a continuous scale for maturity $m > 0$, extending to infinity. Since the solution concerns a continuous version of the system at equilibrium, I drop time subscripts t . To simplify the problem further, I set the money market account to one:

$$\int_0^\infty p(m)dm = 1.$$

These assumptions yield an alternative interpretation for prices $p(m)$, namely that of a probability measure. It offers an interesting application in the vein of the preceding discussion: measure $p(m)$ denotes the assets available should all deposits be withdrawn, simultaneously, at maturity m .

It is also interesting to consider the intuition behind the RoA in the continuous maturity version:

$$\mathcal{R} = - \int_0^\infty p(m) \log p(m) dm, \quad (7.17)$$

where $0 \log 0$ equals zero. Similarly to the discrete version of the model, the RoA mimics the role of entropy in the study of statistical distributions. In addition, it can also be seen as the number of ways a representative bank can pay out a unit of deposits. I leave the precise question of leverage and default for future research, and motivate the maximisation of the RoA as desirable for the sake of economic efficiency.

The usual constraints on durations are similarly stated as

$$\mathcal{D} = \int_0^\infty p(m) m dm, \quad (7.18)$$

$$\mathcal{G} = \int_0^\infty p(m) \log m dm, \quad (7.19)$$

and offer the same interpretation as in the discrete version of the model.

Equilibrium FRC shape Lemma 7.2 derives the equilibrium FRC, following Definition 7.1, with continuous maturity.

Lemma 7.2. *Prices $p(m)$ are in equilibrium when*

$$p(m) = \frac{\beta_g^{\gamma_g+1}}{\Gamma(\gamma_g+1)} m^{\gamma_g} e^{-\beta_g m}, \quad (7.20)$$

where

$$\mathcal{D} = \frac{\gamma_g + 1}{\beta_g},$$

$$\mathcal{G} = -\log \beta_g + \psi(\gamma_g + 1),$$

and $\psi(\cdot)$ is the digamma functions

Proof. A simple proof is available in [Singh & Singh \(1985\)](#). □

Eq. 7.20 is the function for the Gamma distribution, well known in statistics and biology ([Frank 2009](#)). It is the function that maximises the RoA, or entropy, subject to constraints on linear and geometric means, $\mathcal{D}$ and $\mathcal{G}$. It features the geometric scaling of prices $p(m)$ when maturity m is low, but exponential scaling when maturity m is high. This is best seen by taking logarithms,

$$\log p(m) = \log \left(\frac{\beta_g^{\gamma_g+1}}{\Gamma(\gamma_g+1)} \right) + \gamma_g \log m - \beta_g m, \quad (7.21)$$

and noting that the first derivative,

$$\frac{d \log p(m)}{dm} = \frac{\gamma_g}{m} - \beta_g,$$

varies in proportion to γ_g when $m \rightarrow 0$, but in proportion to β_g when $m \rightarrow \infty$.

Special cases The equilibrium shape of the FRC under separate constraints provides further insights. I derive distributions for prices $p(m)$ by either constraining linear duration $\mathcal{D}$, or geometric duration $\mathcal{G}$, separately. The distribution that maximises the RoA $\mathcal{R}$ with a

constraint on mean $\mathcal{D}$ alone is the Exponential distribution

$$p(m) = \beta_e e^{-\beta_e m}, \quad (7.22)$$

where $\beta_e = 1/\mathcal{D}$. Alternatively, the distribution that maximises the RoA $\mathcal{R}$ with a constraint on mean $\mathcal{G}$, instead, is the Pareto distribution

$$p(m) = \frac{\gamma_p k_p^{\gamma_p}}{m^{\gamma_p+1}}, \quad (7.23)$$

where $1/\gamma_p = \mathcal{G} - \log k_p$, and parameter k_p is a constant.

The relationship between these functions and standard models of discounting from economics is best seen by comparing the prices of two annuities. I construct a discount factor for a constant stream of income until time $m = M$ by integrating the exponential distribution,

$$\int_0^M p(m) dm = 1 - e^{-\beta_e M},$$

and the Pareto distribution,

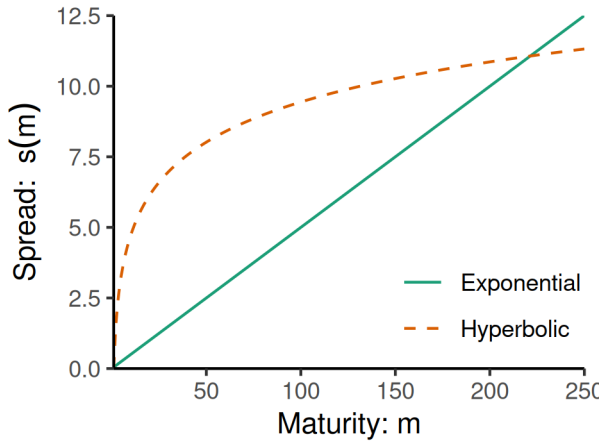
$$\int_k^M p(m) dm = 1 - \left(\frac{k}{M}\right)^{\gamma_p}.$$

These expressions contrast the rate at which an annuity is discounted: by rate β_e in the Exponential distribution, compared to rate γ_p in the Pareto distribution. In economics, the exponential case is referred to as *exponential* discounting, whereas the Pareto case is referred to as *hyperbolic* discounting. The Gamma distribution yields a similar discounting function, that nests both exponential as well as hyperbolic discounting.

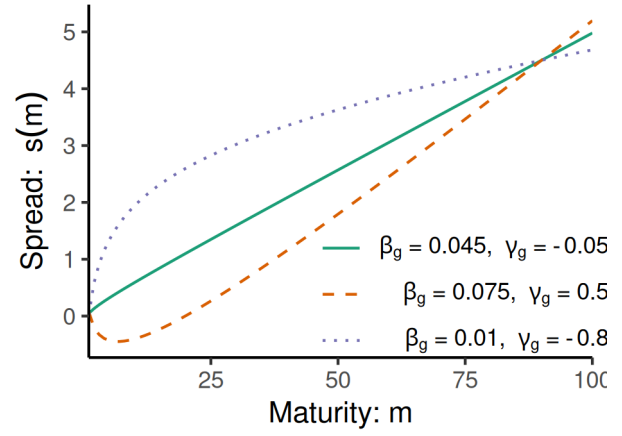
Shapes of the FRC Figure 7.3.1a plots the forward rate spreads for the Exponential distribution, $s(m) = \beta_e m$, corresponding to exponential discounting, and the Pareto distribution, $s(m) = -\log \gamma_p + (1 + \gamma_p) \log m$, corresponding to hyperbolic discounting.¹ I choose parameters $\beta_e = 0.05$ and $\gamma_p = 19/20$ that yield the same linear duration $\mathcal{D} = 20$, to keep both functions comparable. Figure 7.3.1a demonstrates the larger spreads at the front end under hyperbolic discounting, which do not change much at the back end of the FRC. In contrast, spreads increase by a constant amount throughout the FRC under exponential discounting. The economic intuition, reproduced by Figure 7.3.1a, is that hyperbolic discounting places more weight on income streams in the near future, relative to exponential discounting.

Figure 7.3.1b plots the spreads when discounting takes the form of the Gamma distribution. I pick parameter values for β_g and γ_g that fix the linear duration $\mathcal{D}$ to 20, as in Figure 7.3.1a. When the hyperbolic parameter γ_g is small, as per the solid green line in Figure 7.3.1b, the FRC's slope is most similar to the exponential distribution. Alternatively,

¹I produce functions for the spreads by taking forward rates $f(m) = -\log p(m)$, and subtracting a spot rate $s(m) = f(m) - f(0)$ so that the spread near the spot rate is zero, $s(0) = 0$.



(a) Exponential and hyperbolic discounting



(b) Gamma discounting

Figure 7.3.1: FRC shapes under different discount functions: This figure compares the hypothetical interest rate spreads (vertical axis) as a function of maturity (horizontal axis) for an exponential discount function (solid green) to a hyperbolic discount function (dashed orange) on the left. The exponential discount function is evaluated from Eq. 7.22 with parameter β_e set to 0.05, while the hyperbolic function is evaluated from Eq. 7.23 with parameter γ_p set to 19/20. The right figure plots spreads under ‘gamma’ discounting, which manifests as an average of exponential and hyperbolic discounts. Parameter choices used to evaluate the gamma discount function from Eq. 7.21 are listed in the legend. All intercepts are set to zero.

when parameter γ_g is large, such as the dotted purple line in Figure 7.3.1b, the FRC shape is more closely related to that from the Pareto distribution. This demonstrates how *gamma* discounting offers an average of both exponential and hyperbolic discounting, weighted according to parameters γ_g and β_g .

An interesting feature is that gamma discounting admits *negative* hyperbolic discounting. In other words, the hyperbolic portion of the function prefers currency in the future, rather than in the present. This is possible because the exponential part of the discount function eventually dominates the hyperbolic part, and the whole FRC remains integrable. This feature is demonstrated by the dashed orange line in Figure 7.3.1b: the positive rate γ_g induces the FRC to slope downwards at small maturities, causing a dip in its shape.

Summary The model’s behaviour at equilibrium offers interesting insights into the possible shapes of the FRC. While the mechanism behind a positive slope is generally understood to be some form of a discount function, the mechanism behind the FRC’s curvature has not been explored. In this paper, curvature originates in the balancing between exponential and hyperbolic discounting, in a manner I refer to as ‘gamma’ discounting because it mimics the behaviour of a Gamma distribution. Most interesting is the manner in which curvature can cause the FRC to slope downwards at small maturities, when the hyperbolic discounting portion of the gamma function is negative. In light of the motivation in Section 7.3.1, this may take place when banks expect the geometric mean of asset durations to be negative. This goes a long way to explain the patterns for the FRC observed in Figure 7.3.1.

7.3.4 Extreme fluctuations

The gamma function implies that volatility in the FRC scales in a non-trivial manner. This is directly relevant to the observation in Section 7.2 that fluctuations in the FRC peak at a maturity of three to five years.

Proposition 7.1 leads to an expression for the peak in FRC fluctuations with respect to maturity. I state this result in Corollary 7.1.1.

Corollary 7.1.1 (Extreme FRC fluctuation). *Fluctuations in future prices $p_t(m)$ are maximal at maturity $0 < m < M - 1$ if $\gamma > 0$ and $\beta < 0$. The largest fluctuation occurs at maturity*

$$m^* = -\frac{\gamma}{\beta}. \quad (7.24)$$

If $\gamma < 0$ and $\beta > 0$, m^* denotes a minimum.

Proof. See Appendix 7.B.4. □

Corollary 7.1.1 implies that the maximum fluctuation in future prices, relative to the average, occurs at a maturity greater than zero when changes in linear and geometric aggregate durations take different signs. Intuitively, this is where these changes cancel out, geometric duration declines while linear duration increases, and $r_t(m^*)$ is the smallest along the FRC. Alternatively, these changes amplify the price return at m^* , when the geometric duration increases and the linear duration decreases. In both of these cases, the largest changes in future prices along the FRC occur at a point m^* that is greater than zero, and less than the final maturity $M - 1$.

The tradeoff between geometric and linear duration is important to the gamma function, and is also present in the equilibrium price levels for the FRC in Eq. 7.21. It suggests that there is an important meaning behind the manner in which the FRC, and its fluctuation, behave at a time horizon of three to five years. I hypothesise that it may relate to the presence of traders, alongside banks, who act under different constraints, but restrict themselves to short time horizons. This constitutes an interesting question I leave to future research.

7.3.5 Model predictions

The model in Section 7.3 does not consider any restrictions on arbitrage. The main motivation is that the biggest participants in fixed income securities are bankers, and not traders. The main advantage in taking this view is that it uncovers interesting insights about the FRC that previous models, based largely on non-arbitrage conditions, could not make. Naturally, there is a corresponding disadvantage, in that activity by traders likely plays a significant role in the data, which may weaken the empirical support of the model.

Based on the model, I make specific predictions on the shape of the FRC, as well as the magnitude of its fluctuations as a function of maturity. I state them in three hypotheses.

Hypothesis 1: FRC fluctuations follow a linear functional form,

$$r_t(m) = \mu + \kappa f_{t-1}(m) + \beta m + \gamma \log m, \quad (7.25)$$

where parameters μ , κ , β and γ are WLS coefficients, with weights $\omega_{t-1}(m) = p_{t-1}(m)/\mathcal{P}_{t-1}$.

Hypothesis 2: The FRC fits function

$$f(m) = k - \gamma_g \log m + \beta_g m, \quad (7.26)$$

where k is a constant, and parameters β_g and γ_g are WLS coefficients, with weights $\omega_g(m) = p(m)/\mathcal{P}$.

Hypothesis 3: The extreme FRC fluctuation is at maturity

$$m^* = -\frac{\gamma}{\beta}, \quad (7.27)$$

where parameters γ and β are from Eq. 7.25.

These speak directly to the patterns observed in Section 7.2. Section 7.4 evaluates how well the model can explain these phenomena, by fitting the model's parameters to available data on ED futures.

7.4 Results

This section produces parameter estimates for the model in Section 7.3, using the futures data presented in Section 7.2. I focus on testing Hypotheses 1-3, which concern, in turn, i) the source of fluctuations in the FRC, ii) the shape of the FRC and iii) the location of extreme fluctuations along the FRC. In this section, I use data on spreads $s_t(m)$ instead of forward rates $f_t(m)$, since the model makes no assumption for movements in the spot rate $f_t(0)$.

7.4.1 Explaining FRC fluctuations

According to Hypothesis 1, fluctuations in the FRC stem from the re-allocation of liquidity over maturities on two scales; one linear, and one geometric. The maturity of the average contract on a linear scale is 4.53 years, but 2.45 years on a geometric scale. While the ends of the FRC have similar weights for the first and last maturities on both the linear and geometric scales, the weights for contracts with maturities between 2.45 and 4.53 years will be negative (below average) on the linear scale, but positive (above average) on the geometric scale. Eq. 7.25 thus mimics a factor model, where returns on linear maturity β measure the performance of long-dated maturities against short-dated maturities. Similarly, the returns on geometric maturity γ measure the relative performance of maturities in the body of the curve.

I use quarterly observations for ED future spreads to estimate the parameters in Eq 7.25,

$$r_t(m) = \mu_t + \kappa_t s_{t-1}(m) + \beta_t m + \gamma_t \log m + v_t(m), \quad (7.28)$$

where μ_t , κ_t , γ_t and β_t are WLS estimates, with weights $\omega_{t-1}(m) = p_{t-1}(m)/\sum^{M-1} p_{t-1}(m)$,

and $v_t(m)$ is an idiosyncratic error term. I calculate the return in spreads $r_t(m)$ as

$$r_t(m) = \frac{p_t(m)}{p_{t-1}(m)} - 1, \quad \text{where} \quad p_t(m) = e^{-s_t(m)},$$

and $s_t(m)$ is the spread calculated in Eq. 7.1. Since maturities usually increment by one day, these returns are calculated using the previous closing price of the same future – with a constant expiry date – and maturity m takes the value of that future at time t .

The data present certain challenges on the time scales that can be used to estimate model parameters. The key issue is that the maturities available to trade on a given day are different than on the previous day; the maturity of the FRC increments by one day, or more when markets are closed several days in a row. While this may seem trivial at the back end, where maturities are in the order of thousands of days, this can distort measurements at the front end. In particular, the difference in effect, combined with a logarithmic term in the model, can introduce a bias in parameter estimates, since the ‘error’ in maturity is more significant at the front end than the back end.

Since these increments bias parameter estimates, I estimate the parameters in Eq. 7.28 on a quarterly frequency – the common frequency for expiry dates on contracts expiring beyond six months. To do so, I compute mean values for daily variables $p_t(m)$ and m , averaged over the maturity rank of the contract in a given quarter. Lags in prices $p_t(m)$ – used to calculate returns $r_t(m)$, spreads $s_{t-1}(m)$ and weights $\omega_{t-1}(m)$ – are taken from average daily prices of futures in the same maturity rank of the previous quarter.

Figure 7.4.1 presents estimated parameters on a quarterly frequency as time series. Results using daily frequencies, which are likely biased, are available in Appendix 7.C. The solid lines denote each parameter’s point estimate, while the green regions outline the 95% confidence intervals. Figures 7.4.1e and 7.4.1f report the fit of the model for all samples as a histogram, and over time.

Figure 7.4.1a presents estimates for $\hat{\beta}_t$, which measures changes in the slope of the FRC between quarters. These roughly line up with the changes in spreads illustrated in Figure 7.2.2. For example, $\hat{\beta}_t$ spikes with the increase in spreads in 2013 and late 2016, but turns negative as spreads generally decline between 2014 and 2016, then in 2019. Figure 7.4.1b demonstrates that curvature in the FRC did not change significantly around 2015 and early 2016, so fluctuations in spreads largely stemmed from changes in the linear slope coefficient. However, $\hat{\gamma}_t$ turns strongly positive in late 2019, implying that a dip emerged in the FRC. This dip is demonstrated in Figure 7.2.1, and coincides with the ‘taper tantrum’ when the Federal Reserve attempted to unwind its asset position in treasuries.

The large correction on the slope of the FRC in 2016 meant that the return on short-dated futures turned negative relative to long-dated futures. This is demonstrated by the large negative value in $\hat{\kappa}_t$ at that time in Figure 7.4.1c. Parameter κ , which reflects the changes in the profitability of holding an annuity, remains positive on average, but is clearly susceptible to occasional large negative swings. Average fluctuations in spreads unaccounted for by these three components are represented by intercept μ_t , with estimates reported in Figure 7.4.1d. These rise in early 2017 and late 2016, mimicking the peaks in spreads around

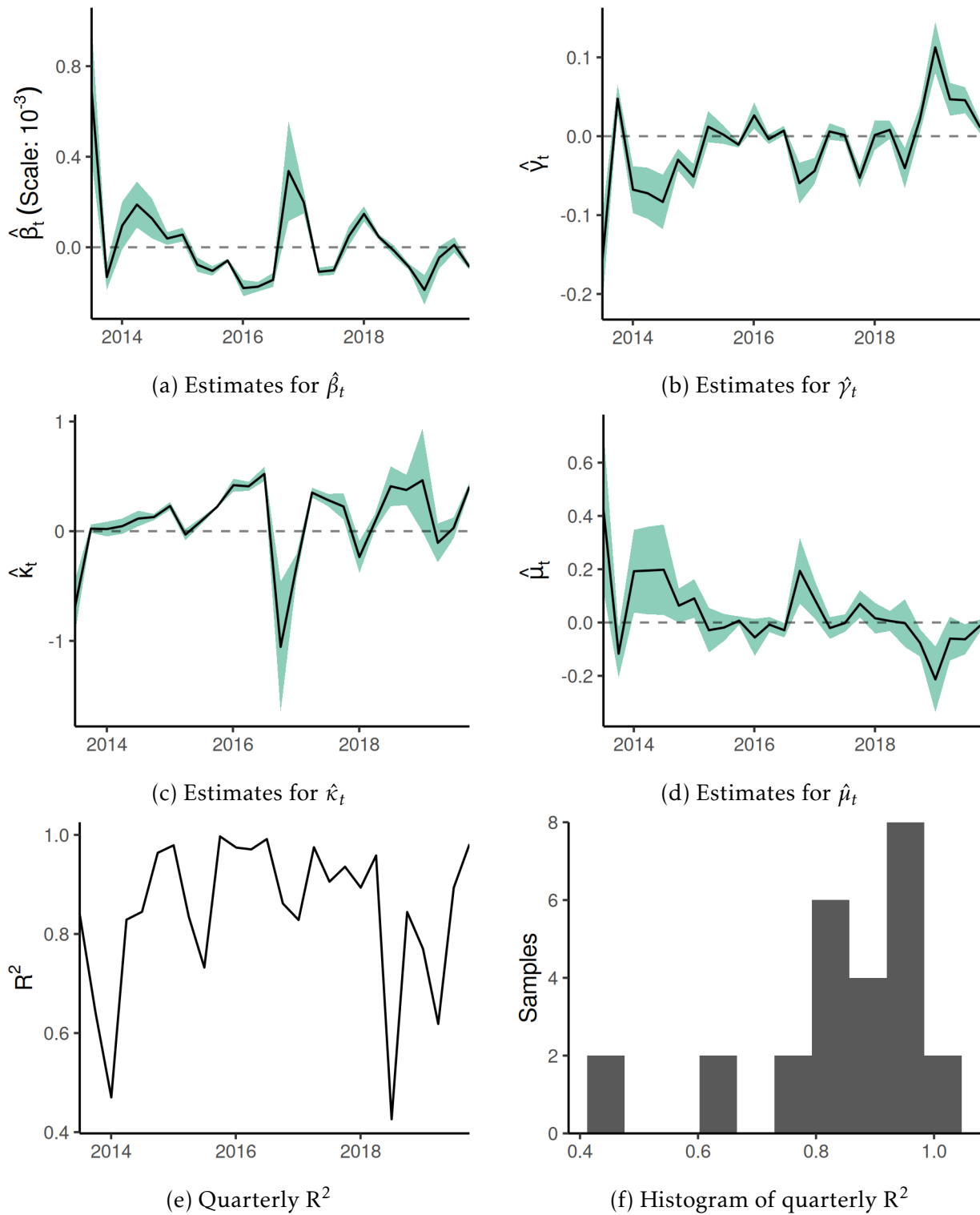


Figure 7.4.1: Quarterly estimates for model parameters suggest the FRC evolved significantly over time: This figure graphs WLS estimates for parameters $\hat{\beta}_t$ (Figure 7.4.1a), $\hat{\gamma}_t$ (Figure 7.4.1b), $\hat{\kappa}_t$ (Figure 7.4.1c) and $\hat{\mu}_t$ (Figure 7.4.1d) to Eq. 7.28. The solid black lines denote point estimates of these parameters for one quarter, with the green region outlining the 95% confidence interval. A horizontal, dashed grey line highlights the distance of point estimates from zero. Figure 7.4.1e plots the R^2 of the model over time, while Figure 7.4.1f graphs a histogram of the R^2 scores.

the same time.

The model presents decent fits for fluctuations in FRC. Alternative models, such as those by Vasicek (1977) and Nelson & Siegel (1987), can almost fully account for fluctuations using more flexible functional forms. In comparison, the more rigid form of Eq. 7.25 still achieves an R^2 value between 0.8 and 1 on most accounts, as seen in Figure 7.4.1f. The model does not appear to perform as well during periods of stress in markets, most notably in the 2019 ‘taper tantrum’ seen in Figure 7.4.1e. This may not be too surprising, as this period presented unusual circumstances for the banking sector. It is not clear how the model would accommodate the impact of quantitative easing or tightening.

7.4.2 Gamma pattern in spreads

Hypothesis 2 proposes that the shape of the FRC can be easily fit using a linear combination of maturities on linear and geometric scales. I use daily observations for ED future prices to estimate the parameters in Eq. 7.26:

$$s_t(m) = \gamma_{g,t} \log m + \beta_{g,t} m + \varepsilon_t(m), \quad (7.29)$$

where $\gamma_{g,t}$ and $\beta_{g,t}$ are WLS estimates, with weights $\omega_{g,t}(m) = p_t(m) / \sum^{M-1} p_t(m)$, and $\varepsilon_t(m)$ is an idiosyncratic error term. Note that $\gamma_{g,t}$ estimates the negative of the corresponding parameter in Eq. 7.26. I fix intercept k to zero, since the curve on spreads will originate from a spot value of zero, $s_t(0) = 0$. I estimate these parameters on daily snapshots of the FRC. Figure 7.4.2 illustrates the results. The solid lines denote each parameters’ point estimates, while the green regions outline the 95% confidence intervals.

Linear slope coefficient $\hat{\beta}_{g,t}$ in Figure 7.4.2a follows the downward trajectory of spreads observed in the sample period. It approaches, but does not cross, zero in 2019, where the FRC stands close to flat. For reference, the value of 1.5×10^{-5} recorded in early 2014 corresponds to an annualised interest rate spread of 0.54%. Sharp changes in $\hat{\beta}_{g,t}$ line up with the volatility of spreads in Figure 7.4.1a caused by movements in the linear slope coefficient. The slope on geometric maturity $\hat{\gamma}_{g,t}$ is positive for most of the sample, which indicates the presence of a hump in the FRC in the medium term. This reverts dramatically in 2019, when rates at the front end inverted.

The fit of the model to the shape of the FRC is close to perfect, with the R^2 in most cases ranging between 0.97 and 1 in Figure 7.4.2d. However, the shape of the FRC in the period following, and including, 2019 is not well explained by the model, and this stands out in Figure 7.4.2c. As mentioned before, the assumptions underlying the model do not allow for the unusual circumstances that governed interest rate derivatives at that time.

7.4.3 Maximal volatility

The evidence for the model’s prediction on maximal FRC deviations in Hypothesis 3 is mixed, due to noisy parameter estimates. Recall that the condition under which Hypothesis 3 holds requires parameters β and γ to take opposing signs. At a quarterly frequency, this yields 16 samples with which to test Eq. 7.27. As a result, I use parameter estimates from

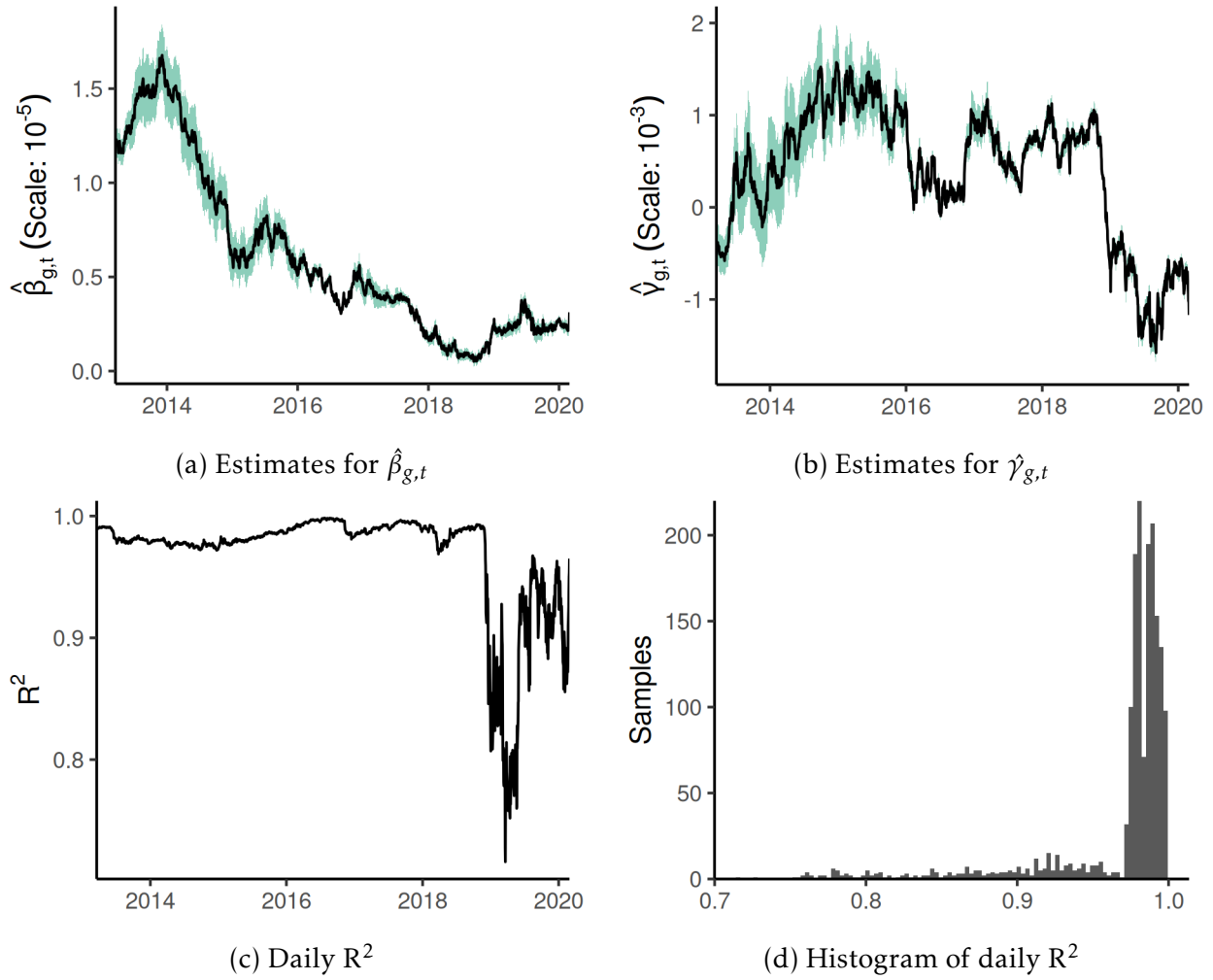


Figure 7.4.2: Daily estimates for model parameters suggest the FRC evolved significantly over time: This figure graphs WLS estimates for parameters $\hat{\beta}_{g,t}$ (Figure 7.4.2a) and $\hat{\gamma}_{g,t}$ (Figure 7.4.2b) in Eq. 7.29, where t denotes a day. The solid black lines denote point estimates of these parameters for one quarter, with the green region outlining the 95% confidence interval. Figure 7.4.2c plots the R^2 of the model over time, while Figure 7.4.2d graphs a histogram of the R^2 scores.

daily samples, despite the fact that they are biased. These estimates are reported separately in Appendix 7.C.

Peak maturity m_t^* is the maturity of the future with the largest growth in spreads $r_t(m)$ on a given day t . I use daily estimates of parameters β_t and γ_t from Eq. 7.28 to run regression

$$\log m_t^* = c^* + a^* \log |\hat{\gamma}_t| + b^* \log |\hat{\beta}_t| + e^*, \quad (7.30)$$

where c^* is constant, a^* and b^* are slope coefficients, and e^* is an error term. I estimate the parameters in Eq. 7.30 using Ordinary Least Squares (OLS) on a sample of observations in which $\hat{\gamma}_t / \hat{\beta}_t < 0$ and $\hat{\gamma}_t > 0$, keeping with the sign restrictions proposed by Corollary 7.1.1. Likewise, I estimate the parameters of

$$\log m_t^- = c^- + a^- \log |\hat{\gamma}_t| + b^- \log |\hat{\beta}_t| + e^-, \quad (7.31)$$

where c^- is constant, a^- and b^- are slope coefficients, and e^- is an error term, on a sample of observations in which $\hat{\gamma}_t/\hat{\beta}_t < 0$ and $\hat{\gamma}_t < 0$. In these instances, I define m^- as the maturity of the future with the smallest growth in spreads $r_t(m)$ on a given day t .

Table 7.4.1: Extreme FRC fluctuations as a function of model parameters

	<i>Dependent variable:</i>	
	$\log m^*$	$\log m^-$
	(1)	(2)
$\log \hat{\gamma}_t $	0.11** (0.05)	0.22*** (0.06)
$\log \hat{\beta}_t $	0.09 (0.06)	0.06 (0.06)
Constant	8.04*** (0.60)	8.31*** (0.67)
Observations	575	571
R ²	0.02	0.03

*p<0.1; **p<0.05; ***p<0.01

Notes: This table presents OLS estimates for the relationship between the maturity of the contract with largest return m^* and the smallest price return (m^-) with parameters $\hat{\gamma}_t$ and $\hat{\beta}_t$ estimated for Eq. 7.28. Column (1) estimates this relationship for the largest return, corresponding to Eq. 7.30, while Column (2) for the smallest return, in Eq. 7.31. These observations are made at a daily frequency. Figure 7.C.1 of Appendix 7.C reports parameter estimates $\hat{\gamma}_t$ and $\hat{\beta}_t$. Standard errors are denoted in brackets.

Table 7.4.1 presents the OLS estimates to Eqs. 7.30-7.31. The statistically significant estimates for a^* and a^- suggest that the geometric duration component in FRC fluctuations helps explain the maturities of futures with extreme returns, whether maximal or minimal. However, the estimated linear duration component does not seem to predict these extremes, judging by the statistically insignificant estimates for b^* and b^- . In addition, Hypothesis 3 asserts that these parameters take values of one, yet the large intercepts indicate ratios $-\hat{\gamma}_t/\hat{\beta}_t$ systematically underestimate the maturity of the extreme return.

7.5 Conclusion

This paper presents convincing evidence that a model based on banking, as opposed to trading, can explain patterns in FRC fluctuations and shapes. The model suggests that the FRC is generated by a process that maximises the RoA for a representative bank, constrained by the expected duration of its assets on both a linear as well as a geometric scale. While the constraint on linear maturity introduces a slope to the FRC, that on geometric maturity introduces curvature, in a similar manner to common statistical models of the FRC.

This model fits data on the volatility of changes in the FRC sufficiently well to pick out patterns that governed interest rate markets between 2013 and 2020, and almost perfectly explains the observed shapes of the FRC. The model fails to explain patterns in the FRC during the ‘taper tantrum’ of 2019, and suggests that the quantitative tightening policies present some form of omitted variable. The precise mechanism by which these flows impact the aggregates studies here remains a subject for future research.

Bibliography

- Campbell, J. Y. & Shiller, R. J. (1991), 'Yield spreads and interest rate movements: A bird's eye view', *The Review of Economic Studies* **58**(3), 495–514.
- Cochrane, J. H. & Piazzesi, M. (2005), 'Bond risk premia', *The American Economic Review* **95**(1), 138–160.
- Dai, Q. & Singleton, K. J. (2000), 'Specification analysis of affine term structure models', *The Journal of Finance* **55**(5), 1943–1978.
- Duffee, G. R. (2002), 'Term premia and interest rate forecasts in affine models', *The Journal of Finance* **57**(1), 405–443.
- Fama, E. F. (2006), 'The behavior of interest rates', *The Review of Financial Studies* **19**(2), 359–379.
- Fama, E. F. & Bliss, R. R. (1987), 'The information in long-maturity forward rates', *The American Economic Review* **77**(4), 680–692.
- Frank, S. A. (2009), 'Common patterns of nature', *Journal of Evolutionary Biology* **22**(8), 1563–1585.
- Frank, S. A. (2018), 'The Price equation program: Simple invariances unify population dynamics, thermodynamics, probability, information and inference', *Entropy* **20**(12).
- Gabaix, X. (2012), 'Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance', *The Quarterly Journal of Economics* **127**(2), 645–700.
- Nelson, C. R. & Siegel, A. F. (1987), 'Parsimonious modeling of yield curves', *The Journal of Business* **60**(4), 473–489.
- Singh, V. P. & Singh, K. (1985), 'Derivation of the Gamma distribution by using the Principle of Maximum Entropy (POME)', *JAWRA Journal of the American Water Resources Association* **21**(6), 941–952.
- Vasicek, O. (1977), 'An equilibrium characterization of the term structure', *Journal of Financial Economics* **5**(2), 177–188.

Appendix

7.A Data appendix

I use the daily settlement files of the Chicago Mercantile Exchange (CME), accessed via their File Transfer Protocol (FTP) site. As of October 2023, the CME stopped publishing daily settlement files on its FTP site. Due to missing observations, I restrict the time window of the sample to daily files downloaded between 20/03/2013 and 28/02/2020.

For ED futures, I select observations with CODE equal to ED. Variable MTH.STRIKE documents the expiry date of the observed future in a given row. The expiration date is defined as the day before the third Wednesday of the month and year quoted in MTH.STRIKE. Variable DATE records the date in which a closing price is published for that future. I compute the maturity m of the future as the difference between these two dates, in days.

Selecting only unexpired futures ($m > 0$), I calculate the price of each future as the variable under column SETT divided by 100. This variable is denoted $P_t(m)$ in the main text, from which I compute forward rate $f_t(m)$ and the ‘compound’ price $p_t(m)$.

7.B Model appendix

7.B.1 Structure of the model’s banking system

This appendix explains the functioning of a representative bank as an entity that sets forward rates to derive fixed income. I describe the balance sheet of a bank in terms of the bonds it buys and shorts, then show how these positions must net on aggregate.

Bond portfolio and balance sheet A bank’s objective is to maximise the income it derives from a portfolio of risk-free bonds. This portfolio of bonds yields a fixed stream of income over M maturities. In order to introduce a scale for interest incomes over maturities, I break down the bond portfolio by its earnings in every period m . The price of the zero-coupon bond that yields \$1 in period m is $B_t(m)$. The bank can buy bonds maturing in $m + 1$, and sell bonds maturing in m , to derive a return that must equal the forward rate at time m :

$$f_t(m) = -\log \frac{B_t(m+1)}{B_t(m)} \quad (7.32)$$

Taking the spot rate as an example, the price of a bond paying \$1 at maturity $m = 0$ is \$1, so that spot rate $f_t(0)$ equals the return on the bond yielding \$1 at maturity $m = 1$ for price

$B_t(1)$. If bond prices and forward rates did not satisfy Eq. 7.32, the bank can finance a long-short position in the bond market either by borrowing at fixed rate $f_t(m)$ and earning a higher rate on bonds, or lending at fixed rate $f_t(m)$ and receiving funds from a cheaper short position on bonds.

At any given time, the total assets, denoted A_t , of the bank equal

$$A_t = \sum_{m=0}^{M-1} A_t(m), \quad (7.33)$$

where $A_t(m)$ denotes the monetary value of a long position in a bond with maturity $m + 1$. Every unit of position $A_t(m)$ therefore earns $1/B_t(m + 1)$ dollars in period $t + m + 1$.

The bank's assets are financed with external deposits Y_t , in addition to short positions on the bond market. The implicit assumption is that these deposits belong to savers that do not have access to the bond market. Instead, they expect the bank to hold their deposit, and receive some interest generated from the projects the bank can invest in. As such, the bank's total liabilities are

$$L_t = Y_t + \sum_{m=1}^{M-1} L_t(m), \quad (7.34)$$

where $L_t(m)$ denotes the monetary value of a short position in a bond with maturity m . As a result, every unit of position $L_t(m)$ costs $1/B_t(m)$ every period $t + m$. Each short position $L_t(m)$ is used to finance a respective long position $A_t(m)$. The initial position in bonds with maturity $m = 1$ must be financed out of deposits Y_t . This means that the Net Asset Value (NAV) is zero,

$$A_t - L_t = 0,$$

in the absence of any bank capital.

Table 7.B.1: Example balance sheet

Assets		
<i>Bonds (long)</i>		
	$m = 1$	$A_t(0)$
	$m = 2$	$A_t(1)$
<i>Total</i>		A_t
Liabilities		
<i>Bonds (short)</i>	$m = 1$	$L_t(1)$
<i>Deposits</i>		Y_t
<i>Total</i>		L_t

Table 7.B.1 presents an example of the bank's balance sheet at time t , with two available maturities. The structure of the bank's balance sheet highlights its ability to trade in the maturities of debt, investing in long term projects, financed through short-term debt. It increases its position in bonds maturing in two periods by shorting bonds that expire in one

period. Furthermore, Table 7.B.1 highlights that the initial value of assets $A_t(0)$ must equal the value of deposits Y_t for the NAV to remain zero.

An important variable in the paper is the return on assets $A_t(m)$, which is just the ratio of bond prices $B_t(m)/B_t(m+1)$ that reflect the long-short position. The significance of assets $A_t(m)$ is unclear when looking at a single bank in isolation. When looking at the money market in general, it is apparent that these are nothing more than the principals of loans that serve to generate a constant level of interest income – *fixed income*, for short.

Market for fixed income For the market as a whole, each short position a bank holds on a bond must show up as a long position on another bank's balance sheet. As a result, the representative bank's entire holding of assets maturing in $t+m+1$ amounts to $A_t(m)$, without a corresponding short position. These assets represent loans to be issued in the future, which earn the guaranteed return $B_t(m)/B_t(m+1)$. Therefore, the interest income generated in every period is

$$X_t(m) = A_t(m-1) \times \frac{B_t(m-1)}{B_t(m)},$$

the principal amount of those loans issued at time $t+m-1$, but repaid in $t+m$, times the return on a long-short position on bonds used to earn that interest.

Aggregate asset values $A_t(m)$ depend on the series of interest income available. In the case of interest rate futures, I make the simplifying assumption that the market in question seeks to price a stream of fixed payments, standardised to one. As a result, $A_t(m-1) = X_t(m) \times B_t(m)/B_t(m-1) = B_t(m)/B_t(m-1)$. This assumption has the desirable property of producing a constant cash flow over time, which can be interpreted as an aggregate annuity the banking sector pays out to households. Effectively, the bank has access to investment opportunities, which it amortises to pay back its depositors. As with typical mortgages, the initial payments primarily reflect the repayment of the principal (deposits), and later payments primarily reflect interest payments.

7.B.2 Lemma 7.1

Maximising $\mathcal{R}_t$ using the Lagrangian method yields expression

$$\mathcal{L} = - \sum_{m=1}^{M-1} p_t(m) \log p_t(m) - \beta \left(\sum_{m=1}^{M-1} p_t(m)m - \mathcal{D}_t \right) - \gamma \left(\sum_{m=1}^{M-1} p_t(m) \log m - \mathcal{G}_t \right) - \mu \left(\sum_{m=1}^{M-1} p_t(m) - \mathcal{P}_t \right),$$

where β , γ and μ are the Lagrangian multipliers for constraints on $\mathcal{D}_t$, $\mathcal{G}_t$ and $\mathcal{P}_t$.

The first order conditions are

$$\begin{aligned}\log p_t(m) &= -1 - \mu - \beta m - \gamma \log m, \\ \mathcal{D}_t &= \sum_{m=1}^{M-1} p_t(m)m, \\ \mathcal{G}_t &= \sum_{m=1}^{M-1} p_t(m) \log m, \\ \mathcal{P}_t &= \sum_{m=1}^{M-1} p_t(m).\end{aligned}$$

7.B.3 Proposition 7.1

The first-order conditions to Eq. 7.11 are

$$\begin{aligned}r_t(m) &\equiv \frac{\Delta p_t(m)}{p_{t-1}(m)} = \mu - \kappa \log p_{t-1}(m) + \beta m + \gamma \log m, \\ \Delta \mathcal{P}_t &= \sum_{m=1}^{M-1} \Delta p_t(m), \\ \partial \mathcal{R}_t &= - \sum_{m=1}^{M-1} \Delta p_t(m) \log p_{t-1}(m), \\ \Delta \mathcal{D}_t &= \sum_{m=1}^{M-1} \Delta p_t(m)m, \\ \Delta \mathcal{G}_t &= \sum_{m=1}^{M-1} \Delta p_t(m) \log m.\end{aligned}\tag{7.35}$$

In order to solve for parameters κ , β , γ and μ , I denote operators for expectations as

$$\mathbb{E}[x(m)] \equiv \sum_{m=1}^{M-1} \omega_{t-1}(m)x(m),$$

variances as

$$\text{Var}[x(m)] \equiv \sum_{m=1}^{M-1} \omega_{t-1}(m)[x(m) - \mathbb{E}[x(m)]]^2,$$

and covariances as

$$\text{Cov}[x(m), y(m)] \equiv \sum_{m=1}^{M-1} \omega_{t-1}(m)[x(m) - \mathbb{E}[x(m)]] [y(m) - \mathbb{E}[y(m)]].$$

The first-order conditions can therefore be written as

$$\begin{aligned}
\frac{\Delta \mathcal{P}_t}{\mathcal{P}_{t-1}} &= \sum_{m=1}^{M-1} \omega_{t-1}(m) r_t(m) &&= E[r_t(m)], \\
\frac{\partial \mathcal{R}_t}{\mathcal{P}_{t-1}} &= - \sum_{m=1}^{M-1} \omega_{t-1}(m) r_t(m) \log p_{t-1}(m) &&= E[-\log p_{t-1}(m) r_t(m)], \\
\frac{\Delta \mathcal{D}_t}{\mathcal{P}_{t-1}} &= \sum_{m=1}^{M-1} \omega_{t-1}(m) r_t(m) m &&= E[r_t(m) m], \\
\frac{\Delta \mathcal{G}_t}{\mathcal{P}_{t-1}} &= \sum_{m=1}^{M-1} \omega_{t-1}(m) r_t(m) \log m &&= E[r_t(m) \log m].
\end{aligned}$$

Substituting Eq. 7.35 for $r_t(m)$ into the expectations yields a system of equations:

$$\begin{aligned}
E[r_t(m)] &= \mu + \kappa E[-\log p_{t-1}(m)] + \beta E[m] + \gamma E[\log m], \\
\text{Cov}[r_t(m), -\log p_{t-1}(m)] &= \kappa \text{Var}[-\log p_{t-1}(m)] + \beta \text{Cov}[-\log p_{t-1}(m), m] \\
&\quad + \gamma \text{Cov}[-\log p_{t-1}(m), \log m], \\
\text{Cov}[r_t(m), m] &= \kappa \text{Cov}[-\log p_{t-1}(m), m] + \beta \text{Var}[m] + \gamma \text{Cov}[m, \log m], \\
\text{Cov}[r_t(m), \log m] &= \kappa \text{Cov}[-\log p_{t-1}(m), \log m] + \beta \text{Cov}[m, \log m] + \gamma \text{Var}[\log m].
\end{aligned}$$

This system is identical to the Weighted Least Squares (WLS) programme, where returns $r_t(m)$ serve as the dependent variables, variables $-\log p_{t-1}(m)$, m and $\log m$ are independent variables, with weights $\omega_{t-1}(m)$. In order to solve for the four unknown parameters, it is much easier to express the system in matrix form:

$$\mathbf{X}'\mathbf{\Omega}\mathbf{R} = \mathbf{X}'\mathbf{\Omega}\mathbf{X} \begin{bmatrix} \mu & \kappa & \beta & \gamma \end{bmatrix}',$$

where an apostrophe denotes a matrix transpose, $\mathbf{X}$ is the $(M-1) \times 4$ matrix of values for $-\log p_{t-1}(m)$, m and $\log m$, and ones in the first column, and m incrementing by one in every row of $\mathbf{X}$,

$$\mathbf{X} = \begin{bmatrix} 1 & -\log p_{t-1}(1) & 1 & 0 \\ 1 & -\log p_{t-1}(2) & 2 & \log(2) \\ 1 & -\log p_{t-1}(3) & 3 & \log(3) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & -\log p_{t-1}(M-1) & M-1 & \log(M-1) \end{bmatrix},$$

matrix $\mathbf{\Omega}$ is a diagonal $(M - 1) \times (M - 1)$ matrix of weights $\omega_{t-1}(m)$,

$$\mathbf{\Omega} = \begin{bmatrix} \omega_{t-1}(1) & 0 & 0 & \dots & 0 \\ 0 & \omega_{t-1}(2) & 0 & \dots & 0 \\ 0 & 0 & \omega_{t-1}(3) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \omega_{t-1}(M-1) \end{bmatrix},$$

and $\mathbf{R}$ is the rank $(M - 1)$ vector of future price returns $r_t(m)$,

$$\mathbf{R}' = [r_t(1) \quad r_t(2) \quad r_t(3) \quad \dots \quad r_t(M-1)]'.$$

As a result, the parameters take the canonical form

$$\begin{bmatrix} \mu & \kappa & \beta & \gamma \end{bmatrix}' = (\mathbf{X}'\mathbf{\Omega}\mathbf{X})^{-1} \mathbf{X}'\mathbf{\Omega}\mathbf{R}.$$

7.B.4 Corollary 7.1.1

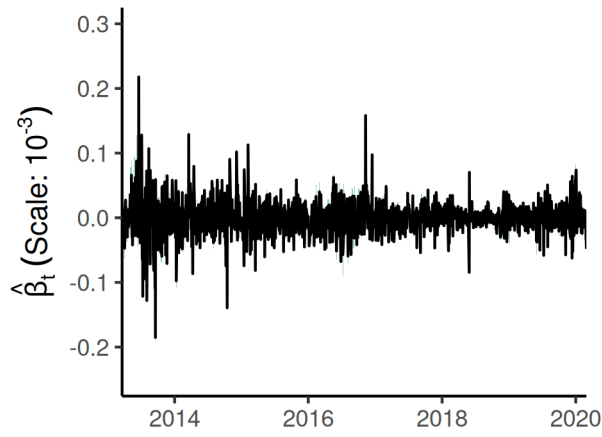
This follows from the first- and second-order conditions of Eq. 7.12:

$$\begin{aligned} \frac{dr_t(m)}{dm} &= \beta + \frac{\gamma}{m}, \\ \frac{d^2r_t(m)}{dm^2} &= -\frac{\gamma}{m^2}. \end{aligned}$$

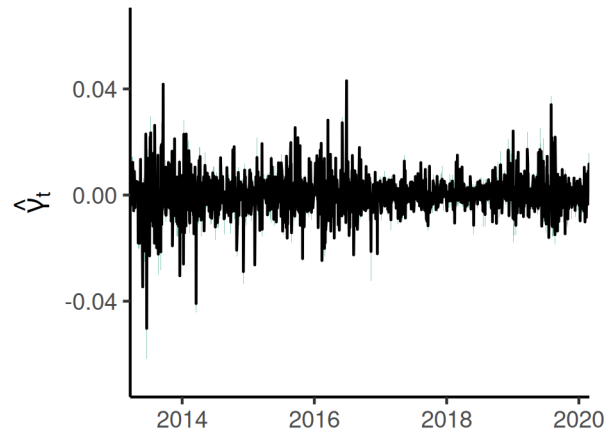
The second-order condition demonstrates that the sign of γ determines whether the turning point is a minimum, or a maximum. The first order condition asserts that parameters γ and β must take opposite signs for the turning point to occur for a maturity above zero, and below $M - 1$.

7.C Daily fluctuations in future prices

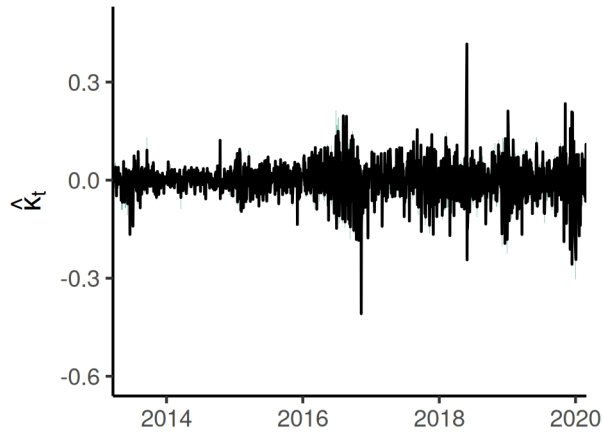
Figure 7.C.1 plots WLS estimates for parameter in Eq. 7.28 using price variables $p_t(m)$ reported on a daily frequency. While the fit of the model, evaluated by the R^2 in Figures 7.C.1e and 7.C.1f, is in line with the version estimated on quarterly samples in Figure 7.4.1, model parameters exhibit significantly more volatility. In addition, estimates for slope coefficients β and γ are likely biased, given that returns are computed with respect to prices of futures with a maturity typically incremented by one day.



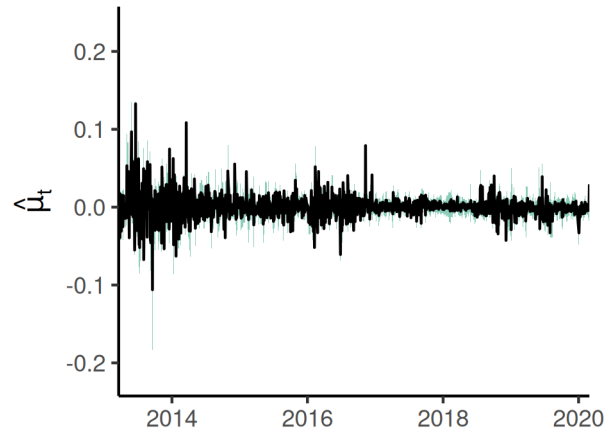
(a) Estimates for $\hat{\beta}_t$



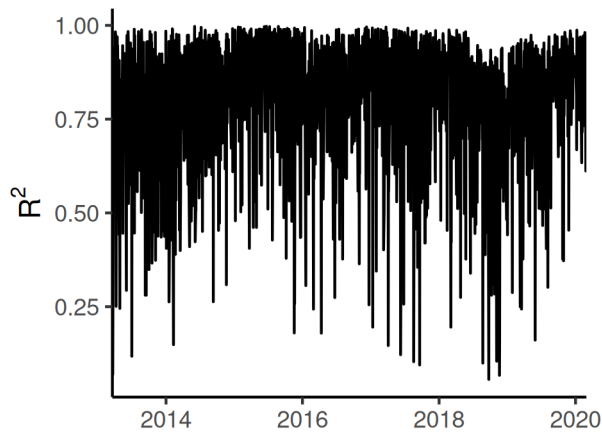
(b) Estimates for $\hat{\gamma}_t$



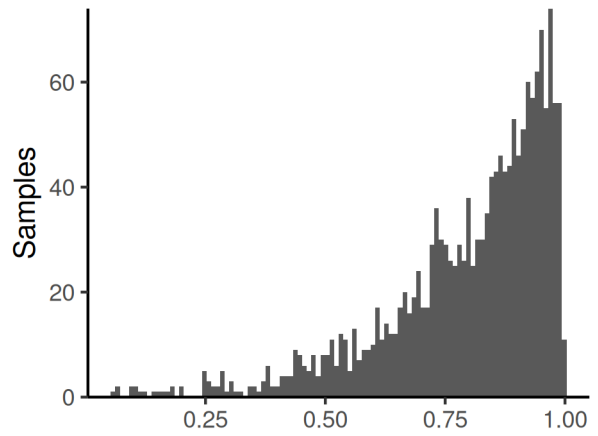
(c) Estimates for $\hat{\kappa}_t$



(d) Estimates for $\hat{\mu}_t$



(e) Daily R^2



(f) Histogram of daily R^2

Figure 7.C.1: Source of fluctuations in the FRC change daily: This figure graphs WLS estimates for parameters $\hat{\beta}_t$ (Figure 7.C.1a), $\hat{\gamma}_t$ (Figure 7.C.1b), $\hat{\kappa}_t$ (Figure 7.C.1c) and $\hat{\mu}_t$ (Figure 7.C.1d) to Eq. 7.28. The solid black lines denote point estimates of these parameters for one day, with the green region outlining the 95% confidence interval. Figure 7.C.1e plots the R^2 of the model over time, while Figure 7.C.1f graphs a histogram of the R^2 scores.

Chapter 8

Changing preferences as a source of stock return variation

Abstract

What can granular data on investors' asset demand tell us about stock return variation? Motivated by the recent literature on demand-based asset pricing, I model the growth rate of portfolio holdings based on evolving asset fundamentals by including demand for asset-specific characteristics in a portfolio optimisation function. Alongside changes in asset characteristics, investors re-allocate wealth according to their evolving demand elasticities. Using the model, I decompose the growth rate of mutual fund holdings by the effect of i) changing stock characteristics, ii) new preferences, and iii) latent demand. I aggregate these components to reconstruct the historical impact of mutual fund investments on stock returns, and find that changing preferences explain at least as much variation in stock prices as changes in fundamentals. This demonstrates the importance of studying heterogeneity in investor preferences, and their evolution, in furthering our understanding of stock market phenomena.

8.1 Introduction

The variety of trading strategies adopted by institutional investors is hard to reconcile with a finance literature focused on market prices. While these prices reflect the aggregated choices, whether informed or not, of countless traders, [Gabaix & Koijen \(2021\)](#) demonstrate that decisions by granular investors cannot be ignored. So do ‘granular’ preferences play a role: firms with large book-to-market equity ratios will be subject to the fortune of value investors.

This paper builds on the recent efforts of [Koijen & Yogo \(2019\)](#) to model portfolio re-allocations triggered by changing asset fundamentals. It includes two innovations that help study investor holdings in a dynamic setting, where heterogeneous preferences change. The key variable of interest is the growth of investors’ stock holdings, instead of static asset demand curves. The first interesting insight from the model is that estimating asset demand curves in a dynamic setting boils down to estimating a factor model for investor portfolios. The growth rate of a holding is proportional to the growth rate of average holdings, changes in asset fundamentals, such as book-to-market equity ratios or market betas, and changes in the investor’s elasticities of demand for different fundamentals. This bears a strong resemblance to asset pricing models that estimate time-varying risk premia of factors in repeated cross-sections of stock returns.

The core model in this paper highlights the trade-off between i) optimising a portfolio according to fundamentals, and ii) the amount of re-balancing, which is constrained by a fixed number of transactions. The first effect is the force that data on firms’ fundamental values exert on the manager’s valuations. The second effect, expressed as square deviations in holding growth rates, is an adjustment cost that introduces inertia to portfolio changes. The sum of those squared deviations serves as an interesting proxy for the aggregate amount of volatility in a portfolio’s components.

This initial framework reproduces the aggregate dynamics of well-known factor models. For example, when the portfolio manager minimises exposure to systemic risk stemming from stock returns’ covariances with the market return, the framework returns the Capital Asset Pricing Model (CAPM) of [Sharpe \(1964\)](#) and [Lintner \(1965\)](#). This paper studies the inertia produced, not just by adjustment costs, but by the portfolio manager’s changing preferences. To that end, I incorporate the asset demand system studied by [Koijen & Yogo \(2019\)](#) into the portfolio management problem. [Koijen & Yogo \(2019\)](#) use a demand system approach to capture the heterogeneity in investor-level investment decisions.

Demand matters since observed fundamental value can be surprising in a setting with noise, and perhaps unreliable. I introduce a new objective function that regulates the portfolio manager’s exposure to *surprises* in fundamental news. The function introduces diminishing returns to fundamental value at the stock level, as in [Koijen & Yogo’s \(2019\)](#) demand system and the investment mandates in [Gabaix & Koijen \(2021\)](#). I call this term *entropy production*, because it gauges the extent to which a portfolio’s exposure to fundamental value remains diversified – and robust to surprises.

This paper highlights that past allocations are informative of future portfolio re-allocations.

This is because changes in fundamentals compel the manager to re-allocate only if it shifts their demand curve. Alternatively, changes in their preferences can trigger re-allocations despite no fundamental news. Therefore, the past allocation is an important variable to condition on to study the impact of fundamentals on holding growth rates. I further decompose past allocations between the share explained by past asset characteristics versus the unexplained, latent demand component. [Koijen & Yogo \(2019\)](#) emphasise the importance of latent demand on stock returns in their empirical analysis. The bottom line is that the paper decomposes holding growth rates at the investor level by three components: i) the effect of changing fundamental values, ii) the effect of changing preferences, and iii) the effect of latent asset demand.

Application The model asserts a structure on stock returns, relying on disaggregated holding growth rates by a cross-section of investors. I decompose the variation in returns explained by investor-level i) changes in value derived from fundamentals, ii) changes in preferences, iii) mean reversion in latent demand. I assess the importance of each component for stock return variation using quarterly data on mutual fund holdings from the Center for Research in Security Prices (CRSP). While the data capture only a fraction of the investment universe, they provide valuable insights into the role of varied investment mandates, classified by CRSP as, for example, growth- versus income-focused mutual funds. I use firm data from Compustat and stock price data from CRSP to construct variables for five fundamental characteristics, following the five factor model of [Fama & French \(2015\)](#).

I construct a proxy for stock price growth by aggregating holding growth rates by mutual funds, weighing their holdings by their reported total net assets. I find that this ‘aggregated’ proxy explains three quarters of variation in stock price growth, even though mutual fund assets cover roughly 18% of stock market equity. Using the repeated cross-sections for portfolio holdings at the mutual fund level, I first estimate the latent demands by adapting the asset demand specification of [Koijen & Yogo \(2019\)](#). In the second step, I estimate the coefficients to the dynamic asset demand model using estimated latent demands, as well as lags and first differences in the stock characteristics. I find large heterogeneity in estimated coefficients, which remains difficult to disentangle when categorising funds by their reported objectives.

The different components to holding growth rates open the doors to predicting stock returns. Changes in fundamental characteristics explain up to 15.1% of stock price growth variation through growth in mutual fund holdings, so predicting changes in firm fundamentals can help predict a similar fraction of stock price variation. Similarly, preferences account for up to 24.3% of stock price variation, which exceeds the share of variation explained by changing fundamental characteristics. Explaining changes in strategies and preferences constitutes an interesting avenue for novel research. Around half of variation in stock returns stems from unexplained variation in holding growth rates, which highlights the fact that funds likely follow a wide variety of signals, many of which may not be observed.

Related literature Different literatures enlighten on the role of entropy production in the cross-section of portfolio holdings. Since a representative investor's asset demand is reflected by the market equity of firms, the model produces a factor not dissimilar from the FFF 'Small-Minus-Big' (SMB) portfolio.¹ In this fashion, this paper integrates recent advances in the study of asset demand into more traditional factor models (Koijen & Yogo 2019, Gabaix & Koijen 2021, Pavlova & Sikorskaya 2022).

Variation in stock returns is typically modelled from the perspective of one representative investor, who evaluates commonly observed risk factors (Sharpe 1964, Lintner 1965, Fama & French 1993, 2015). Certain avenues of research incorporate investor psychology into financial models to capture a variety of reactions to financial news (De Long et al. 1990, Hirshleifer 2015, Bordalo et al. 2020). Their predictions still focus on central tendencies in investment decisions, typically among a few groups of investor types.

Taking a data-focused approach, Koijen & Yogo (2019) demonstrate the theoretical and empirical importance of disaggregated data on investor holdings in reconstructing aggregate stock price dynamics. This paper borrows their approach to model *heterogeneous* demand for characteristics. The authors detail the importance of mean reversion of latent asset demand at the investor level. In a recent paper, Gabaix & Koijen (2021) demonstrate the distortions from mandates held by large mutual funds on the share of capital invested in stocks, versus bonds. Their amplifying mechanism produces significant impacts on overall valuations even when new capital flows are relatively small. Furthermore, Pavlova & Sikorskaya (2022) find that the stock prices increase(decrease) based on their inclusion(deletion) from benchmarks used by fund managers. In addition to the lower expected returns on stocks more prominently featured in those benchmarks, stock prices are evidently affected by preferences of large investors.

In contrast to the asset demand model of Koijen & Yogo (2019), this paper proposes a dynamic asset demand model that predicts the growth rate of portfolio holdings as a function of characteristics-based demand. In particular, investors are compelled to decrease holdings in firms that reveal positive changes in fundamental value, so that their portfolio's exposure to fundamentals remains well diversified. The foundational assumption that justifies this behaviour is that investors exhibit a preference for variety in their portfolio's exposure to fundamentals, so that fundamentals from one asset to another are not perfect substitutes.

A literature in portfolio management studies the tradeoff between a portfolio's depth – its aggregate Sharpe ratio – against its breadth – the number of stocks held (Grinold 1989, Ding & Martin 2017). Chien et al. (2012) study excess stock price volatility due to investors re-balancing their portfolios infrequently. They emphasise the interaction such intermittent re-balancing has with other market actors, who need to anticipate more aggregate risk during downturns. Calvet et al. (2009) study administrative data on portfolios of Swedish households, and document the strong heterogeneity with which individual households re-balance their portfolios. This paper links portfolio re-balancing to the characteristics-based

¹A particularly desirable property of market capitalisation as a second factor is its role as 'zero-covariance' diversification; among a sequence of totally uncorrelated asset returns, optimal allocations tend to be uniform (Samuelson 1967).

approach of [Kojien & Yogo \(2019\)](#), using a utility function that compels the portfolio manager to keep a diversified source of fundamental value through their exposure to different assets.

Another angle enters from the learning literature, in that the portfolio manager benefits from introducing some ‘shrinkage’ with which to estimate valuations from noisy data. In this fashion, the model links to the literature on reinforcement learning ([Camerer & Ho 1999](#), [Malmendier & Nagel 2011](#)), as well as algorithmic trading ([Cover 1991](#)). This behaviour is similar to the use of *robustness* in [Hansen & Sargent \(2008\)](#), who propose to replace expected value in rational agents’ objectives with the entropy of value. In this way, they optimise for a distribution of possible outcomes, which is desirable when the agent takes into account the likelihood their model is mis-specified. Their efforts are closely related to [Sims’s \(2003\)](#) use of *rational inattention*, which is motivated by the limited capacity economic agents possess in processing data. Again, entropy is a natural way with which to describe this channel capacity.

The key message of this paper is that asset demand introduces entropy production into portfolio managers’ optimisation programme. The result is a preference for diversification, which, on aggregate, results in a factor for price returns linked to their overall market value. This size effect is documented first by [Banz \(1981\)](#), and included in the attempts by [Fama & French \(1992, 1993\)](#) to isolate the few factors that maximally describe the cross-section of stock returns. Interestingly, [Pollet & Wilson \(2008\)](#) study an intermediate result, and show the improved performance achieved by funds that diversify more in response to growing larger.

Structure The paper is structured as follows. Section 8.2 introduces the portfolio model that links the evolution of asset demand to the growth rate of holdings. This model produces a decomposition for variation in stock price growth, which I test empirically in Section 8.3. Section 8.4 concludes.

8.2 Modelling a portfolio’s evolution

The motivation behind the model is to study market equities, and their changes, as a function of characteristics-based demand at the portfolio level. The main output is a linear model that expresses the growth rate of portfolio investments as a function of i) the characteristics valued by the portfolio manager – the asset demand function of [Kojien & Yogo \(2019\)](#), ii) the evolution of preferences and iii) latent demand.

I start with a simple version to demonstrate the foundational mechanisms using portfolio management. This version is interesting on its own since it reproduces the CAPM as a special case. I then introduce entropy production to study equilibrium changes in portfolio shares as a function of changing asset demand. This behaviour manifests as a moment condition in the cross-section of stock returns corresponding to the size factor.

Notation The model includes an initial period and final period, $t - 1$ and t , to describe a cross-section of M stocks indexed by j . Since the focus is on choice between risky assets, I do not include a risk-free security, although many of the results for holding growth rates can be in excess of an arbitrary risk-free rate. While the model applies for a cross-section of portfolio managers with possibly heterogeneous beliefs, I set it up from the perspective of one manager only for notational ease.

At any given time, the total portfolio equity in all M firms is $Y_t = \sum_{j=1}^M Y_{j,t}$, where $Y_{j,t}$ denotes the portfolio equity of firm j . The first variable of interest is the equity for firm j as a fraction of total equity $y_{j,t} \equiv Y_{j,t}/Y_t$, which is strictly positive under a short-selling constraint faced by the manager. Portfolio equities evolve over time by a holding growth rate $g_{j,t} \equiv \Delta Y_{j,t}/Y_{j,t-1}$, where Δ denotes the difference operator $\Delta x_t \equiv x_t - x_{t-1}$. These holding growth rates only reflect changes in equity, without any dividend payments.

A bar denotes the valuation-weighted average of a variable throughout the paper, $\bar{x}_t \equiv \sum_{j=1}^M y_{j,t-1} x_{j,t}$, using past valuations to stay consistent when dealing with average growth rates. Thus, the total growth rate $\bar{g}_t \equiv \sum_{j=1}^M y_{j,t-1} g_{j,t} = \Delta Y_t/Y_{t-1}$ is the holding growth rate averaged across j . Since the model uses several permutations of variances and covariance, I standardise their notation using

$$\sigma^2(\mathbf{x}_t) \equiv \sum_{j=1}^M y_{j,t-1} (x_{j,t} - \bar{x}_t)^2$$

for the variance of a vector $\mathbf{x}_t$ of variables $x_{j,t}$ weighted by portfolio equities $y_{j,t-1}$, and

$$\sigma(\mathbf{z}_t, \mathbf{x}_t) \equiv \sum_{j=1}^M y_{j,t-1} (z_{j,t} - \bar{z}_t)(x_{j,t} - \bar{x}_t).$$

for covariances between values $z_{j,t}$ and $x_{j,t}$ weighted by market equities $y_{j,t-1}$. As such, I define holding growth variation as the weighted sum of squared holding growth rates.

Definition 8.1. *Holding growth variation is*

$$\sigma^2(\mathbf{g}_t) = \sum_{j=1}^M y_{j,t-1} (g_{j,t} - \bar{g}_t)^2,$$

the value-weighted, average squared deviation in holding growth rate $g_{j,t}$.

It may help to treat the portfolio manager as a social planner in charge of the market portfolio, in which case holding growth rates coincide with stock price returns, and holding growth rate variation is the cross-sectional variance in stock price returns. Holding growth variation, defined by Definition 8.1, is a measure of realised volatility I use instead of the model-implied volatilities that are commonly used to compute portfolio tracking errors. I later discuss how holding growth variation relates to the adjustment cost the manager faces when re-balancing their portfolio.

Fundamental value Firms reveal news on their fundamental value $d_{j,t}$. For now, I start with a single characteristic used to gauge fundamental value. In practice, I use characteristics from the five factor Fama-French Factor (FFF) model (Fama & French 2015). Same as with holding statistics, the average fundamental value is weighted by market equities, $\bar{d}_t = \sum_{j=1}^M y_{j,t-1} d_{j,t}$, and dispersion in fundamental value in the weighted cross-sectional variance in fundamental value $d_{j,t}$ is $\sigma^2(\mathbf{d}_t)$.

Welfare A portfolio manager derives welfare from maximising the fundamental value of their portfolio

$$\mathcal{W}_t = \mathbb{E}[d_{j,t}],$$

The foundational assumption is that the manager is required to liquidate an unknown portion of her portfolio with some probability. In that event, they derive welfare $\sum_{j=1}^M A_{j,t} d_{j,t}$ where $A_{j,t} \leq Y_{j,t}$ is the required liquidation amount, of which each unit grants $d_{j,t}$.

The portfolio manager faces a decision on re-allocations $\Delta Y_{j,t}$, which change her welfare between $t-1$ and t :

$$\begin{aligned} \Delta \mathcal{W}_t &= \Delta \left(\sum_{j=1}^M Y_{j,t} d_{j,t} \right) \\ &= \sum_{j=1}^M \Delta Y_{j,t} d_{j,t} + \sum_{j=1}^M Y_{j,t-1} \Delta d_{j,t}. \end{aligned} \tag{8.1}$$

Note that the manager derives welfare from inflating holdings for all stocks uniformly by some constant $\Delta Y_{j,t} = \Delta Y_t / M$. I address this issue by fixing the change in portfolio equity to an exogenous value ΔY_t , so that improvements in welfare stem from the re-allocations of capital between assets.

Competing investors This model naturally extends to a setting where multiple investors with heterogeneous beliefs form a stock market. In particular, $Y_{i,j,t}$ would denote the portfolio holding of an investor i in stock j , who re-allocates according to their – potentially private – information $d_{i,j,t}$. Koijen & Yogo (2019) take this view by including a market clearing constraint. In order to focus attention on the key mechanism for this paper, I only focus on the allocation decision of the manager in isolation.

8.2.1 Model with one factor

In this section, I gauge the change in portfolio equities $\Delta Y_{j,t}$ between two periods driven by the fundamental value revealed at the start of period t , $d_{j,t}$. In this way, the fundamental values for a cross-section of public firms trigger portfolio re-allocations.

This model yields what looks like the traditional factor structure of returns, expressed for portfolio-level holding growth rates, where the factor driving the cross-section of returns

is the characteristic the manager seeks to target. If the characteristic is exposure to the market return, the result is the classic CAPM of [Sharpe \(1964\)](#) and [Lintner \(1965\)](#).

Objective: fundamental value The portfolio manager is tasked with re-balancing their portfolio allocations to incorporate information $d_{j,t}$ about their fundamental value. Given that the characteristic is constant, $d_{j,t-1} = d_{j,t}$, substituting $\Delta d_{j,t} = 0$ into Eq. 8.1 isolates the improvement in welfare due to better positioning:

$$\mathcal{P} = \sum_{j=1}^M \Delta Y_{j,t} d_{j,t}. \quad (8.2)$$

Quantity $\mathcal{P}$ is the projection of fundamental values $d_{j,t}$ onto portfolio re-allocations $\Delta Y_{j,t}$. This total is analogous to some ‘price discovery’ on behalf of the manager, who re-allocates wealth to maintain a level of exposure to the fundamental value of each stock. For simplicity, I generally refer to $\mathcal{P}$ as fundamental ‘profit’, because it reflects the improvement in the portfolio’s average fundamental value through better positioning.

Constraint: holding growth variation The manager is assumed to face an adjustment cost when changing portfolio allocations, which I represent as a constraint on the total squared deviations of portfolio holding growth rates. The total adjustment between the set of old and new valuations is thus the sum of squared distances:

$$\mathcal{V}^2 = \sum_{j=1}^M \frac{(\Delta Y_{j,t})^2}{Y_{j,t-1}} = \sum_{j=1}^M Y_{j,t-1} g_{j,t}^2. \quad (8.3)$$

In this context, $\mathcal{V}^2$ is a measure for the non-dimensional distance between the market portfolios at times $t - 1$ and t . This lends an intuitive perspective on holding growth variation as the distance between the market portfolio at time t and the portfolio at time $t - 1$. The potential profit from re-allocations can be high when old allocations $Y_{j,t-1}$ are far from optimal. If the manager is compelled to re-allocate more aggressively to meet a given level of profit $\mathcal{P}$, she will raise holding growth variation as a by-product.

Constraint: liquidity A final constraint fixes the change in total portfolio equity to

$$\Delta Y_t = \sum_j \Delta Y_{j,t}, \quad (8.4)$$

labelled as the portfolio’s *liquidity*. This constraint regulates the degree to which the portfolio manager is able to make new investments without funding them out of other positions. In the data, this variable corresponds to the change in total net assets of mutual fund portfolios.

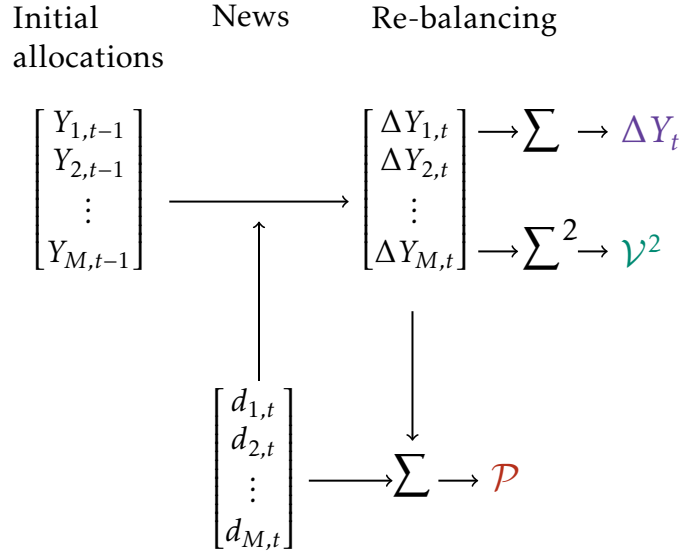


Figure 8.2.1: **Visualising the relationship between profit $\mathcal{P}$ and holding growth rate volatility $\mathcal{V}^2$:** this paper models volatility as a by-product of portfolio re-allocations $\Delta Y_{j,t}$, triggered by the revelation of fundamental values $d_{j,t}$. Returns on holding fundamentals $d_{j,t}$, $\mathcal{P}$, are in equilibrium with liquidity ΔY_t and holding growth rate volatility $\mathcal{V}^2$. Arrows indicate the chain of causality, and symbol $\sum$ indicates summation across vector elements.

Solving the programme Figure 8.2.1 demonstrates the core mechanism of the simple model. Initial valuations $Y_{j,t-1}$ and fundamentals $d_{j,t}$ are given at the start of the period. The portfolio manager has to choose new valuations $Y_{j,t}$, which generate deviations $\Delta Y_{j,t}$. Fixing liquidity ΔY_t and the sum of squared deviations, in the form of holding growth rate volatility $\mathcal{V}^2$, constrains the manager's ability to increase profit $\mathcal{P}$ by better positioning her portfolio according to fundamentals.

The system for portfolio re-allocations $\Delta Y_{j,t}$ is in equilibrium when profit is maximised, subject to the constraint on holding growth variation and liquidity. It amounts to solving the Lagrangian

$$\mathcal{L} = \sum_{j=1}^M \Delta Y_{j,t} d_{j,t} - \frac{1}{2\kappa} \left(\sum_{j=1}^M Y_{j,t-1} \left(\frac{\Delta Y_{j,t}}{Y_{j,t-1}} \right)^2 - \mathcal{V}^2 \right) - \rho \left(\sum_{j=1}^M \Delta Y_{j,t} - \Delta Y_t \right), \quad (8.5)$$

where the objective is profit $\mathcal{P}$, with constraints on volatility $\mathcal{V}^2$ and liquidity ΔY_t . Parameters κ and ρ are the relevant multipliers. I state the solution to the maximisation problem in Proposition 8.1.

Proposition 8.1 (Single factor model). *Profit $\mathcal{P}$ is maximised when portfolio equity grows at rate*

$$g_{j,t} = \rho + \beta(\mathbf{g}_t, \mathbf{d}_t) d_{j,t}, \quad (8.6)$$

where

$$\beta(\mathbf{g}_t, \mathbf{d}_t) = \frac{\sigma(\mathbf{g}_t, \mathbf{d}_t)}{\sigma^2(\mathbf{d}_t)}, \quad \sigma(\mathbf{g}_t, \mathbf{d}_t) = \sigma(\mathbf{g}_t)\sigma(\mathbf{d}_t), \quad \rho = \bar{g}_t - \beta(\mathbf{g}_t, \mathbf{d}_t)\bar{d}_t,$$

$\sigma^2(\mathbf{d}_t)$ is the cross-sectional variance in fundamental news, and $\sigma(\mathbf{g}_t, \mathbf{d}_t)$ is the weighted covariance of fundamentals and returns.

Proof. See Appendix 8.A.1

□

Eq. 8.6 in Proposition 8.1 is just a linear regression of stock holding growth rates on fundamental values at the portfolio level. The subtle difference is that the average overlap between holding growth rates and fundamentals – encapsulated by covariance term $\sigma(\mathbf{g}_t, \mathbf{d}_t)$ – is related to quantity $\mathcal{P}$ the portfolio’s manager actively maximises by growing their allocations at variable rate $g_{j,t}$.

Since holding growth variation $\sigma(\mathbf{g}_t)$ is set by $\mathcal{V}^2$, profit is ultimately constrained by the triangle inequality

$$\sigma(\mathbf{g}_t, \mathbf{d}_t) \leq \sigma(\mathbf{g}_t)\sigma(\mathbf{d}_t).$$

Therefore, the line of reasoning for the manager is that she should adjust her portfolio according to i) overall dispersion in fundamental values and ii) her tolerance to adjustment costs by incurring more variable re-allocations. She can subsequently use those aggregates to determine how her portfolio can maximise exposure to fundamentals, by constructing the slope coefficient $\beta(\mathbf{g}_t, \mathbf{d}_t)$. The intercept ρ reflects changes in total portfolio equity.

This logic follows that of finance tradition, which asserts that stock holding growth rates reflect the exposure to factor innovations that feature in the covariance of holding growth rates (Ross 1976, Fama & French 1993). In this setting, the beta accompanying a given characteristic is tied to each portfolio manager’s possibly varying preferences for portfolio adjustments $\mathcal{V}^2$. This is the one source of heterogeneity on the condition that fundamentals $\mathbf{d}_{j,t}$ are common to all managers.

In the full version of the model, I introduce portfolio equities as a second characteristic for the manager to optimise holding growth rates over, thus yielding a kind of ‘Small-Minus-Big’ factor. In doing so, I integrate the characteristics-based approach of Kojien & Yogo (2019) into what looks like a standard factor model. The discussion in Section 8.2.4 demonstrates how exposure to market risk as the measure for fundamental value yields the CAPM as a special case.

8.2.2 Demand as a robustness criterion

In Proposition 8.1, allocations adjust according to data on fundamental values. I now extend the framework to incorporate asset demand. I then discuss how asset demand acts as the portfolio manager’s memory of past preferences, embedded in initial portfolio allocations, so that the manager is skeptical of adjusting to new data. This use of demand is inspired by Kojien & Yogo (2019), who model the cross-section of portfolio equities using a demand system.

Demand and market equity Assume that the portfolio manager can liquidate portfolio holdings, earning a payoff set according to an exponential utility function with fundamental

value as variable input:

$$\mathcal{U}(d_{j,t}) = 1 - e^{-\lambda d_{j,t}},$$

where λ is the relevant elasticity. Setting the equity share of stock j equal to the marginal utility derived from exposure to fundamental value $d_{j,t}$ yields

$$y_{j,t} = k e^{-\lambda d_{j,t}}, \quad (8.7)$$

where k is a constant, and λ the demand elasticity with respect to fundamental d_j . [Kojen & Yogo \(2019\)](#) estimate this system using time-varying parameters for each fund separately, in order to investigate the heterogeneity and temporal variation in the characteristics sought out by different types of investors – households, pension funds, mutual funds, banks etc. I will refer to equity shares $y_{j,t}$ as the portfolio manager's demand for asset j . The growth in demand is

$$\log \frac{y_{j,t}}{y_{j,t-1}},$$

and reflects diverging marginal utilities in a cross-section of stocks, forced by fundamentals $d_{j,t}$. The portfolio manager is now tasked with re-allocating the market portfolio according to changes in asset demand triggered by fundamental news. Welfare improves according to

$$\Delta \mathcal{W} = \sum_{j=1}^M \Delta Y_{j,t} \log \frac{y_{j,t}}{y_{j,t-1}}. \quad (8.8)$$



Figure 8.2.2: **Re-pricing caused by (a) fundamentals vs. (b) demand for fundamentals.**

The significance of replacing fundamental values in Eq. 8.1 with demands is to add a layer to the portfolio manager's decision to re-allocate, which I demonstrate visually in Figure 8.2.2. Depending on the specific form of demand, allocations may be invariant with respect to certain forms of the payoff function. The payoff function in Eq. 8.7 is one of the simpler options, in which asset demands are invariant to a shift parameter, as well as a multiplier. I discuss this in more detail after outlining the implication this has for the portfolio manager's programme.

Maximising welfare with entropy production To see the role of changing demand in the model, substitute Eq. 8.7 into Eq. 8.8:

$$\begin{aligned}\Delta\mathcal{W} &= \sum_{j=1}^M \Delta Y_{j,t} (\log k - \lambda d_{j,t}) - \sum_{j=1}^M \Delta Y_{j,t} \log y_{j,t-1} \\ &= (\Delta Y_t) \log k - \lambda \mathcal{P} + \partial \mathcal{E}, \\ \partial \mathcal{E} &\equiv - \sum_{j=1}^M \Delta Y_{j,t} \log y_{j,t-1},\end{aligned}\tag{8.9}$$

where I define $\partial \mathcal{E}$ as ‘entropy production’. Re-allocations based on changes in demands now incorporate entropy production in addition to profit from fundamentals $\mathcal{P}$ and liquidity ΔY_t . I contrast this mechanism to the simple one studied in the preceding section in Figure 8.2.2. Welfare improvements are not just a function of better positioning to fundamentals, but also the extent to which the portfolio manager has to change her mind regarding how she derives utility from exposure to fundamentals. Profit $\mathcal{P}$ remains in the objective, and encapsulates the direct impact of fundamentals in the same fashion as in the simple version of the model.

Why entropy? In the same way that entropy governs model misspecification in Hansen & Sargent (2008), entropy production in the present model is a natural measure for the adjustment in the portfolio manager’s frame of reference by re-valuing the parameters to their utility function. These parameter adjustments can change the manner in which the manager allocates new wealth to stocks.

Whether this additional step in deciding allocations increases, or decreases the amount of aggregate holding growth rate volatility will be seen to depend on aggregate outcomes as well as the specific scaling of fundamental values $d_{j,t}$. This scaling is incorporated by the demand function in Eq. 8.7. For instance, if fundamentals improve uniformly for all stocks, so that fundamental $d_{j,t} = a + d'_{j,t} \forall j$ relative to some benchmark value $d'_{j,t}$, there is no need to re-allocate. Demand shares $y_{j,t}$ remain constant thanks to an adjustment in parameter k , that ensures all demands sum to one. This was the case trivially in Proposition 8.1, which depends on fundamental dispersion, since adding a constant to a random variable does not increase its variance.

However, demands in Eq. 8.7 are also invariant with respect to a multiplier a , such that $d_{j,t} = a \times d'_{j,t}$, due to an adjustment in λ . This is not the case in the model without asset demand, where fundamental dispersion would change by a factor a^2 . We therefore have a situation where fundamental dispersion may well have increased, and yet asset demands are the same, so no re-allocations take place.

Persistent asset demand Entropy measures inertia exhibited by asset demand in Eq. 8.7. Figure 8.2.3 demonstrates how persistence in asset demand counteracts the impact of fundamental news on final valuations. Despite the higher fundamental value revealed on behalf of the orange stock, the portfolio manager exhibits a preference for smooth allocations

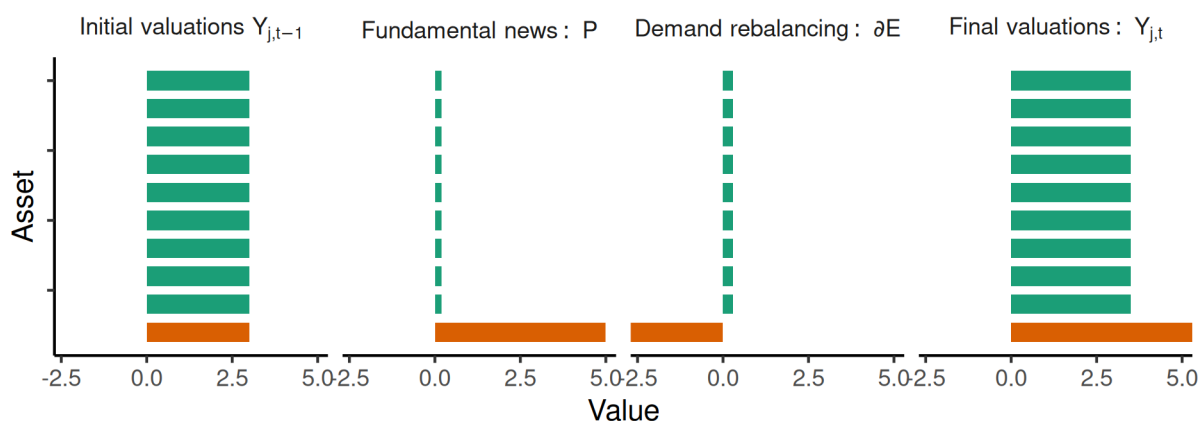


Figure 8.2.3: **Persistent asset demand contributes volatility to changes in asset prices driven by fundamental news:** in this cross-section of 10 hypothetical assets, which are uniformly valued at the start of the period (leftmost plot), the orange asset receives highly positive news on fundamental valuations (center-left plot). Since the portfolio manager exhibits a preference for the demand system at the start of the period, she re-balances the portfolio by removing a fraction of funds from the orange asset, and allocating those to the green assets in the cross-section (center-right plot). The final distribution of valuations is therefore a combination of changes driven by fundamental news, and asset demand.

– a classic motivator for utility in economics. This preference compels them to re-balance the allocations made with respect to fundamentals alone, selling off the orange stock and buying more of the green ones, so that the final portfolio is less concentrated in the orange stock. Measured as entropy production, this inertia contributes volatility to holding growth rates *on top of* dispersion in fundamental news.

There are practical reasons for persistence in asset demand, for example transaction costs, but fundamentally the model requires the manager to *change her preferences* regarding market equity as a function of incoming news on her targeted fundamental characteristic. [Gabaix & Koijen \(2021\)](#) motivate persistent asset demand at the fund level by noting that holdings are not perfectly responsive to news, and funds pursue fixed investment mandates.

However, it is clear that there are strong parallels to the field of robust control, and the exercise in this paper mimics the implementation of robust control methods on behalf of rational agents explored in [Hansen & Sargent \(2008\)](#). In their setting, it is not enough for an agent to incorporate relevant news, but they also make decisions that are least exposed to model misspecification. I provide additional motivation for entropy as an additional SMB factor in Section 8.2.4.

New programme Separating the manager’s welfare objective into three aggregates, namely liquidity, profit and entropy production, outlines which quantities can be used for optimisation. In the programme, I set entropy production as the quantity to maximise, and set profits $\mathcal{P}$ and liquidity ΔY_t as the corresponding constraints. The implication is that the manager seeks to re-allocate based on past asset demand $-\log y_{j,t-1}$, but she cannot afford to deviate from fundamental news. This effectively mimics Bayesian updating, whereby the

manager lends some weight to a prior distribution of demands $y_{j,t-1}$, then adjusts according to new data on fundamentals $d_{j,t}$.

Instead of maximising profit from fundamentals $\mathcal{P}$, the manager now seeks to maximise entropy production $\partial\mathcal{E}$ to satisfy her asset demand. Note that this amounts to introducing a second ‘size’ characteristic to the single-factor model in Proposition 8.1. Profit from fundamentals now acts as a constraint, as the manager has to adjust her demand according to news on fundamentals to the extent that her elasticity of demand λ is high. I solve the problem via the Lagrangian

$$\mathcal{L} = -\sum_{j=1}^M \Delta Y_{j,t} \log y_{j,t-1} - \lambda \left(\sum_{j=1}^M \Delta Y_{j,t} d_{j,t} - \mathcal{P} \right) - \frac{1}{2\kappa} \left(\sum_{j=1}^M Y_{j,t-1} \left(\frac{\Delta Y_{j,t}}{Y_{j,t-1}} \right)^2 - \mathcal{V}^2 \right) - \rho \left(\sum_{j=1}^M \Delta Y_{j,t} - \Delta Y_t \right), \quad (8.10)$$

where the objective is entropy production $\partial\mathcal{E}$, with constraints on profit $\mathcal{P}$, volatility $\mathcal{V}^2$ and liquidity ΔY_t . Parameters λ , κ and ρ are the relevant multipliers. Proposition 8.2 gives the solution to the factor model by including asset demand as a second characteristic.

Proposition 8.2 (Factor model with demand). *Entropy production $\partial\mathcal{E}$ is maximised when*

$$g_{j,t} = \rho + \beta_d d_{j,t} - \beta_y \log y_{j,t-1}, \quad (8.11)$$

where

$$\begin{aligned} \beta_d &= \frac{\sigma(\mathbf{d}_t, \mathbf{g}_t) \sigma^2(-\log \mathbf{y}_{t-1}) - \sigma(\mathbf{d}_t, -\log \mathbf{y}_{t-1}) \sigma(\mathbf{g}_t, -\log \mathbf{y}_{t-1})}{\sigma^2(-\log \mathbf{y}_{t-1}) \sigma^2(\mathbf{d}_t) - \sigma^2(\mathbf{d}_t, -\log \mathbf{y}_{t-1})}, \\ \beta_y &= \frac{\sigma(\mathbf{g}_t, -\log \mathbf{y}_{t-1}) \sigma^2(\mathbf{d}_t) - \sigma(\mathbf{d}_t, -\log \mathbf{y}_{t-1}) \sigma(\mathbf{d}_t, \mathbf{g}_t)}{\sigma^2(-\log \mathbf{y}_{t-1}) \sigma^2(\mathbf{d}_t) - \sigma^2(\mathbf{d}_t, -\log \mathbf{y}_{t-1})}, \\ \rho &= \bar{g}_t - \beta_d \bar{d}_t - \beta_y \mathcal{E}_{t-1}, \\ \mathcal{E}_{t-1} &\equiv -\sum_{j=1}^M y_{j,t-1} \log y_{j,t-1}. \end{aligned}$$

Proof. See Appendix 8.A.2. □

As opposed to Proposition 8.1, Proposition 8.2 includes initial portfolio equities $-\log \mathbf{y}_{t-1}$ as an additional characteristic. The result effectively models holdings growth rate by a weighted linear model, with fundamentals $\mathbf{d}_t$ and log equity shares $-\log \mathbf{y}_{t-1}$ as explanatory variables. The portfolio manager can solve this problem with knowledge of the covariance between initial allocations and fundamental values, which are both known and given exogenously.

Eq. 8.11 is a two-factor model of holding growth rates, where a size-like characteristic regulates the degree to which holding growth rates on large holdings are allowed to increase. The size anomaly was first investigated by Banz (1981). Under the lens of Ross’s (1976) Arbitrage Pricing Theory, this may be justified by the fact that small firms face different risks than large firms. Here, no assumptions were made on risk as a function of holdings size.

In fact, to the degree that riskiness is incorporated in the arbitrary value $d_{j,t}$, the covariance between fundamental value and the size of portfolio shares is also arbitrary. There exists a separate, robustness-like requirement for portfolio allocations to satisfy a smooth payoff function in Eq. 8.7. This is interesting in its own right, but what can be said about investment decisions at the individual level?

8.2.3 Dynamic asset demand

The role of robustness is for the portfolio manager to shield themselves against model misspecification. In the present context, this can be introduced in the form of time-varying demand elasticities $\beta_{d,t}$. Entropy production is a useful tool to track the degree to which changes in demand elasticities compel the portfolio manager to re-allocate.

Proposition 8.2 reveals the interesting structure behind holding growth rates. Namely, consider a hypothetical set of allocations that fail to trigger any re-allocation $\log y_j^* = \log y_{j,t-1} = \log y_{j,t}$. In that circumstance, holding growth rates on all stocks are equal, $g_{j,t} = \bar{g}_t$, and

$$\log y_j^* = k + \frac{\beta_{d,t}}{\beta_y} d_{j,t}, \quad (8.12)$$

where k is a constant, and I allow elasticity $\beta_{d,t}$ to vary with time t . Initial allocations $\log y_{j,t-1}$ can therefore be modelled as a function of variation in demands explained by characteristics, plus some latent demand $e_{j,t-1}$:

$$\log y_{j,t-1} = k + \frac{\beta_{d,t-1}}{\beta_y} d_{j,t-1} + e_{j,t-1}. \quad (8.13)$$

Eq. 8.13 is similar to the asset demand function specified by [Kojien & Yogo \(2019\)](#), except that they retain the non-linear form from Eq. 8.7. The authors subsequently use a large data set on portfolio holdings to estimate the demand elasticities – $\beta_{d,t-1}/\beta_y$ in Eq. 8.13 – with a repeated cross-section of investor holdings. In the current context, substituting Eq. 8.13 into Eq. 8.11, and introducing time subscripts to allow parameters to vary over time, yields an equation for the dynamic evolution of portfolio investments:

$$g_{j,t} = \tilde{\rho}_t + \Lambda_t d_{j,t-1} + \beta_{d,t} \Delta d_{j,t} - \beta_{y,t} e_{j,t-1}, \quad (8.14)$$

where $\tilde{\rho}_t$ is a constant, and coefficient $\Lambda_t = \beta_{y,t} \Delta (\beta_{d,t}/\beta_{y,t})$ reflects the holding growth rate explained by a change in demand elasticities for a given level of fundamental value $d_{j,t-1}$. This equation is of particular interest when studying the growth rate of portfolio investments, and includes three key components.

Endogenous market movements The first is the effect of changing preferences, $\Lambda_t d_{j,t-1}$. It constitutes a level effect on holding growth rates due to a change in the responsiveness of the manager's allocations to the past level of fundamentals. An example of this is a scenario where the value of one firm is unchanged, whereas those of other companies drop.

In this scenario, the demand elasticity is higher, since the resource is overall more scarce. The manager sells off stock in the unchanged firm in order to keep a diversified portfolio. Conversely, if other firms experience an increase in valuations, the holding growth rate on the unchanged investment will also grow to keep up with the other allocations.

The second effect is the direct impact of improved fundamentals, $\beta_{d,t}\Delta d_{j,t}$. This term encapsulates the declining marginal utility of holding a stock that provides more fundamental value. This change effect is distinct from the level effect that depends on changing preferences, in that it regulates over-exposures into highly valuable companies by gradually penalising the growth rate within a portfolio.

The third effect is the impact of [Koijsen & Yogo's \(2019\)](#) latent demand, $\beta_{y,t}e_{j,t-1}$. This term can be understood as a fixed effect revealed at the initial stage, in that it measures a level of demand for the asset that is not justified by fundamental value. Without any observable change in latent demand, it penalises the holding growth rate on that stock within the portfolio. The reason is the same as before: large holdings are difficult to grow in excess of smaller holdings as long as the manager prefers a diversified portfolio. The preference for diversification is embodied in their asset demand function, taken from Eq. 8.7.

Key predictions The first desirable outcome is to gauge the degree to which each component contributes to variation in holding growth rates, and, by extension, to holding growth rates. Besides a variance decomposition, the model makes a prediction on the persistence of latent demand, and by extension a window for return predictability. The dynamic asset demand model suggests the presence of an indirect impact from fundamentals on investment activity, through the re-balancing mechanism outlined in Eq. 8.14. While the direction in which preferences change may go in one direction or the other, the requirement of positive adjustment costs restricts elasticity $\beta_{y,t} > 0$. I state this as a prediction.

Prediction: portfolio holding growth rates $g_{j,t}$ are negatively correlated with past latent demand $e_{j,t-1}$, so that $\beta_{y,t} > 0$.

These predictions are consequential to our understanding of market volatility. In the spirit of [Gabaix & Koijsen \(2021\)](#), this model introduces extra factors into the return cross-section stemming from a host of market participants re-adjusting their portfolios to meet their investment mandates, in addition to the impact of fundamental news alone. I test them in Section 8.3 using data on mutual fund holdings.

8.2.4 Discussion

I introduce the model for a portfolio holding M assets, held by a portfolio manager tasked with maximising her exposure to firms' fundamental value. The result is an equilibrium condition on holding growth variation as a function of the manager's ability to change allocations. Introducing a robustness-like effect by introducing an asset demand system produces extra factors for returns at the portfolio level, as she constantly re-balances allocations even in the absence of news for certain companies. The structural parameters that govern the problem are aggregate risk aversion, with respect to penalising holding growth

variation, and the elasticity of demand with respect to fundamental value, which penalises deviation of allocations from fundamentals.

Comparison to factor models and the CAPM I begin with a simple version of the model in which the portfolio manager changes allocations, thus prices, to match data on fundamental values alone. She acts as a dictator who derives welfare from allocating resources as closely to fundamental values as possible. The only structural parameter is the degree to which the portfolio manager produces price volatility to match fundamental news, which is analogous to an aggregate risk aversion parameter.

This relationship is a factor model. Typically, $\beta(\mathbf{g}_t, \mathbf{d}_t)$ is measured as a time varying return to holding a stock with fundamental value $d_{j,t}$. Introducing additional factors will reduce the contribution of each fundamental to its variation left unexplained, as is the case in the full model. One important factor to consider is variation in market returns.

Proposition 8.1 reproduces the CAPM when the relevant fundamental value is the covariance of return $r_{j,t}$ with market return $\bar{r}_t$, $d_{j,t} = \sigma_t(r_j, \bar{r})^2$.² Setting a representative manager's holding growth rate to the stock's return $g_{j,t} = r_{j,t}$, the average covariance

$$\bar{d}_t = \sum_{j=1}^M y_{j,t-1} \sigma_t(r_j, \bar{r}) = \sigma_t^2(m)$$

is the implicit variance of the market return (Fama & French 2004). Finally, the renowned 'market beta' is given by $\beta(r_j, \bar{r}) \equiv \sigma_t(r_j, \bar{r}) / \sigma_t^2(m) = d_{j,t} / \bar{d}_t$, the proportional risk each unit invested in j contributes to the market portfolio. To see this, note that

$$\sigma(\mathbf{r}_t) = \sigma(\mathbf{d}_t) \times \frac{\bar{r}_t}{\bar{d}_t} \quad (8.15)$$

when substituting $r_{j,t} = (d_{j,t} / \bar{d}_t) \bar{r}_t$ into Definition 8.1 for holding growth variation. Substituting Eq. 8.15 into Eq. 8.6 and re-arranging yields

$$r_{j,t} = \beta(r_j, \bar{r}) \times \bar{r}_t,$$

which is the CAPM of Sharpe (1964) and Lintner (1965) without including the risk free rate and the expectation operator.³

In principle, the full model in Proposition 8.2 is just a two-factor model, and demand $\log y_{j,t-1}$ can be treated as a size characteristic. Again setting holding growth rate to the stock's return $g_{j,t} = r_{j,t}$, the model reconstructs the returns to a portfolio with weights $-\log y_{j,t-1} - \mathcal{E}_{t-1}$ in the term $\sigma(\mathbf{r}_t, -\mathbf{log} \mathbf{y}_{t-1})$. For the representative investor, this is no more than the return on a SMB portfolio included by the FFF model, with the difference that weights are based on the logarithm of size.

²The covariance must vary with time, as indicated by its subscript. Otherwise, the resulting 'market beta' of asset j will mechanically vary over time.

³It makes sense to take a forward-looking aggregate using an expectations operator when assuming that the market beta is constant over time.

Other factors Fama & French (2004, 2015) review likely fundamental characteristics that drive the cross section of asset returns. For example, firms with varying book value of assets relative to their market value are differently affected by interest rates. This is the underlying motivation behind their FFF model (Fama & French 1993, 2015). It would be interesting for future research to incorporate these additional factors in the structural model for volatility using alternative specifications for demand, for example by appealing to the term structure of equity returns (van Binsbergen & Koijen 2017).

8.3 Decomposing price fluctuations

This section tests the key predictions of the model in Section 8.2 using data on mutual fund holdings. The empirical framework can be summarised as a regression specification for the growth rate of investor holdings, in nominal terms, in addition to the demand system estimated by Koijen & Yogo (2019). The purpose is to provide a framework under which stock returns from price changes can be attributed to i) changing fundamentals, ii) changing preferences and iii) latent demand by mutual funds.

8.3.1 Data

Mutual fund data I use the CRSP mutual fund database to extract a panel of portfolio holdings for US mutual funds. These data are required by the Securities and Exchange Commission (SEC) to be reported at a quarterly frequency. In case of duplicates, I only keep the latest values reported in each quarter for a given fund's holdings and total net assets. I filter out any index and exchange traded funds. For each holding in quarter t , I impute its value $Y_{i,j,t}$ by multiplying the portfolio share by the reported total net assets of the fund. I then first-difference it with its value in the previous quarter, if it exists, and take a ratio to compute the holdings value growth rate

$$g_{i,j,t} = \Delta Y_{i,j,t} / Y_{i,j,t-1}.$$

Finally, I drop funds which report 30 or fewer holdings in a given quarter.

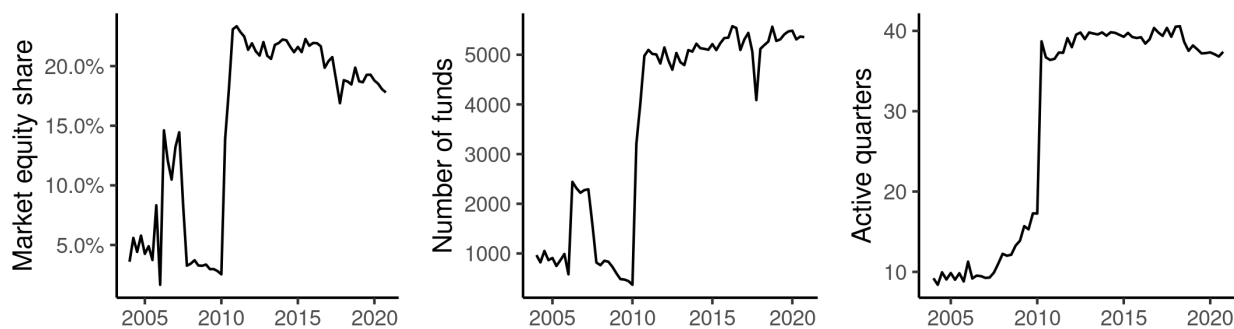


Figure 8.3.1: **Coverage of CRSP mutual fund data levels off in 2011:** this paper uses mutual fund data starting from 2011:Q1 in order to mitigate the sparse selection of funds in the early periods of data coverage.

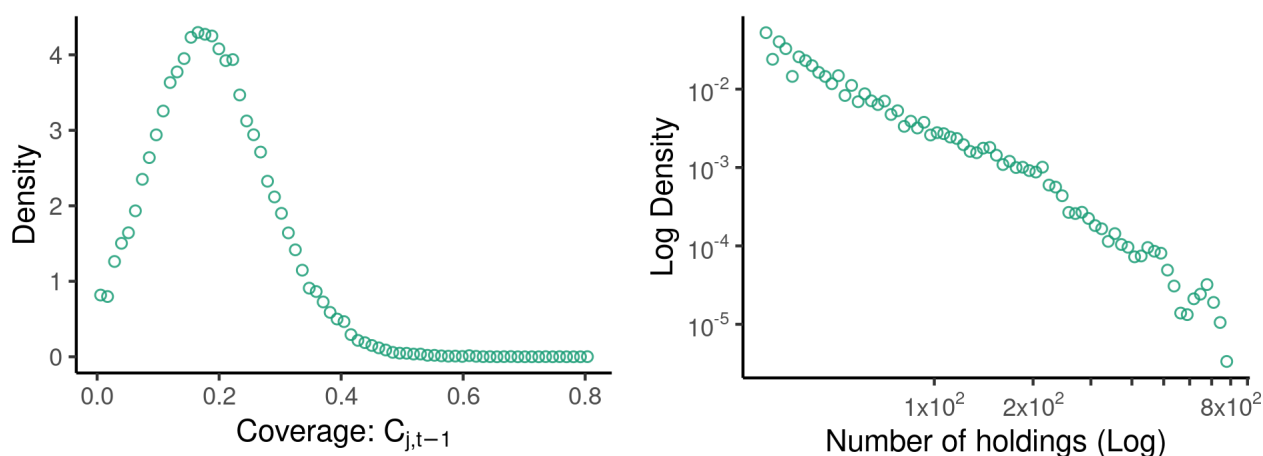


Figure 8.3.2: **(Left) Mutual funds hold around 20% of market equity in modal large companies and (Right) their number of holdings follows an approximate power-law distribution:** lag in coverage (Left), defined as the sum of observed mutual fund holdings as a share of market equity for each stock, is used to construct proxy stock returns according to Eq. 8.17. The number of holdings (Right) is a snapshot of the observation counts available to estimate the holding growth rates, according to Eq. 8.19, at the fund-quarter level of disaggregation.

Figure 8.3.1 illustrates the coverage of the CRSP mutual fund data, in terms of the number of portfolios observed, the total equity held as a fraction of total market equity, and the number of quarters an average fund provides data for. There is a perceptible break in 2011, even though the SEC presumably required disclosures of holdings starting in 2001. Overall, the magnitude of the data is remarkable; even though it only includes a fraction of market participants reported in [Koijen & Yogo \(2019\)](#), these data still account for about 18% of the US stock market. At the stock level, coverage in terms of market equity held across funds mostly varies between 10% and 40% for stocks with a market equity greater than one billion USD, as seen in the Left of Figure 8.3.2. The average fund covered by the data reports holdings for roughly 40 quarters. When removing all observations with incomplete data on portfolio holdings and their corresponding characteristics, the resulting panel includes 1.48×10^7 stock-fund-quarter observations, 8,839 mutual funds, holding 9,483 stocks among them. The number of holdings per fund-quarter, plotted on the Right of Figure 8.3.2, appears to follow a power-law. This indicates that the majority of funds will hold few stocks in their portfolio, in the order of a few dozen, but some hold many hundreds of stocks in any given quarter.

Stock data I use data from the CRSP for common stocks traded on the three major US exchanges: the NYSE, the NYSE MKT (previously AMEX), and the NASDAQ. I build a panel of observations for market equity by firm j at the end of quarter t using prices and outstanding shares reported by CRSP, as detailed in Appendix 8.B.1. Appendix 8.B.1 also details the construction of other variables and additional cleaning procedures. The main variable of interest, besides stock returns and market equity, is the covariance of daily excess returns with the daily excess market return – where the risk-free rate and market returns are provided

by Kenneth R. French's data library.⁴ Importantly, I only include stock-quarter observations where the stock is listed both in the first and last available trading days of that quarter, in order to avoid the distortions caused by firms de-listing or joining an exchange.

Firm data I merge data from Compustat into the panel of stock returns. I use both quarterly as well as annual versions of Compustat, given that certain variables are reported at different frequencies. I detail the cleaning procedure, as well as a list of variables used, in Appendix 8.B.2. For estimation, I use: i) the logarithm of book-to-market equity, using reported liquidation values from Compustat and the corresponding stock's market equity at the end of quarter t , ii) the profit rate, as the fraction of quarterly operating income after tax and interest payments over liquidation value, and iii) the investment rate, as the annual growth rate of the logarithm of gross property, plant and equipment plus intangible capital and inventories. For each stock, I lag and first difference each characteristic by reported quarter – including the market covariance – to produce the required variables to estimate Eq. 8.14.

8.3.2 Empirical framework

Stock prices from the bottom-up The idea that aggregate outcomes can be built up from granular, investor-level actions was introduced in Section 2.1; here I make the link tangible for mutual fund holdings and stock price growth. The price growth on stock j in quarter t can be decomposed via

$$r_{j,t} = \sum_{i=1}^{\mathcal{N}} \frac{Y_{i,j,t-1}}{Y_{j,t-1}} g_{i,j,t}, \quad (8.16)$$

where $Y_{i,j,t-1}/Y_{j,t-1}$ is the past market equity share in stock j held by investor i , and $g_{i,j,t} \equiv \Delta Y_{i,j,t}/Y_{i,j,t-1}$ their holding growth rate in stock j . Eq. 8.16 assumes that the entire universe of $\mathcal{N}$ investors is observed. Notably, stock price growth excludes any dividends distributed to shareholders, an important distinction from the usual stock returns in asset pricing.

In practice, I use data on N mutual funds, who hold a market equity share

$$\hat{Y}_{j,t} = \sum_{i=1}^N Y_{i,j,t} \leq Y_{j,t}, \quad N \leq \mathcal{N}$$

in sum. Here, indices $1 \leq i \leq N$ refer to mutual funds only, whereas the remaining $N < i \leq \mathcal{N}$ indicate unobserved investors. This does not cover the entire investment universe, and coverage

$$C_{j,t} = \frac{\hat{Y}_{j,t}}{Y_{j,t}}$$

may vary by stock and time, a fact I document later in this section. However, I construct a proxy for stock price growth $\hat{r}_{j,t}$, by adapting Eq. 8.16 to include only activity observed in

⁴https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

the mutual funds data:

$$\hat{r}_{j,t} = \sum_{i=1}^N \frac{Y_{i,j,t-1}}{\hat{Y}_{j,t-1}} g_{i,j,t}. \quad (8.17)$$

Roadmap In total, there are two stages to predicting stock price growth $r_{j,t}$. The first stage depends on the relationship between aggregate and proxied returns,

$$r_{j,t} = \delta + \psi \hat{r}_{j,t} + \varepsilon_{j,t}, \quad (8.18)$$

where δ is an intercept, ψ a slope coefficient, and $\varepsilon_{j,t}$ returns unexplained by mutual fund holding growth rates.

Table 8.3.1: Stock price and mutual fund holdings

	<i>Dependent variable:</i>	
	$r_{j,t}$	
	Unweighted	Weighted by coverage
$\hat{\psi}$	0.76 (0.004)	0.87 (0.003)
$\hat{\delta}$	0.01 (0.001)	0.02 (0.001)
Obs.	26,716	26,716
R ²	0.63	0.72
Res. Std. Error	0.11	0.03

Notes: This table provides OLS estimates for the relationship between mutual fund holding growth rates in stock j , weighed by that fund's lagged holding size, and the actual returns of that stock in quarter t . This relationship is seen in Eq. 8.18, where the slope coefficient is estimated to be $\hat{\psi}$, and the intercept $\hat{\delta}$. The second column weighs observations according to the lagged coverage ratio of observed mutual fund holdings as a share of the stock's total market equity, plotted in Figure 8.3.1. Standard errors are reported in brackets.

To motivate what follows, I test the relationship between stock returns and their proxy constructed from the aggregated holdings growth at the mutual fund level, according to Eq. 8.18. Due to the sensitivity of the results for small-cap stocks, I restrict the data to stocks with market equity exceeding one billion USD. Table 8.3.1 presents the OLS estimates to the slope $\hat{\psi}$ and intercept $\hat{\delta}$ in Table 8.3.1. In contrast to the first column, the second column uses a weighted OLS scheme, where the weights are equal to coverage $C_{j,t-1}$. The weighted estimates ought to be more reliable, because the relationship between the return proxy from mutual fund holdings growth should hold more sway over the stock's actual price return if funds hold a larger share of the stock's market equity. The remarkable result is not only that the slope coefficient is close to one, but the R² of the regression is 0.72, despite the fact that coverage is rarely above 50%, as seen in Figure 8.3.2.

The second stage involves estimating holding growth rates using Eq. 8.14. This is the key innovation of this paper. Holding growth rates are predicted to have a level and change rela-

tionship with respect to stock characteristics, but also depend on old latent demands $e_{i,j,t-1}$. In order to adapt Eq. 8.14 to the current empirical setting, I estimate weighted Ordinary Least Squares (OLS) estimates for the coefficients of

$$g_{i,j,t} = \tilde{\rho}_{i,t} + \sum_{a=1}^A \Lambda_{a,i,t} d_{a,j,t-1} + \sum_{a=1}^A \beta_{a,i,t} \Delta d_{a,j,t} - \beta_{y,i,t} \hat{e}_{i,j,t-1} + v_{i,j,t}, \quad (8.19)$$

which requires estimated latent demands $\hat{e}_{i,j,t}$, plus an idiosyncratic error term $v_{i,j,t}$. Lags in characteristics $d_{a,j,t-1}$ and their changes $\Delta d_{a,j,t}$ are assumed to be observed by all funds in period t .

I finalise the roadmap for this section before explaining the procedure used to estimate latent demands. Coefficient estimates in Eq. 8.19 can then be used to construct the aggregate components to proxied stock price growth in Eq. 8.17. In total, the components aggregate into stock price growth using Eq. 8.18, such that

$$r_{j,t} = \tilde{\delta} + \underbrace{\sum_{a=1}^A L_{a,t} d_{a,j,t-1}}_{\text{Preferences}} + \underbrace{\sum_{a=1}^A B_{a,t} \Delta d_{a,j,t}}_{\text{Fundamentals}} - \underbrace{K_t}_{\text{Latent Demand}} + \tilde{\varepsilon}_{j,t}, \quad (8.20)$$

where $L_{a,t}$, $B_{a,t}$ and K_t are the aggregated coefficients from the holdings growth rate estimation:

$$\begin{aligned} L_{a,j,t} &= \psi \sum_{i=1}^N \frac{Y_{i,j,t-1}}{\hat{Y}_{j,t-1}} \hat{\Lambda}_{a,i,t}, \\ B_{a,j,t} &= \psi \sum_{i=1}^N \frac{Y_{i,j,t-1}}{\hat{Y}_{j,t-1}} \hat{\beta}_{a,i,t}, \\ K_{j,t} &= \psi \sum_{i=1}^N \frac{Y_{i,j,t-1}}{\hat{Y}_{j,t-1}} \hat{\beta}_{y,i,t} \hat{e}_{i,j,t-1}, \end{aligned} \quad (8.21)$$

while intercept $\tilde{\delta}$ and residual $\tilde{\varepsilon}_{j,t}$ are modified to account for the intercept and residuals from the holdings growth rate equations. One note is that differences in funds' equity shares among available stocks introduce variation in aggregated coefficients $L_{a,j,t}$, $B_{a,j,t}$ by stock j . The notes under each term's brace in Eq. 8.20 indicate the possible sources of stock price growth variation, which include the effect of i) changing preferences, ii) changing fundamental characteristics, and iii) latent demand. Before delving into the data used to estimate the coefficients for Eqs. 8.19 and 8.20, I detail the estimation of latent demands and holding growth rates as a function of characteristics.

Estimating latent demands I estimate latent demands in a similar fashion to [Kojien & Yogo \(2019\)](#), albeit using the logarithmic version of their preferred specification. While they document that this leads to some inefficiency, and potential bias, I intend to implement their

full procedure in the near future. The static asset demand specification given by Eq. 8.13 includes A possible characteristics:

$$\log \frac{y_{i,j,t}}{y_{i,0,t}} = k_{i,0,t} \log y_{j,t} + \sum_{a=1}^A \gamma_{i,a,t} d_{a,j,t} + e_{i,j,t}, \quad (8.22)$$

where fundamental values $d_{a,j,t}$ are assumed to be public knowledge, and $j = 0$ denotes an arbitrary holding that serves as the benchmark for the estimation procedure. An important detail is that all parameters are allowed to vary by time and fund, as per Koijen & Yogo (2019). The main idea is that each portfolio is treated as a separate cross-section, and the asset demand systems are estimated for each investor and quarter in isolation.

Log market equity share $\log y_{j,t} = \log Y_{j,t}/Y_t$ features as a variable that determines the intercept for the demand system. As a case scenario, elasticities $\gamma_{i,a,t}$ should equal zero, and intercept $k_{i,0,t}$ one, for an index fund with holdings that match market equity shares. The characteristics to assess fundamental value with are the same as in Koijen & Yogo (2019), namely the five stock characteristics from Fama & French (2015), except for size which is implicitly included in market equity shares: log book-to-market equity, investment growth rate, profit rate, and market covariance.

Hypotheses Characteristics valued by fund managers, gauged by elasticities $\beta_{a,i,t} < 1$, may significantly contribute to variation in stock price growth. This speaks to a long-standing issue regarding the excess volatility of stock returns relative to fundamentals. Here, different funds may compete more, or less, for the same assets, depending on the degree of overlap in their preferences. Eq. 8.20 clarifies how changing fundamentals may be amplified if aggregate demand elasticities happen to be high. This line of research is actively pursued (Gabaix & Koijen 2021).

The case for diversification, which is central to this paper, is outlined by Prediction from Section 8.2, which asserts a negative relationship between latent demand and holding growth rate, $\beta_{y,i,t} > 0$. Without an objective that compels the fund to increase the breadth of their portfolio, there is no reason for the level of past holdings to drive growth in corresponding holdings, beyond what is justified by changes in fundamentals. Moreover, mutual funds may find themselves compelled to divest out of positions they are unable to justify. Koijen & Yogo (2019) attribute this phenomenon to a mean-reversion in latent demands.

As discussed in Section 8.2, coefficient $\Lambda_{a,i,t}$ captures changes in fund i 's sensitivity to characteristic $d_{a,j,t-1}$. These preference changes can be understood as a level effect, because they compel the fund to grow their holdings at variable rates despite the characteristic itself remaining constant. This paper does not offer a prediction on the sign of these coefficients, since preferences can evolve in different ways. However, it opens the door to interesting research on the evolution of investment mandates under competition (Brock & Hommes 1997, Farmer 2002). Their estimated contribution to stock price growth via Eq. 8.20 would be a valuable motivation to expand that line of research.

I first describe the available data on mutual fund holdings and stocks, then present esti-

mates for Eqs. 8.18-8.22. Finally, I discuss the results in line with these three hypothetical sources of stock price variation and predictability.

8.3.3 How do portfolios grow?

I first estimate Eq. 8.22 using OLS to retrieve fitted values for latent demands $\hat{e}_{i,j,t}$. Generally, I only observe holdings held in two subsequent periods. This might raise concerns regarding a selection effect, by which a fund opens or closes a new position in a manner not captured by the model. For now, I estimate the coefficients to Eq. 8.19 using OLS, and lags of latent demands $\hat{e}_{i,j,t-1}$.

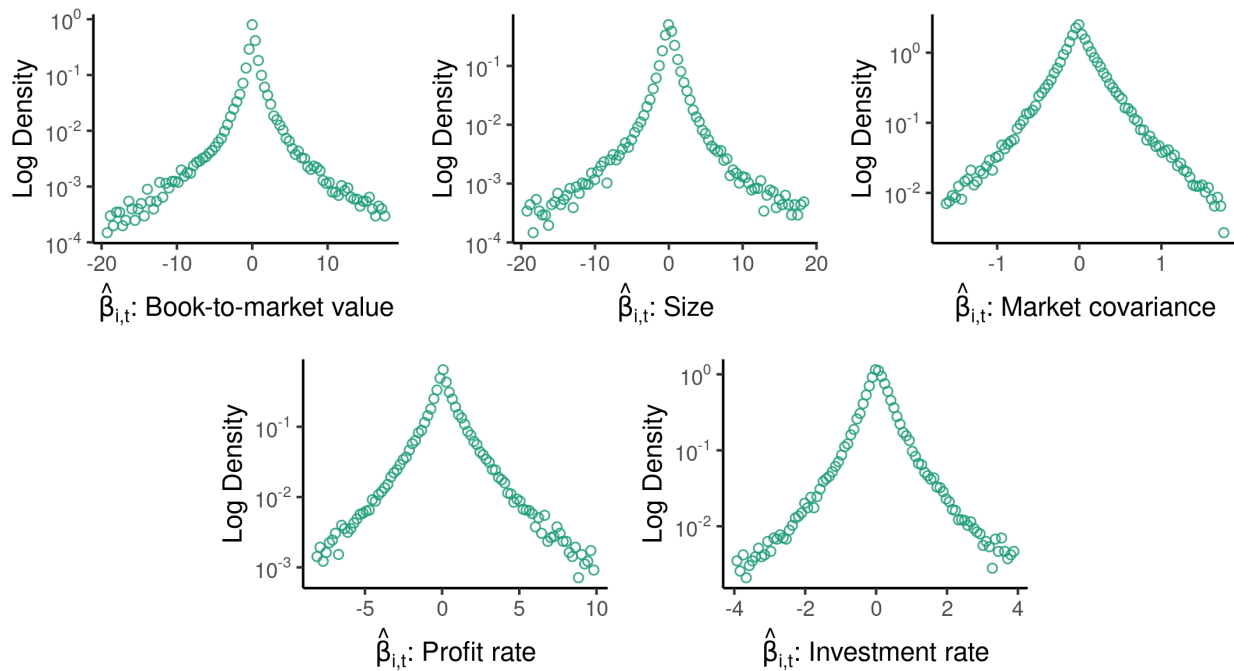


Figure 8.3.3: **Estimated demand coefficients for fundamentals $\hat{\beta}_{a,i,t}$ display strong heterogeneity:** OLS estimates of the relationship between the change in characteristics and the holding growth rate of fund i in quarter t for log book-to-market equity (Top left), the logarithm of market equity share (Top middle), market covariance (Top right), profit rate (Bottom left) and investment rate (Bottom right). 99% of observations shown.

Figure 8.3.3 plots the distributions of the coefficients estimated for changes in characteristics, which correspond to pseudo-demand elasticities $\beta_{a,i,t}$ in Eq. 8.19. These plots communicate considerable heterogeneity, with estimates taking large positive and negative values. It is difficult to discern what a mean estimate might be, which makes sense given that mutual funds hold a variety of mandates and strategies. This heterogeneity also persists in Figure 8.3.4, which plots estimated values for coefficients corresponding to the level effect in the lag of a fundamental characteristic, $\Lambda_{a,i,t}$ in Eq. 8.19. The variable that stands out for a supposed change in preference is the coefficient for book-to-market value, which is typically much smaller than the change effects estimated in Figure 8.3.3. This may indicate that funds target book-to-market values to different degrees, yet this preference does not change much over time. Future research could shed some light on the manner in which mutual fund investments may move over the state space of preference over time.

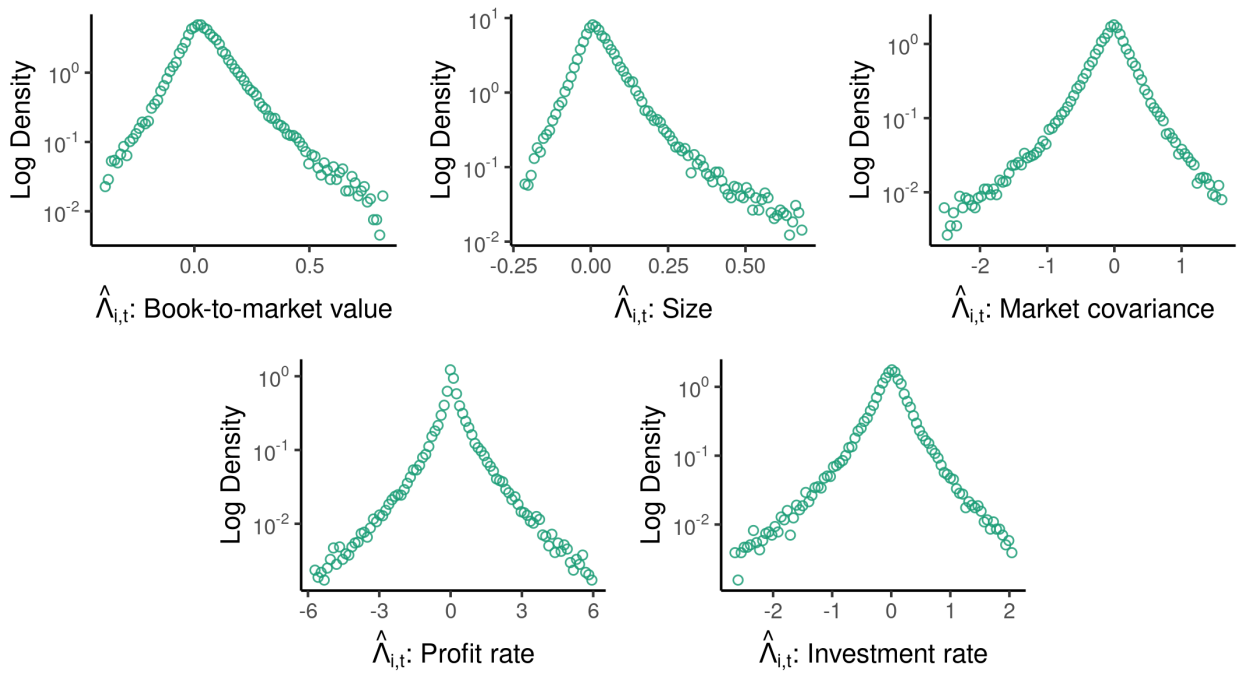


Figure 8.3.4: **Estimated demand coefficients for changes $\hat{\Lambda}_{i,t}$** : OLS estimates of the relationship between the lag in characteristics and the holding growth rate of fund i in quarter t for log book-to-market equity (Top left), the logarithm of market equity share (Top middle), market covariance (Top right), profit rate (Bottom left) and investment rate (Bottom right). 99% of observations shown.

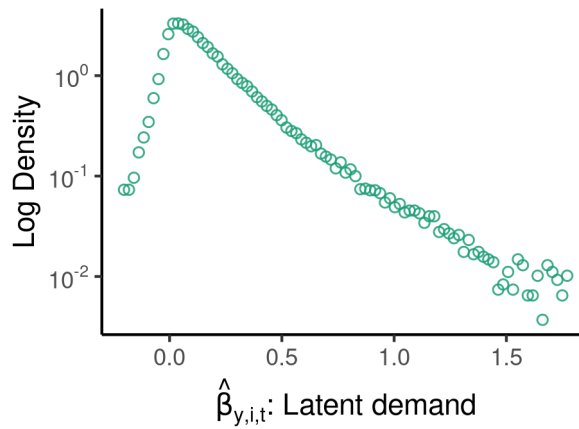


Figure 8.3.5: **Estimated adjustment costs $\hat{\beta}_{y,i,t}$ are overwhelmingly positive**: adjustment cost $\beta_{y,i,t}$ reflects the average effect of a mutual fund's latent demand for an asset in the previous quarter, on its holding growth rate in the current quarter. 99% of observations shown.

Figure 8.3.5 plots the OLS estimates for the effect of lagged latent demand on the growth rate of mutual fund holdings. In the model from Section 8.2, these parameters correspond to the adjustment costs faced by portfolio managers, a constraint which can bind to varying degrees. The data provide strong evidence that this parameter is indeed negative, although, again, it demonstrated substantial heterogeneity among funds' portfolios. Relating it back to the model, a high value for estimate $\hat{\beta}_{y,i,t}$ suggests that a fund aggressively divests out of positions which are not justified by their fundamental values, as measured by the five FFF

characteristics, and the fund's estimated preferences. In contrast, a lower value indicates that the fund tends to sit on the outsized position.

Since the data offer a total of 47,260 portfolios to estimate Eq. 8.19 with, I summarise the results in two ways. First, I plot distributions for all estimated coefficients, pooling them together for all quarter-fund samples. To better highlight some of the interesting heterogeneity in mutual fund investment behaviour, I summarise the key results by reporting the average coefficient for each quarter by *type* of fund, categorised by CRSP's objective codes. Specifically, I group coefficient estimates for funds categorised as i) growth funds (code EDYG), ii) income funds (code EDYI), iii) mixed funds (code EDYB) and iv) an 'other' category of remaining funds.

Denoting each of the four fund types by subscript q , I aggregate coefficient estimates weighed by the relevant fund's reported holding value in the previous quarter, as a share of the total reported in that quarter for the same fund type,

$$\hat{\Lambda}_{q,a,t} = \sum_{i=1}^N \frac{Y_{q,i,t-1}}{\sum_{i=1}^N Y_{q,i,t-1}} \hat{\Lambda}_{q,a,i,t}, \quad \hat{\beta}_{q,a,t} = \sum_{i=1}^N \frac{Y_{q,i,t-1}}{\sum_{i=1}^N Y_{q,i,t-1}} \hat{\beta}_{q,a,i,t}, \quad \hat{\gamma}_{q,a,t} = \sum_{i=1}^N \frac{Y_{q,i,t-1}}{\sum_{i=1}^N Y_{q,i,t-1}} \hat{\gamma}_{q,i,a,t}.$$

Combined, they convey the heterogeneity as well as the central tendency that describes the holding growth rates of mutual funds.

Results Figure 8.3.6 presents the coefficients aggregated for each quarter and fund type, weighed by funds' holding values, using standard box plots. These help discern certain patterns. The sign of the coefficients line up with fund types in some cases. For example, growth funds grow holdings more in response to growth in market shares, compared to income funds, as seen in estimates for $\hat{\beta}_{q,a,t}$ in Figure 8.3.6a. Growth funds also allocate a greater share of their holdings to stocks that move with the market in Figure 8.3.6c's estimates for $\hat{\gamma}_{q,a,t}$. Taken together, those estimates suggest that growth funds stand apart from other funds in the degree to which they follow market and stock momentum. Their holding shares also favour firms that display high investment growth rates, as seen in Figure 8.3.6d. Relative to income funds, growth funds also display a tendency to hold fewer stocks which earn a high rate of profit, as per Figure 8.3.6e. This speaks to their focus on companies seeking future profits at the expense of running short-term losses.

As discussed earlier, it is hard to comment on the lag effect of characteristics on holding growth rates – the component related to preference changes in Eq. 8.20. In Figure 8.3.6, estimates for $\hat{\Lambda}_{q,a,t}$ cluster around zero in all cases. This suggests that mutual fund investment patterns are relatively stable over time with respect to their sensitivity to different stock characteristics. For future research, it would be interesting to pin down exogenous variation in these coefficients to determine what might trigger funds to change their observed responsiveness to these characteristics.

Figure 8.3.6f presents strong evidence that mutual funds penalise the holding growth rates for firms in which they already hold elevated latent demand. The corresponding elasticity, which relates to multiplier $\beta_{y,i,t}$ to adjustment costs, is always positive for quarterly aggregates. These findings lend considerable support to the Prediction made in Section 8.2,

which attributes this relationship to a preference for a diversified portfolio. To accomplish this, the mutual fund must grow their positions in small holdings faster than those of large holdings. [Koijen & Yogo \(2019\)](#) present a similar finding, noting that latent demands exhibit strong mean reversion in their repeated cross-sections.

Summarising the uncertainty around these estimates is tricky, given the sheer amount of coefficients estimates. The range of coefficients in Figure 8.3.6 goes some of the way in describing the possible realisations of these parameters between investors and over time. To get a better idea of the overall fit of the model in Eq. 8.19, Figure 8.3.7 plots the distribution of all adjusted R^2 values estimated for all fund-quarter sub-samples. The model explains around 0.2 to 0.3 of the variation of holding growth rates, which is good considering the rather restricted set of explanatory variables. One concern is that the number of explanatory variables may lead to over-fitted models, especially for certain funds that only report a few dozen holdings.

8.3.4 Aggregating stock returns

I re-construct the proxied returns from Eq. 8.19 using the fitted values for all coefficients. These fitted values result in a nested model for stock price growth using the proxy in Eq. 8.18, which I detail in Eq. 8.20. I re-estimate coefficients to Eq. 8.18, but substitute observations for holding growth rates using the components to Eq. 8.20. Another difference is that I do not filter out stock with market equity less than one billion USD, and instead compute the coefficients using weighted OLS, where weights correspond to the market equity share in the previous quarter, $y_{j,t-1}$.

Table 8.3.2 presents weighted OLS estimates for slope coefficients to the components for holding growth rates nested into the equation for stock price growth. To give a sense for the degree of uncertainty, the first column presents estimated components to Eq. 8.20 using the five FFF model, whereas the second column uses characteristics from the three FFF model, which excludes profit rate and investment rate. Removing those two characteristics significantly increases the number of observations, and while the results will not be directly comparable to those from the five FFF version, that will lend some insights into the degree of variability.

Given that proxied returns already matched stock returns closely in Table 8.3.1, it is unsurprising that the coefficients for both versions of the model are close to one. The additional data from small cap stocks for the 5 FFF model does not seem to worsen the R^2 , although this is largely attributed to the market equity weights. The coefficients are estimated to be close to one when using characteristics to the 3 FFF model, although the overall fit, as judged by an R^2 of 0.57.

Variance decomposition These nested models do not give an accurate picture with regards to stock price variation *explained* by either changes in fundamentals, preferences, or latent demand. This is because variation in fundamental characteristics, even though they map perfectly onto returns, displays only a fraction of the variation in stock prices. To get a better idea for the variation in stock prices explained by mutual fund preferences, changing

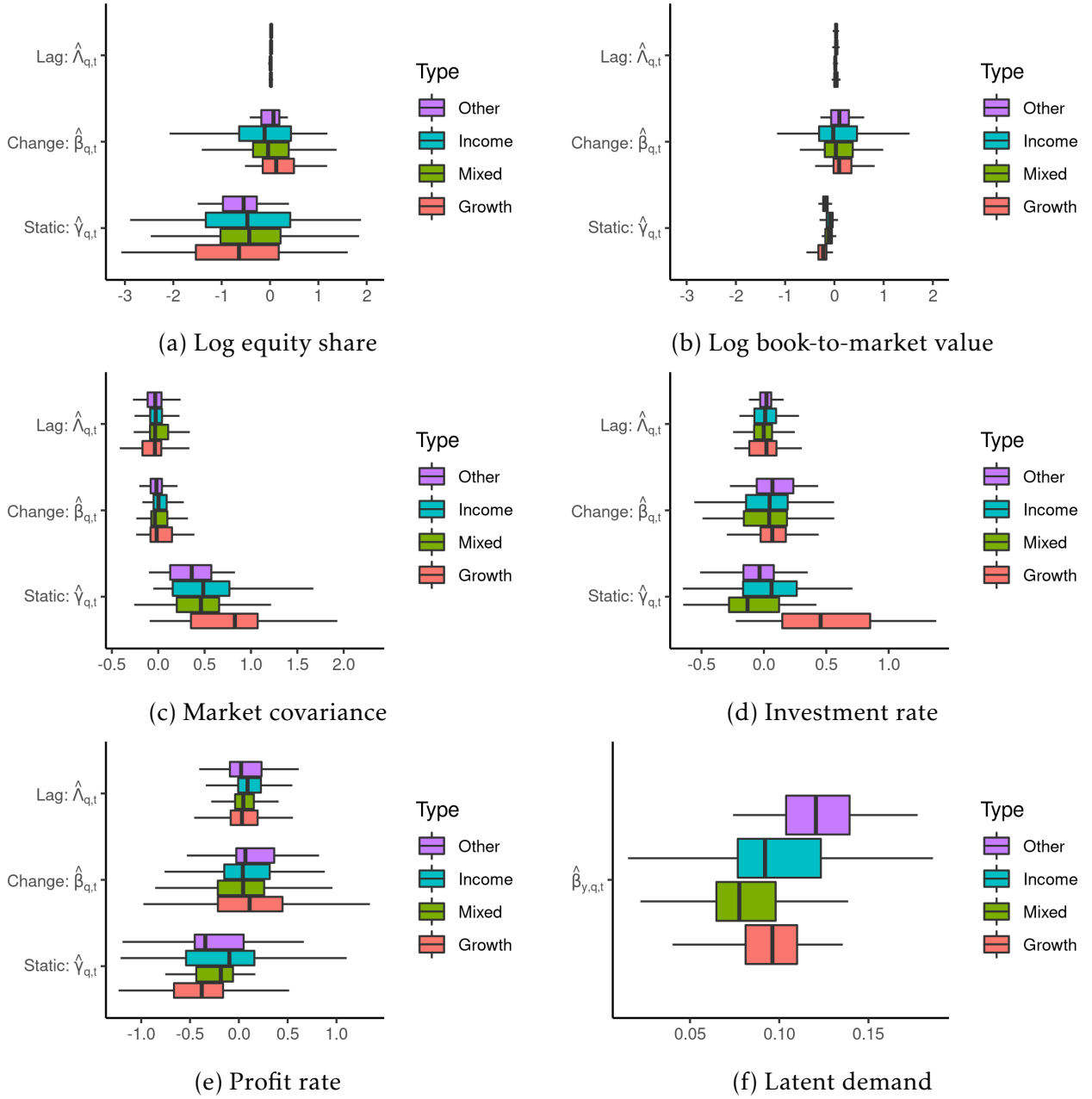


Figure 8.3.6: OLS coefficients for the average response of holdings and their growth rates by characteristic and mutual fund type: these figures present box plots for the aggregated coefficient estimates across mutual funds in each quarter. The averages were computed using the shares of funds' total holding values reported in the previous quarter, and are grouped by certain fund types q that correspond to the CRSP mutual fund objective code (labelled CRSP_OBJ_CD). Each characteristic, labelled by the title, is accompanied by an estimated coefficient from Eq. 8.19 for the effect of its lag ($\hat{\Lambda}_{q,a,t}$) and change ($\hat{\beta}_{q,a,t}$) on holding growth rates, as well as the elasticity of holding shares in response to the characteristic ($\hat{\gamma}_{q,a,t}$) for Eq. 8.22. The bottom right panel includes aggregated coefficients for latent demands ($\hat{\beta}_{y,q,t}$). Observations exceeding the 1st or 99th percentiles are not included for visualisation purposes.

fundamentals or latent demand, I decompose the variance of stock price growth in Eq. 8.20

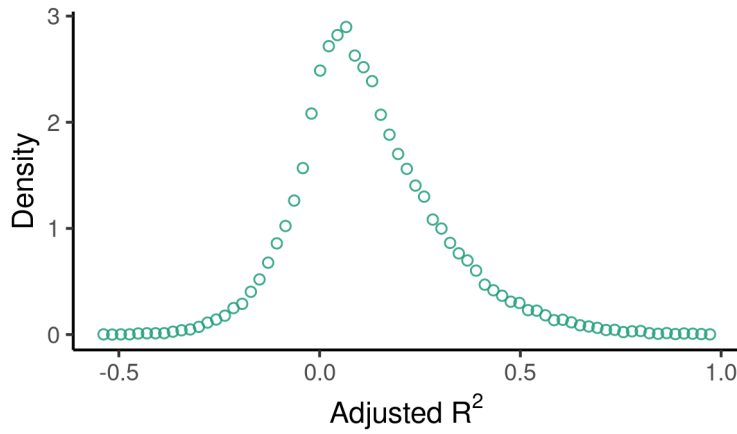


Figure 8.3.7: **Adjusted R^2 values average between 0.2 and 0.3:** This figure plots a histogram for the adjusted R^2 of the holding growth rate model in Eq. 8.19, estimated separately for 47,260 quarterly mutual fund portfolios.

Table 8.3.2: Nesting models for mutual fund holdings to regress on stock prices

	<i>Dependent variable:</i>	
	$r_{j,t}$	
	(1)	(2)
Intercept	0.85 (0.003)	0.95 (0.004)
Preferences	0.87 (0.003)	0.97 (0.004)
Fundamentals	0.86 (0.003)	0.96 (0.004)
Latent demand	0.76 (0.01)	0.86 (0.01)
Residual holdings	0.70 (0.004)	0.83 (0.003)
$\hat{\delta}$	0.02 (0.0004)	0.02 (0.001)
Obs.	41,526	118,351
Adj. R^2	0.75	0.57
Res. Std. Error	0.001	0.001

Notes: This table presents OLS estimates for the relationship between components to growth rates in mutual fund holdings in stock j and the quarterly price growth in stock j . These components, outlined in Eq. 8.20, include the effect of i) an intercept, by which all holdings grew by a uniform rate, ii) changing preferences, whereby funds hold larger or smaller shares in stocks that exhibit higher values for certain characteristics than other stocks, iii) changes in fundamentals, which compel the fund to increase or decrease a holding as a function of changing values for characteristics, iv) latent demand, the holding share unexplained by characteristics in the previous period, and v) a common intercept across all stocks. The characteristics used to construct these components conform to the five factors from the FFF model, namely the market equity share, the book-to-market value, the stock's return covariance with the market return, the profit rate, and the investment rate, in Column (1). Column two constructs these components using only the market equity share, book-to-market value and stock return covariance with the market, which correspond to the three factor FFF model. Standard errors are reported in brackets.

by each component:

$$\sigma^2(\mathbf{r}_t) = \underbrace{\sigma(\boldsymbol{\rho}_t, \mathbf{r}_t)}_{\text{Intercept}} + \underbrace{\sum_{a=1}^A \sigma(B_{a,t} \Delta \mathbf{d}_{a,t}, \mathbf{r}_t)}_{\text{Fundamentals}} + \underbrace{\sum_{a=1}^A \sigma(L_{a,t} \mathbf{d}_{a,t-1}, \mathbf{r}_t)}_{\text{Preferences}} - \underbrace{\sigma(\mathbf{K}_t, \mathbf{r}_t)}_{\text{Latent demand}} + \underbrace{\sigma(\mathbf{v}_t, \mathbf{r}_t)}_{\text{Holding residual}} + \underbrace{\sigma^2(\boldsymbol{\varepsilon}_t)}_{\text{Stock residual}},$$

weighted by the stocks' market equity shares in the previous quarter y_{t-1} . The relevant components include preferences, and fundamentals and latent demand, as per Eq. 8.20, as well as the intercept estimated for mutual funds' holding growth rates, plus the unexplained residuals for holding growth rates and the residuals in stock price growth.

Table 8.3.3: Contributions to stock return variation from growth in mutual fund holdings

	5 Factors	3 Factors
Fundamentals		
Book-to-market	3.26	-2.82
Size	3.02	6.15
Market covariance	3.87	1.20
Profit rate	1.61	
Investment rate	3.32	
<i>Total</i>	15.08	4.53
Preferences		
Book-to-market	3.87	2.31
Size	3.17	0.55
Market covariance	11.80	11.47
Profit rate	2.20	
Investment rate	3.27	
<i>Total</i>	24.30	14.34
Latent Demand	2.40	1.99
Intercept	1.97	6.30
Residual		
Fund holdings	48.98	56.31
Stock returns	9.25	22.84
<i>Total</i>	58.22	79.15

Notes: This table decomposes the cross-sectional variance of quarterly stock price growth by their covariance with different components to mutual fund holding growth rates. These components are outlined in Eq. 8.20.

Table 8.3.3 outlines the variance decomposition of stock price growth by component, separately for the 5 FFF characteristics and the 3 FFF characteristics. The individual characteristics, on their own, contribute rather small amounts to stock price variation. One interesting feature is the large contribution stemming from mutual funds growing holdings that displayed high levels of covariance with the market return. This may speak to mutual funds' changing risk appetites over the course of the cycle. A second observation is that changes in fundamentals contribute roughly the same, if not less, price variation relative to preferences. Recall that fundamentals capture the change effect of characteristics on holdings, whereas preferences capture the level effect; re-balancing according to given fundamentals is therefore reflected in stock price variation, via the tendency of mutual funds to change holding shares in stocks to match a new set of.

Importantly, a significant amount of price variation remains unexplained. However, this is not because stock prices grow more erratically than mutual fund holdings. Rather, about one half of stock price variation is tied to growth rate in mutual fund holdings unexplained

by either fundamentals or preferences for characteristics.

8.3.5 Discussion and limitations

These results do not provide causal evidence that mutual fund investment behaviour has price impact. Importantly, market clearing features in the crudest way possible. Mutual fund investments likely experience some interaction with other market participants. Rather, the exercise is to describe the cross-section of returns from the perspective that a vast heterogeneity in investment mandates provides different sources with which to explain stock returns, besides using fundamental characteristics alone. On that front, the framework is promising, and opens the door to further research that leans into various predictive elements of mutual fund behaviour.

Certain aspects of the empirical exercise can be improved. The selection of stocks into mutual fund portfolios likely biases the estimates for fund-level coefficients, and this extensive margin also provides more information with which to make these estimators more efficient. Another important note is that the driving motivation behind disaggregating returns is the equilibrium dynamics of portfolio shares. The application is currently missing a similar treatment for mutual fund sizes.

Finally, mutual funds may not necessarily be an important driver of daily price fluctuations, since they constitute the market's 'buy side'. Most price fluctuations would be driven by 'sell side' participants, who do not hold inventory overnight, not to mention on the quarterly frequency. The impact of mutual fund investments may therefore be more pronounced in certain liquidity-constrained situations, time scales beyond the quarterly frequency, or through the extensive margin discussed above. Unexplained stock price variation is higher when excluding profit rate and investment rate variables, however this is largely due to a selection effect for small stock that do not report their earnings with the same regularity as larger companies.

8.4 Conclusion

The growth rates of mutual fund holdings are accurately modelled by a model that endogenises portfolio re-balancing in response to fundamentals, producing a level and change effect from stock fundamentals. Although a portfolio is directly more valuable after the fundamental value of one company increases, changes in preferences and an intent to remain diversified trigger re-balancing to avoid being over-exposed to a few firms. Empirically, a larger fraction of variation in stock price growth is explained by the level effect, which reflects evolving preferences aggregated across mutual funds, compared to the change effect, which reflects changes in asset characteristics. While a majority of stock price variation remains unexplained, it mostly stems from variation in holding growth rates that is not explained by fundamentals or preferences, as opposed to stock price variation. Therefore, the main empirical limitation is the search for variables which better explain the evolution of mutual fund holdings.

As demonstrated in the volume of recent research on the topic, investigating asset demand is productive in the study of stock price volatility and their co-movements. In this paper, latent asset demand compels mutual funds to re-balance aggressively in order to meet their preferred allocations according to observable characteristics. In principle, it amounts to no more than the inclusion of an additional characteristic in what is otherwise a straightforward single-factor model. Indeed, the core theoretical framework reproduces the [Sharpe \(1964\)](#) and [Lintner \(1965\)](#) CAPM as a special case, where demand adjusts to meet preferences on the level of exposure to market risk. However, this simple framework does not lend itself well to a scenario where changes in portfolio allocation create inertia, in the form of adjustment costs faced by the manager of the market portfolio. This inertia is closely related to uncertainty, or ‘surprises’ in Information Theory.

This research speaks to the great opportunities in using data on heterogeneous investors to better understand intricate behaviours that are currently missed by the aggregate models dominating finance. In particular, the manner in which investors adapt and change their strategies is shown to contribute a significant amount of stock price fluctuations. This opens the door to a host of novel questions regarding granular origins of large market events.

Bibliography

- Banz, R. W. (1981), ‘The relationship between return and market value of common stocks’, *Journal of Financial Economics* **9**(1), 3–18.
- Bordalo, P., Gennaioli, N. & Shleifer, A. (2020), ‘Memory, Attention, and Choice’, *The Quarterly Journal of Economics* **135**(3), 1399–1442.
- Brock, W. A. & Hommes, C. H. (1997), ‘A rational route to randomness’, *Econometrica* **65**(5), 1059–1095.
- Calvet, L. E., Campbell, J. Y. & Sodini, P. (2009), ‘Fight or flight? Portfolio rebalancing by individual investors’, *The Quarterly Journal of Economics* **124**(1), 301–348.
- Camerer, C. & Ho, T.-H. (1999), ‘Experienced-weighted attraction learning in normal form games’, *Econometrica* **67**(4), 827–874.
- Chien, Y., Cole, H. & Lustig, H. (2012), ‘Is the volatility of the market price of risk due to intermittent portfolio rebalancing?’, *American Economic Review* **102**(6), 2859–96.
- Cover, T. M. (1991), ‘Universal portfolios’, *Mathematical Finance* **1**(1), 1–29.
- De Long, J. B., Shleifer, A., Summers, L. H. & Waldmann, R. J. (1990), ‘Noise trader risk in financial markets’, *Journal of Political Economy* **98**(4), 703–738.
- Ding, Z. & Martin, R. D. (2017), ‘The fundamental law of active management: Redux’, *Journal of Empirical Finance* **43**, 91–114.

- Fama, E. F. & French, K. R. (1992), 'The cross-section of expected stock returns', *The Journal of Finance* **47**(2), 427–465.
- Fama, E. F. & French, K. R. (1993), 'Common risk factors in the returns on stocks and bonds', *Journal of Financial Economics* **33**(1), 3–56.
- Fama, E. F. & French, K. R. (2004), 'The Capital Asset Pricing Model: Theory and evidence', *Journal of Economic Perspectives* **18**(3), 25–46.
- Fama, E. F. & French, K. R. (2015), 'A five-factor asset pricing model', *Journal of Financial Economics* **116**(1), 1 – 22.
- Farmer, J. D. (2002), 'Market force, ecology and evolution', *Industrial and Corporate Change* **11**(5), 895–953.
- Gabaix, X. & Koijen, R. S. (2021), In search of the origins of financial fluctuations: The inelastic markets hypothesis, Technical report, National Bureau of Economic Research.
- Grinold, R. C. (1989), 'The fundamental law of active management', *Journal of Portfolio Management* **15**(3), 30–37.
- Hansen, L. P. & Sargent, T. J. (2008), *Robustness*, Princeton University Press, Princeton, N.J.
- Hirshleifer, D. (2015), 'Behavioral finance', *Annual Review of Financial Economics* **7**, 133–159.
- Koijen, R. S. J. & Yogo, M. (2019), 'A demand system approach to asset pricing', *Journal of Political Economy* **127**(4), 1475–1515.
- Lintner, J. (1965), 'The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets', *The Review of Economics and Statistics* **47**(1), 13–37.
- Malmendier, U. & Nagel, S. (2011), 'Depression babies: Do macroeconomic experiences affect risk taking?', *The Quarterly Journal of Economics* **126**(1), 373–416.
- Pavlova, A. & Sikorskaya, T. (2022), 'Benchmarking Intensity', *The Review of Financial Studies* **36**(3), 859–903.
- Pollet, J. M. & Wilson, M. (2008), 'How does size affect mutual fund behavior?', *The Journal of Finance* **63**(6), 2941–2969.
- Ross, S. A. (1976), 'The arbitrage theory of capital asset pricing', *Journal of Economic Theory* **13**(3), 341–360.
- Samuelson, P. A. (1967), 'General proof that diversification pays', *The Journal of Financial and Quantitative Analysis* **2**(1), 1–13.
- Sharpe, W. F. (1964), 'Capital asset prices: A theory of market equilibrium under conditions of risk', *The Journal of Finance* **19**(3), 425–442.

- Sims, C. A. (2003), 'Implications of rational inattention', *Journal of Monetary Economics* **50**(3), 665–690.
- van Binsbergen, J. H. & Koijen, R. S. (2017), 'The term structure of returns: Facts and theory', *Journal of Financial Economics* **124**(1), 1–21.

Appendix

8.A Proofs

8.A.1 Proof to Proposition 8.1

Solving the first order conditions of Eq. 8.5 with respect to re-allocations $\Delta Y_{j,t}$ yields an expression for the excess holding growth rate for each stock j , as a percentage:

$$\kappa(g_{j,t} - \bar{g}_t) = d_j - \bar{d}_t, \quad (8.23)$$

where a bar denotes a value-weighted mean, $\bar{x}_t \equiv \sum_{j=1}^M y_{j,t-1} x_{j,t}$. Squaring both sides, and multiplying by $y_{j,t-1}$ then summing across j yields:

$$\kappa = \frac{\sigma(\mathbf{d}_t)}{\sigma(\mathbf{g}_t)}. \quad (8.24)$$

Alternatively, multiplying both sides of Eq. 8.23 by $\Delta Y_{j,t}$ yields:

$$\kappa = \frac{\sigma(\mathbf{d}_t, \mathbf{g}_t)}{\sigma^2(\mathbf{g}_t)}, \quad (8.25)$$

which is the coefficient of a weighted regression of fundamentals on holding growth rates. Setting the right-hand sides of Eq. 8.24 and Eq. 8.25 and re-arranging yields the solution.

8.A.2 Proof to Proposition 8.2

For tractability, I re-write Eq. 8.10 as:

$$\mathcal{L} = \sum_{j=1}^M Y_{j,t-1} \left(\frac{\Delta Y_{j,t}}{Y_{j,t-1}} \right)^2 - 2\kappa \left(- \sum_{j=1}^M \Delta Y_{j,t} \log y_{j,t-1} - \partial \mathcal{E} \right) - 2\lambda \left(\sum_{j=1}^M \Delta Y_{j,t} d_{j,t} - \mathcal{P} \right) - 2\rho \left(\sum_{j=1}^M \Delta Y_{j,t} - \Delta Y_t \right), \quad (8.26)$$

which yields the same solution at the maximum.

The first order conditions are

$$\frac{\Delta Y_{j,t}}{Y_{j,t-1}} \equiv g_{j,t} = \rho - \kappa \log y_{j,t-1} + \lambda d_{j,t}, \quad (8.27)$$

and

$$\mathcal{P} = \sum_{j=1}^M \Delta Y_{j,t} d_j, \quad \partial \mathcal{E} = - \sum_{j=1}^M \Delta Y_{j,t} \log y_{j,t-1} \quad \Delta Y_t = \sum_{j=1}^M \Delta Y_{j,t}.$$

Multiplying Eq. 8.27 by $Y_{j,t-1}$, and summing across j solves for

$$\rho = \bar{g}_t + \kappa \sum_{j=1}^M y_{j,t-1} \log y_{j,t-1} - \lambda \bar{d}_t,$$

using the fact that $y_{j,t-1} = Y_{j,t}/Y_{t-1}$ and $\bar{g}_t = \Delta Y_t/Y_{t-1}$. Setting $\mathcal{E}_{t-1} \equiv -\sum_{j=1}^M y_{j,t-1} \log y_{j,t-1}$ yields an expression for the excess holding growth rate for each stock j , as a percentage:

$$g_{j,t} - \bar{g}_t = \kappa (-\log y_{j,t-1} - \mathcal{E}_{t-1}) - \lambda (d_{j,t} - \bar{d}_t). \quad (8.28)$$

The joint result for holding growth rate moments is derived from Eq. 8.28. Multiplying both sides by $Y_{j,t-1} d_j$ and dividing by Y_{t-1} yields

$$\begin{aligned} \frac{\mathcal{P}}{Y_{t-1}} - \bar{g}_t \bar{d} &= \kappa \left(- \sum_{j=1}^M y_{j,t-1} d_j \log y_{j,t-1} - \mathcal{E}_{t-1} \bar{d} \right) + \lambda \sum_{j=1}^M y_{j,t-1} d_{j,t} (d_{j,t} - \bar{d}_t), \\ \Rightarrow \sigma(\mathbf{d}_t, \mathbf{g}_t) &= \kappa \sigma(\mathbf{d}_t, -\mathbf{log} \mathbf{y}_{t-1}) + \lambda \sigma^2(\mathbf{d}_t), \end{aligned}$$

where $\sigma(\mathbf{x}_t, \mathbf{z}_t) = \sum_j y_{j,t-1} (x_{j,t} - \bar{x}_t)(z_{j,t} - \bar{z}_t)$ denotes the value-weighted covariance between x_t and z_t . Multiplying both sides of Eq. 8.28 by $\Delta Y_{j,t}$ and dividing by Y_{t-1} yields

$$\begin{aligned} \frac{\mathcal{V}^2}{Y_{t-1}} - \bar{g}_t^2 &= \kappa \left(- \sum_{j=1}^M y_{j,t-1} g_{j,t} \log y_{j,t-1} - \mathcal{E}_{t-1} \bar{g}_t \right) + \lambda \sum_{j=1}^M y_{j,t-1} g_{j,t} (d_{j,t} - \bar{d}_t), \\ \Rightarrow \sigma^2(\mathbf{g}_t) &= \kappa \sigma(\mathbf{g}_t, -\mathbf{log} \mathbf{y}_{t-1}) + \lambda \sigma(\mathbf{g}_t, \mathbf{d}_t). \end{aligned}$$

Multiplying both sides of Eq. 8.28 by $-\log y_{j,t-1}$ and dividing by Y_{t-1} yields

$$\begin{aligned} \frac{\partial \mathcal{E}}{Y_{t-1}} - \mathcal{E}_{t-1} \bar{g}_t &= \kappa \left(\sum_{j=1}^M y_{j,t-1} (-\log y_{j,t-1} - \mathcal{E}_{t-1})^2 \right) + \lambda \sum_{j=1}^M y_{j,t-1} g_{j,t} (-\log y_{j,t-1}) (d_{j,t} - \bar{d}_t), \\ \Rightarrow \sigma(\mathbf{g}_t, -\mathbf{log} \mathbf{y}_{t-1}) &= \kappa \sigma^2(-\mathbf{log} \mathbf{y}_{t-1}) + \lambda \sigma(\mathbf{d}_t, -\mathbf{log} \mathbf{y}_{t-1}). \end{aligned}$$

Solving by substitution yields

$$\begin{aligned} \lambda &= \frac{\sigma(\mathbf{d}_t, \mathbf{g}_t) \sigma^2(-\mathbf{log} \mathbf{y}_{t-1}) - \sigma(\mathbf{d}_t, -\mathbf{log} \mathbf{y}_{t-1}) \sigma(\mathbf{g}_t, -\mathbf{log} \mathbf{y}_{t-1})}{\sigma^2(-\mathbf{log} \mathbf{y}_{t-1}) \sigma^2(\mathbf{d}_t) - \sigma^2(\mathbf{d}_t, -\mathbf{log} \mathbf{y}_{t-1})}, \\ \kappa &= \frac{\sigma(\mathbf{g}_t, -\mathbf{log} \mathbf{y}_{t-1}) \sigma^2(\mathbf{d}_t) - \sigma(\mathbf{d}_t, -\mathbf{log} \mathbf{y}_{t-1}) \sigma(\mathbf{d}_t, \mathbf{g}_t)}{\sigma^2(-\mathbf{log} \mathbf{y}_{t-1}) \sigma^2(\mathbf{d}_t) - \sigma^2(\mathbf{d}_t, -\mathbf{log} \mathbf{y}_{t-1})}. \end{aligned}$$

Note that $\lambda = \beta_d$ and $\kappa = \beta_y$ in the notation of the Proposition. These are the coefficients to a weighted linear model with two dependent variables, where the variables are lagged equity share $-\log y_{j,t}$ and fundamentals $d_{j,t}$.

8.B Data appendix

8.B.1 CRSP stock variables

This paper uses the following variables from CRSP, with a set of accompanying filters:

- **DATE:** trading date. This corresponds to the calendar date indexed by subscript t in which an observation is made.

Filter: I use observations spanning December 15th, 1973 to December 31st, 2020. This is because coverage for firms listed on the NASDAQ begins December 14th, 1973.

- **PERMNO:** firm identification number. This identifier is issued by CRSP, and for the purposes of this paper constitutes firm identifiers denoted by subscript j .

Filter: When two observations are available on the same trading day for one firm identifier, I keep the first line of observation and discard the rest to mitigate any duplicates.

- **SHRCD:** the share code.

Filter: This paper only includes common stock, which are labelled as 10 and 11.

- **EXCHCD:** exchange code.

Filter: This paper only includes data from the three US exchanges, namely the NYSE (1), the NYSE MKT (2) and the NASDAQ (3).

- **PRC:** last available price. This value is used for the share price $P_{j,t}$ to compute daily market capitalisation of firm j .

Filter: I include observations led by a dash, which indicates an imputed value using the average bid-ask spread.

- **SHROUT:** shares outstanding. This value is used for the number of shares $Q_{j,t}$ to compute the market equity of firm j . This variable can vary due to stock splits, but this split will also be reflected in the unadjusted price $P_{j,t}$ so that the market capitalisation is unaffected.

- **RET:** returns. This variable includes dividends. I remove values that are non-numeric, which refer to different instances of missing observations.

8.B.2 Compustat variables

This paper uses the following variables from Compustat, with a set of accompanying filters:

- **DATADATE:** reporting date. This variable reports the month and year corresponding to company report.

Filter: I use observations reported on December 31st. These correspond to data on company operations from the preceding fiscal year.

- **FYEAR:** fiscal year. I use this variable to assign financial variables to the corresponding calendar year.
- **GVKEY:** firm identification number. This identifier is issued by Compustat.
- **INDFMT:** industry format. This identifier records the reporting standard used for the corresponding financial variables, either as FS or INDL.

Filter: When two observations for financial variables are available for the same year and company, I complete observations if any are missing under industry format INDL using observations reported under industry format FS. I subsequently remove duplicated entries by company-year, prioritising observations under the INDL format.

- **OIBDP:** operating income before depreciation.
- **TXT:** total income taxes.
- **XINT:** interest and related expenses.
- **PPEGT:** property, plant and equipment gross of tax.
- **INTAN:** intangible assets.
- **INVT:** inventories.

Filter: I drop observations for which no value is reported for operating income, tax and interest payments.

- **CEQL:** liquidation value of company assets. I use this variable to construct the book value of firms.

Chapter 9

Lessons and future direction

9.1 Making sense of distributions

The distribution of an economic variable carries important information regarding its behaviour. This holds for the distribution of capital to determine stock returns (Chapter 8), the distribution of assets to determine interest rates (Chapter 7), and for the distribution of output to determine productivity (Chapter 6). In each case, distributions can be understood as a diffusion process constrained by important aggregates.

Economic theory offers an abundance of intuition for diffusion of various quantities or prices. The most conspicuous idea is encapsulated in ‘convex preferences’ for utility models, whereby an agent’s ranking of preferences promotes a variety of outcomes relative to extreme amounts of few outcomes. This thesis serves to connect these ideas to statistics using a shared mathematical framework. In statistics, variety manifests in the production of entropy, where constraints on aggregates reduce the frequency of certain outcomes relative to others. Therein emerges a critical duality between utility theory in economics and entropy in statistics – between Pareto Efficiency and the Maximum Entropy Principle.

This relationship is perhaps most evident in finance, where risk and higher-order moments are of natural concern. Chapter 8 re-casts factor models through the lens of asset demand functions, and Chapter 7 combines exponential and hyperbolic discounting to price interest rate futures. In both cases, entropy manifests as a component to profits earned by a hypothetical representative agent. The statistical perspective promises to take us further.

9.2 Insights into extremes

Social contagion I revisit Chapters 2 and 3 to indicate an agenda for future research. They document the large stock price fluctuations that originate from social contagion in investor sentiment. The binary nature by which sentiments are shared has an impact on the behaviour of the aggregate.

The suggested scaling for individual sentiments $\phi \in (-1, 1)$ as a function of a signal r is

$$\phi = \tanh(r + \alpha \bar{\phi}), \quad (9.1)$$

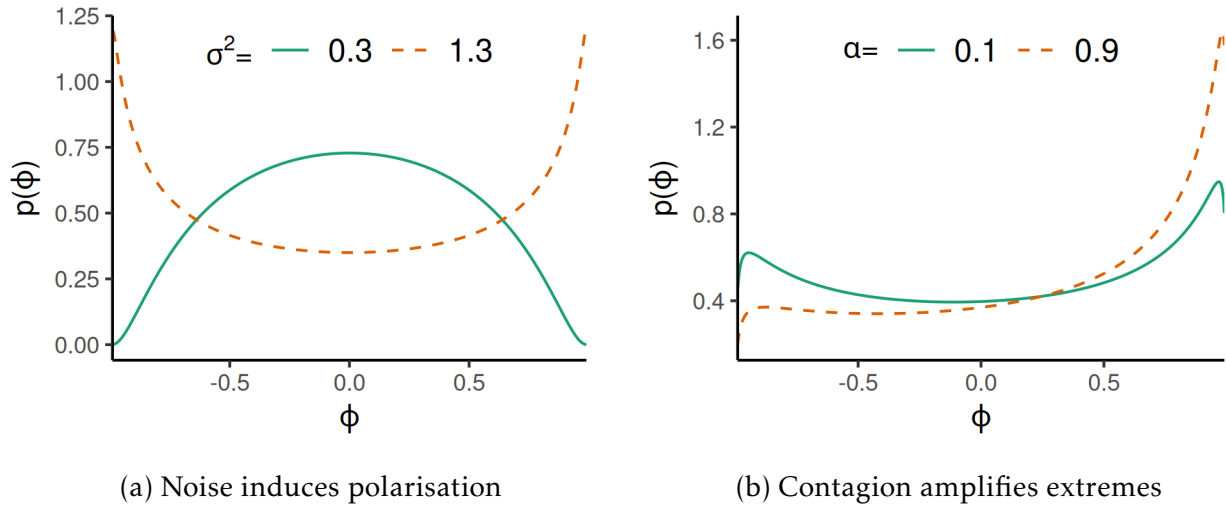


Figure 9.2.1: **Sentiments scale with signal:** These figures plot the distribution of sentiments expressed as a function of a randomly generated signal $r \sim \mathcal{N}(\mu, \sigma^2)$. Figure 9.2.1a demonstrates that high levels of noise σ^2 result in a bi-modal distribution of sentiments, according to Eq 9.2. Figure 9.2.1b illustrates the convergence of sentiments to one extreme with higher levels of contagion α , following Eq. 9.3, while fixing mean signal μ to 0.1.

where $\alpha > 0$ is a parameter, and

$$\bar{\phi} = \int_{-1}^1 \phi p(\phi) d\phi$$

is the population average sentiment. Assuming that individuals independently receive signals from a Gaussian distribution $r \sim \mathcal{N}(\mu, \sigma^2)$, it is straightforward to show that the distribution for sentiments is

$$p(\phi) = \frac{1}{\sigma(1 - \phi^2)\sqrt{2\pi}} \exp \left\{ -\frac{[\text{atanh}(\phi) - \mu - \alpha\bar{\phi}]^2}{2\sigma^2} \right\} \quad (9.2)$$

using a change of variable.

Figure 9.2.1a illustrates this distribution with a mean signal and sentiment equal to zero, but varying degrees of noise σ^2 with which the signal is obtained. When the amount of noise is low, individuals generally agree that the signal is about zero. In contrast, higher noise drives significant disagreement in sentiments, despite the average signal remaining zero. Instead of dampening sentiment, as is the case in traditional portfolio algorithms, noisier data drives significant polarisation.

The consequences of social contagion can be seen when setting the variance of the signal to one ($\sigma = 1$). Solving for mean sentiment $\bar{\phi}$ yields

$$p(\phi) = \frac{1}{(1 - \phi^2)\sqrt{2\pi}} \exp \left\{ -\frac{[\text{atanh}(\phi) + \text{sgn}(\bar{\phi}) \log(1 - |\bar{\phi}|)]^2}{2} \right\}, \quad (9.3)$$

where $\text{sgn}(x)$ takes value -1 if $x < 0$, and 1 if $x > 0$, and $\bar{\phi}$ is the solution to

$$\bar{\phi} = \begin{cases} e^{\mu + \alpha \bar{\phi}} + 1 & \text{if } \mu < 0, \\ e^{-\mu - \alpha \bar{\phi}} - 1 & \text{if } \mu > 0, \\ 0 & \text{if } \mu = 0. \end{cases}$$

Figure 9.2.1b illustrates how increasing contagion parameter α can lead to a strong amplification of average sentiments, even under a modest mean signal $\mu = 0.1$. This average is pushed up by those with negative signals, persuaded the true signal is positive, and by the growing conviction of those with positive signals that they are correct.

This exercise demonstrates the importance of the manner in which economic variables, such as sentiment, scale. This is especially true in the study of extremes, which often stem from some amplification mechanism such as social contagion. These extremes are entirely concealed if the researcher only accesses mean estimates of economic aggregates. Doing so leaves them blind to qualitatively different outcomes, even when those aggregates are identical.

Labour productivity Chapter 4 documents that labour productivity at the firm level closely follows a Lévy alpha-stable distribution. This distribution is a result of the Generalized Central Limit Theorem, but it can also be derived to maximise the entropy under a special constraint. This constraint is placed on the mean for the characteristic function of a random variable (Frank 2009). An interpretation of the Fourier transform to an economic quantity constitutes an open problem, but promises new solutions to old problems.

9.3 Final notes

Equilibrium in the models of Chapters 7 and 8 constitutes a special case that emerges when all constraints are fully satisfied. Outside of equilibrium, the constraints exert a force that generates variance in price returns, and thus volatility. This speaks to the intuition of core finance models that differences in returns must compensate for a variety of risks in their version of equilibrium. There is a sense in which the statistical view proposed by this thesis de-emphasises the importance of an efficient market to achieve that equilibrium. Rather, these dynamics can emerge from changes in important constraints that compel investors to trade one asset for another. Furthermore, these results do not hold in expectation, and instead apply to realised returns and volatility.

In all, this thesis constitutes a scratch on the surface to understand the many distributional features of economic variables. The full limits of the tools developed here for the study of volatility and extremes in economics remain to be seen.

Bibliography

Frank, S. (2009), 'Common patterns of nature', *Journal of Evolutionary Biology* 22(8), 1563–1585.