Rational Inattention in Macroeconomics



Alistair Macaulay

Nuffield College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy in Economics

Trinity 2020

Abstract

In many markets, acquiring and processing the information needed to make optimal decisions costs agents time, effort, and/or money. This thesis is concerned with the macroeconomic implications of the way households behave in the face of these costs, which is known as rational inattention. The thesis consists of three self-contained chapters.

The first and second chapters both explore limited household attention to the choice between different savings products. Regulators (e.g. FCA, 2015) have found that households who gather more information when choosing savings products obtain higher interest rates, and the chapters explore two implications of this.

The first chapter studies how the attention decision varies over the business cycle, and the impact that then has on the cycle. I find in both theory and data that attention to savings product choice is countercyclical. This amplifies business cycle fluctuations, because it introduces a countercyclical wedge between policy rates and the interest rate faced by households. In an estimated DSGE model this amplification is substantial, increasing the variance of consumption by 17%. Policies that reduce the cost of information weaken this amplification and so reduce business cycle volatility.

The second chapter explores the implications of rational inattention to savings choices for inequality and fiscal policy. In a model related to that in chapter 1, I find that poorer households pay less attention to saving choices because they have less to gain from higher interest rates. They therefore face lower interest rates, so have less asset income and less desire to save. I show that this feedback between wealth and attention renders the representative agent steady state of the model unstable with respect to idiosyncratic household income shocks. The stable steady state has a two-agent structure, with poor households at the borrowing constraint paying no attention, while wealthy households pay attention and earn high returns. As the steady state inequality is driven by endogenous attention choices, it responds to policy. I find that aggregate consumption has a smaller and more persistent response to large fiscal expansions in this model than a comparable model where inequality is driven by exogenous discount factor heterogeneity.

The third chapter, co-authored with Martin Ellison, moves away from attention to savings choices, considering the information households choose to acquire about the future of the labour market. We find that introducing rational inattention into a model with uninsurable unemployment risk can generate multiple steady states, where with full information the model has a unique steady state. Large mutations in the distribution of household beliefs can shift the economy between steady states. Labour market beliefs in the model are heterogeneous and persistent, consistent with survey evidence.

Acknowledgements

This work was made possible by the help and support of a great many people. First and foremost, I am indebted to my supervisor Martin Ellison. His guidance brought out the best in my good ideas, and showed me how to distinguish those from the bad. I also thank Michael McMahon for his invaluable comments and advice.

Along the way, Bassel Tarbush and Ian Crawford have been constant sources of support. They have been instrumental in turning me into a researcher, and have been generous with their time and advice at each step of this journey.

I also benefited from numerous conversations with the rest of the Oxford Macroeconomics group. Many members of the faculty, post-docs and DPhil students gave their time to discuss and comment on my work, and I am extremely grateful for their contributions. Outside of Oxford, I was hosted by Filip Matějka at CERGE-EI in 2018, and Julio A. Blanco and John Leahy at the University of Michigan in 2019. I learned a great deal from those visits, and I am very grateful to all three hosts and their students.

The Economic and Social Research Council and the David Walton Distinguished Doctoral Scholarship provided generous financial support, for which I am very thankful.

My friends have been instrumental in supporting me through this process. In particular I would like to thank Alex and the rest of the Rose Lane 1 crowd. Merrilyn, Julian and the others at Nuffield lunch also deserve thanks.

I am also indebted to my parents and sister Chloe. They have proof-read more academic writing than could ever have been expected of them, and without their support this thesis could not have happened.

Finally, I would like to thank Elizabeth. We shared the high points, and she helped me through the low points. Long may that continue.

Contents

1 Cyclical Attention to Saving **1**

1.1 Introduction . . . 2
1.2 Partial Equilibrium Model . . . 7
1.2.1 Savings Products . . . 7
1.2.2 Households . . . 8
1.2.3 Implications . . . 13
1.3 Data . . . 14
1.3.1 Data Sources . . . 14
1.3.2 Institutional Background . . . 16
1.4 Empirical Results . . . 17
1.4.1 Interest Rate Dispersion . . . 17
1.4.2 Constructing φ: a summary statistic for household choice . . . 22
1.4.3 Properties of φ . . . 24
1.5 Quantitative Assessment . . . 27
1.5.1 Model . . . 27
1.5.2 Estimation . . . 30
1.5.3 Results: Amplification from attention . . . 31
1.5.4 Discussion and Policy Implications . . . 33
1.6 Conclusion . . . 34

2 The Attention Trap: Rational Inattention, Inequality, and Fiscal Policy 36

2.1 Introduction . . . 37
2.2 Model . . . 41
2.2.1 Savings Products . . . 42
2.2.2 Households . . . 42
2.2.3 Firms . . . 44
2.2.4 Fiscal and Monetary Policy . . . 46
2.2.5 Market Clearing . . . 46
2.3 Steady States . . . 47
2.3.1 Existence of two steady states . . . 47

2.3.2 Instability with respect to idiosyncratic shocks . . . 50
2.4 Fiscal Policy . . . 52
2.4.1 The Comparison Model . . . 53
2.4.2 Debt funded tax cuts . . . 54
2.5 Low Interest Rate Dispersion . . . 58
2.6 Conclusion . . . 60

3 A Rational Inattention Unemployment Trap 61
3.1 Introduction . . . 62
3.2 Static Model . . . 66
3.2.1 Households . . . 66
3.2.2 Rational Inattention Problem . . . 67
3.2.3 Rational Inattention Solution . . . 70
3.2.4 Aggregate Consumption . . . 71
3.2.5 Firms . . . 72
3.2.6 Equilibrium . . . 73
3.2.7 Requirements for multiplicity . . . 75
3.2.8 Dynamic Solution . . . 77
3.2.9 Stability under belief mutations . . . 83
3.3 HANK model . . . 86
3.3.1 Households . . . 87
3.3.2 Firms . . . 88
3.3.3 Rational Inattention . . . 88
3.4 Survey Expectations . . . 90
3.5 Conclusion . . . 93

Bibliography 94

APPENDIX 103

A Appendix to Chapter 1 104
A.1 Alternative Mechanisms . . . 104
A.1.1 The core mechanism in alternative models . . . 104
A.1.2 Loans and misallocation . . . 106
A.2 Proofs . . . 108
A.2.1 The household FOCs are sufficient for utility maximisation . . . 108
A.2.2 Equilibrium interest rates rise when attention increases . . . 110
A.3 Persistent bank costs . . . 111

A.3.1 Modelling persistent bank costs . . . 111
A.3.2 Persistence of interest rate rankings in the data . . . 115
A.4 Relationship between attention and φ . . . 116
A.4.1 N=2 banks . . . 116
A.4.2 N>2 banks . . . 117
A.5 Robustness for section 1.4 . . . 120
A.5.1 Alternative measures of φ . . . 120
A.5.2 Market size cannot explain fluctuations in φ_t . . . 121
A.6 Estimation Details . . . 124
A.6.1 Data Sources and Treatment . . . 124
A.6.2 Foreign VAR . . . 124
A.6.3 Estimation Results . . . 125

B Appendix to Chapter 2 **129**
B.1 Proofs . . . 129
B.1.1 Derivation of equation 2.3 . . . 129
B.1.2 The household first order conditions are sufficient for utility maximisation . . . 130
B.1.3 Proposition 1 . . . 133
B.1.4 Consumption responds less to tax cuts in the attention trap model 136
B.2 Steady State Transition Simulation . . . 139
B.3 Comparison Model Equations . . . 140
B.4 Calibration for sections 2.4 and 2.5 . . . 141
B.5 Intratemporal Redistribution . . . 142

C Appendix to Chapter 3 **144**
C.1 Aggregate consumption function graphs for the static model . . . 144
C.2 Static Model: Firms . . . 144
C.3 HANK model details . . . 146
C.3.1 Household problem . . . 146
C.3.2 Firm problem . . . 147
C.3.3 Calibration . . . 148
C.3.4 Household problem with rational inattention . . . 149
C.3.5 Solution . . . 150
C.3.6 Transitions . . . 153

List of Tables

1.1 Relative consumption response to shocks with and without variable attention 32

3.1 Prior belief transition matrices 83
3.2 Persistence of hiring rate expectations 93

A.1 Payoff matrix with observed previous state 112
A.2 Bank transition matrices 116
A.3 Business cycle correlations of alternative choice statistics 121
A.4 Proportion of households holding relevant fixed-rate bonds 2006-2010 . . 122
A.5 Proportion of households holding larger fixed-rate bonds 2006-2010 . . . 123
A.6 Education of households holding relevant fixed-rate bonds 2006-2010 . . . 123
A.7 Description of estimated parameters 126
A.8 Estimated parameters in baseline model 127
A.9 Estimated parameters in full information model 127

C.1 Prior belief transition matrices 153

List of Figures

1.1 Distribution of interest rates on fixed-rate bonds, October 2010 19
1.2 Determinants of φ, October 2010 23
1.3 φ over time, 6 month moving averages. 24
1.4 φ against unemployment and average interest rates 25

2.1 Consumption and interest rate effects of a debt-funded aggregate tax cut 56
2.2 Aggregate consumption response to expansionary fiscal policy with high and low interest rate dispersion 59

3.1 Optimal decision rule under rational inattention 70
3.2 Aggregate morning consumption function with various information costs 72
3.3 Equilibrium under full information and rational inattention 74
3.4 Prior beliefs seen in steady state with $\psi = 0.002$ and $\rho = 0.9$. 81
3.5 Steady state morning consumption function 82
3.6 HANK steady states under perfect information and rational inattention . 89
3.7 Distribution of hiring rate expectations, Survey of Consumer Finances . . 91

A.1 Long run equilibrium in persistent cost model, as attention varies 114
A.2 Short run equilibrium in persistent cost model, as attention varies 115
A.3 Consumption and output variance decomposition, full information vs. rational inattention 128

B.1 Model region where representative agent steady state is unstable with respect to idiosyncratic shocks 136
B.2 Transition paths from representative-agent to unequal steady state 140
B.3 Effects of a redistributive tax shock 143

C.1 Aggregate morning consumption function with more information costs . . 144
C.2 Decision rule for $\Pi = 1.005$, $b_s = 0$, with a uniform prior belief. 150
C.3 Prior beliefs seen at steady state with $\psi = 0.0025$, $\rho = 0.9$ 151
C.4 Decision rules for $\Pi = 1.005$, $b_s = 0$. 152
C.6 Hiring rate simulations 154

Chapter 1

Cyclical Attention to Saving*

Abstract

This paper explores the business cycle implications of limited household attention to choosing between different savings products. In a model with heterogeneous banks, savers pay more attention to their bank choice when the marginal utility of income is high. This implies that attention rises in contractions. I find evidence for such countercyclical attention using a novel combination of data on retail savings markets in the UK. In the data, banks offer heterogeneous interest rates on very similar products, and savers more reliably choose products closer to the top of the rate distribution during contractions. Countercyclical attention amplifies shocks to consumption: after a contractionary shock, attention rises, so savers experience higher interest rates, which causes a further fall in consumption. In a quantitative New Keynesian model, this amplification is estimated to be large. Countercyclical attention increases the variance of consumption by 17%, and amplifies some key shocks by more than 25%. Policies that reduce the costs of comparing between financial products have substantial stabilization effects.

*I am grateful to Richard Harrison for kindly sharing the code for Harrison and Oomen (2010), and to Moneyfacts for giving me permission to use their data.

1.1 Introduction

In the majority of dynamic macroeconomic models the interest rate is crucial in determining how shocks propagate through the economy, in part because it regulates the consumption of intertemporally maximising households. The interest rate is usually taken as given by households in these models, but regulators have noted that in reality savers face a range of rate-bearing products, and that they could increase the interest rate they earn on their savings by 'shopping around' for the best product (FCA, 2015).

In this paper I ask if the extent of shopping around, or *attention* to product choice, varies systematically with the business cycle. I find in both theory and data that attention is countercyclical. This substantially amplifies shocks in an estimated business cycle model, because of the effects of attention on the interest rate that households experience.

I first develop a simple model to explore the interaction of rationally inattentive savers[1] and deposit-taking banks. Profit-maximising banks face heterogeneous costs, and in the face of incomplete attention from savers they offer heterogeneous interest rates. If a household pays more attention, they increase their probability of choosing a bank offering a high interest rate, and so they increase the average interest rate they face. The key drivers of attention in this environment are the marginal utility of future income and the extent of interest rate dispersion.

The marginal utility of income drives the countercyclical behaviour of attention, which in turn implies that variable attention amplifies business cycle fluctuations. Consider, for example, a shock that causes consumption to fall. The marginal utility of income rises, and so households pay more attention to their choice of savings product: intuitively, it becomes more important to extract every possible dollar of interest income out of their savings, and so they pay more attention in order to achieve that. That means they face higher interest rates relative to the distribution of offered rates. In addition, with all savers paying more attention the deposit market is more competitive, causing banks to offer higher interest rates. Through these two channels, the household therefore faces higher interest rates than they would have done if attention had stayed constant, and higher interest rates cause consumption to fall even further through a standard consumption Euler equation. Countercyclical attention therefore amplifies the consumption fall.

I find evidence of countercyclical attention to savings using a novel combination of data on savings markets in the UK. Detailed product-level data reveals substantial dispersion in the interest rates banks offer on a set of extremely similar products at any point in

[1] I follow the discrete choice rational inattention framework of Matějka and McKay (2015). I show in appendix A.1.1 that a similar model with endogenous search effort without congestion externalities, as in McKay (2013), gives the same qualitative implications as in the main body of the paper. The extension of persistent interest rates (see appendix A.3.1) would however be intractable in this alternative setup.

time. Linking this with data on the average interest rates achieved by savers opening new products in this set, I show that savers on average choose products higher up in the interest rate distribution in contractions, as predicted by the model. For this analysis I focus on fixed interest rate products, as their simplicity gives me the best chance of ruling out that rate dispersion and saver decisions are being driven by unobservable product differentiation. This should be viewed as a useful laboratory in which to study household behaviour: none of the mechanisms I explore are specific to this market, or to the UK[2].

The existence of interest rate dispersion is an important prerequisite for attention to affect the interest rate households face. I obtain panel data on the savings products available in the UK by digitising monthly editions of Moneyfacts, a magazine for UK financial advisers. There is substantial dispersion in offered interest rates even among products which are identical across the wide range of product features reported. Considering institutional details of the UK savings market, I argue that unobserved product heterogeneity is unlikely to explain the majority of this dispersion. Instead, I argue that much of this interest rate dispersion persists in equilibrium because of an information friction: it is costly for households to acquire information about the set of products on offer. The existence of the Moneyfacts data is itself a justification for the information cost interpretation. Financial advisers (and indeed the Bank of England and several other regulators) would not need to pay for such a magazine if the information was easy to obtain elsewhere[3].

The model predicts that savers should experience higher interest rates relative to this distribution of offered rates in contractions, as they increase their attention. This is precisely what I find in the data. Data from the Bank of England gives the average interest rate achieved on new accounts opened each month for specific sets of savings products with particular characteristics. Identifying the set of products with those characteristics in the Moneyfacts data, I find that the position of the rate households achieve within the distribution of offers is indeed countercyclical. When the unemployment rate is high, and the level of average interest rates in the market is low[4], households on average choose products that are higher up within the distribution of interest rates.

To quantify the importance of countercyclical attention for shock transmission, I build a medium-scale small open economy New Keynesian model of the UK based on that of

[2]The mechanism through which variable attention to savings products affects the business cycle does not necessarily apply in the same way to loans. In addition, I only consider the effects of countercyclical attention to savings on consumers, leaving aside the question of what this means for the allocation of credit. I discuss these additional channels further in appendix A.1.2.

[3]Staff at the regulator found that shopping around decisions were indeed driven by an analysis of the costs and benefits, including time spent shopping and likely interest rate gains (Cook et al., 2002).

[4]In the estimated model I allow for distinct channels through which each of these could affect attention. I find that for most contractionary supply shocks - after which unemployment and interest rates both rise - attention increases. The marginal utility of income is the dominant driver of attention.

Harrison and Oomen (2010), featuring many of the frictions that have become standard in quantitative macroeconomics. To this I add the interaction from the simple model: heterogeneous banks sell domestic bonds to rationally inattentive households. I estimate the model using standard macroeconomic data and key series from the savings data in the empirical part of the paper.

This quantitative exercise is possible because of the novel theoretical approach developed in the simple model. Existing macroeconomic models with limited shopping around for prices or interest rates mostly have households engaging in costly search following Burdett and Judd (1983) (e.g. McKay (2013), Kaplan and Menzio (2016)), but outside of simple cases these are not usually tractable enough to estimate. I retain many of the qualitative features of the Burdett-Judd model[5], while keeping the model sufficiently tractable that the interaction of households and banks can be embedded into a quantitative DSGE model, and solved and estimated using standard techniques.

I find that variable attention amplifies the consumption impact of most shocks[6], as in the simple model. This effect is substantial: the consumption response in the estimated model (cumulated over a year) to government spending and TFP shocks is 43% and 28% larger respectively than if attention is held at steady state. These two shocks explain the largest shares of consumption variation. Overall, the variance of consumption is 17% larger in the baseline model than if variable attention is shut off in this way.

The presence of this amplification has an important policy implication. The extra volatility due to variable attention can be substantially reduced if the marginal cost of information is reduced. Halving the cost of information reduces the variance of consumption by 10%. Policies aimed at providing households with information and facilitating easy product comparisons in this market could therefore lead to lower business cycle volatility.

An additional implication of countercyclical attention is that it can explain a substantial portion of the risk premium shocks typically found to be important in estimated macroeconomic models. Changes in attention affect the model in the same way as risk premium shocks: they change the interest rate faced by the household relative to the policy rate from the central bank. The difference is that attention is an endogenous choice variable. It is not that risk premium shocks cause recessions, but that other kinds of contractionary shock cause attention to rise. An estimation of the model without information costs finds that risk premium shocks explain 25% and 19% of the variance of

[5]The exception is that firms (banks) in my model are not identical, so interest rate heterogeneity is only partly determined by attention, and partly by cost heterogeneity. However, this cost heterogeneity is in fact useful to help the model match the behaviour of interest rate dispersion over the business cycle.

[6]A small number of shocks have large effects of on interest rate dispersion that just outweigh the marginal utility effect discussed above, in which cases attention falls with consumption, weakening the shock slightly. Only one of these accounts for a non-negligible share of consumption and output volatility.

consumption and output respectively. With inattention these shares both fall to 10%. The extra volatility is attributed to supply shocks, notably TFP and price markup shocks, and there is also a larger role given to government spending shocks.

Related Literature. There is a very large literature studying how information frictions affect the business cycle. Many of these papers study frictions in the information agents receive about continuously distributed exogenous shocks (e.g. Maćkowiak and Wiederholt, 2015)[7], or in agents' information about the reaction functions of other agents and the relationships of endogenous variables to shocks (e.g. Eusepi and Preston, 2011)[8]. Unlike these papers, the friction I study is over the discrete choice of which bank to use for saving each period.

Specifically, I draw on Matějka and McKay (2015), who show that an agent facing a discrete choice problem with information costs as in the rational inattention literature (Sims, 2003) will choose information such that choice probabilities resemble those from a multinomial logit model, with a 'twist' that reflects the influence of prior beliefs. This form of inattention is used by Dasgupta and Mondria (2018) to study trade shocks in a model in which countries are inattentive to which country to import each good from, and by Acharya and Wee (2019) to examine information frictions in hiring decisions in a search-based labour market model. I extend the rational inattention literature by showing that information frictions in product choices can have substantial implications for the business cycle.

Another way of modelling the friction in financial product choice would be to use costly search or shopping effort. Coibion, Gorodnichenko and Hong (2015) find that households spend more time and effort shopping for groceries when unemployment rises, echoing my findings that attention to savings product choices rises in contractions. Similarly, the model in Kaplan and Menzio (2016) has unemployed households searching harder for low goods prices, so average search effort rises in recessions. The choice of how much attention to pay to the savings product choice in this paper can be seen as an extension of this literature to financial products, which have particular importance for the business cycle as they influence the intertemporal allocation of consumption.

I also contribute to the literature on the drivers of the business cycle, by showing that countercyclical attention provides a structural interpretation for the risk premium shock that is commonly found to be important in estimated macroeconomics models[9] (e.g.

[7]In general, agents are assumed to understand the model, so tracking exogenous or endogenous variables are equivalent as agents can perfectly map between them. Other examples of models of this kind include Adam (2007, 2009), Angeletos and La'O (2009, 2020), Auclert, Rognlie, and Straub (2020), Lorenzoni (2010), Maćkowiak and Wiederholt (2009), Paciello and Wiederholt (2014), and Reis (2006a, 2006b).

[8]See Eusepi and Preston (2018) for a review of these models.

[9]The lack of a clear structural explanation for this shock is one argument Chari, Kehoe and McGrattan

Smets and Wouters (2007), Christiano, Eichenbaum, and Trabandt (2015)). Attention, however, is not exogenous, but is a response to other variables. Explaining risk premium shocks in this way increases the estimated role of supply shocks in the business cycle.

Several other papers have also studied the role of search or information frictions in financial products. A literature starting with Arrow (1987) finds that information frictions are helpful in explaining wealth inequality, as wealthier households have more incentive to process information about saving and investment choices, and so make better choices and earn higher rates of return on average than less wealthy households[10]. In contrast, I focus on a model with a representative household and show that the same information friction amplifies the effect of aggregate shocks on consumption.

Other papers have considered the importance of deposit market frictions for the business cycle. Diebold and Sharpe (1990) and Driscoll and Judson (2013) document significant stickiness in the pass-through from wholesale interest rates to retail deposit rates. Drechsler, Savov and Schnabl (2017) find that this limited pass-through is critical in the transmission of monetary policy, through the effects of policy on bank balance sheets. The mechanism I explore focuses on the effects of deposit frictions on households through their intertemporal consumption decisions, so is a complement to this channel.

Several empirical studies have found evidence of inattention in retail financial markets. Martín-Oliver, Salas-Fumás, and Saurina (2009) and Branzoli (2016) show that interest rate dispersion and the incidence of choice 'mistakes' (choosing an unambiguously dominated product) respectively are less common when consumers have more incentive to pay attention to their product choice. Adams et al. (2019) find substantial inattention to savings product choices in a large information-provision experiment with savers at five retail banks in the UK. I contribute to this literature by studying how that inattention varies over the business cycle, and showing the macroeconomic consequences of that variation.

In a closely related paper, Yankov (2018) shows that estimated search costs are substantial in a structural model of the US deposit market based on Burdett and Judd (1983). I use similar data on the menu of interest rates available for the UK, but I extend his work by combining it with data on how households choose within that menu. The more model-free empirical approach this allows also suggests that costly information is important for household decisions over savings products, bolstering Yankov's argument[11]. I also differ from Yankov by focusing on the business cycle variation in attention, rather

(2009) give against the Smets-Wouters model. Fisher (2015) shows theoretically that it can be interpreted as a shock to the utility of holding safe assets.

[10]Campanale (2007), Lei (2019), and Lusardi, Michaud and Mitchell (2017) show that this is quantitatively important for explaining wealth inequality. McKay (2013) studies the importance of the wealth-search link in the welfare consequences of social security privatisation.

[11]Interestingly, Yankov finds that attention rises when interest rates rise, in contrast to my results. I discuss explanations for this difference in section 1.4.

than whether search frictions can explain observed dispersion.

The rest of the paper is organised as follows: I develop a partial equilibrium model of rationally inattentive households and heterogeneous banks in section 1.2. In section 1.3 I detail the data sources I use, and some institutional background on UK savings markets. I examine this data, showing the dispersion in interest rates and studying household choices within that distribution in section 1.4. In section 1.5 I quantify the impact of variable attention on the business cycle by estimating a medium-scale New Keynesian model of the UK incorporating the interaction modelled in section 1.2. Section 1.6 concludes.

1.2 Partial Equilibrium Model

In this section I build a simple partial equilibrium model of rationally inattentive households and heterogeneous banks. Households can pay a utility cost to obtain more information about which of a finite set of banks is offering the best interest rate each period. With more information they will achieve a higher interest rate relative to the distribution of rates on offer. I show that attention (the quantity of information processed) is driven by the marginal utility of income and the dispersion of interest rates. Contractions therefore cause attention to rise, because the marginal utility of income rises. Higher attention makes the deposit market more competitive, so banks increase the rates they offer, further increasing the interest rates households experience. Higher interest rates cause further falls in consumption through the household Euler equation.

1.2.1 Savings Products

To generate interest rate dispersion in the model, I assume that households buy government bonds through banks, some of whom are more efficient than others. Inefficient (high cost) banks offer lower interest rates than their efficient competitors in equilibrium.

There are N banks. Each period t, each bank n buys bonds from the government and sells them on to individuals, both at price 1. In the next period, the government pays the bank $1 + i_t^{CB}$ per bond bought, and the bank pays $1 + i_t^n$ to the individuals it sold to. Bank n also pays a transaction cost χ_t^n per bond. In this partial equilibrium exercise the policy rate is exogenous, but it is endogenous in the quantitative model in section 1.5.

Bank n chooses the interest rate they offer to individuals i_t^n to maximise profits, taking into account that their market share[12] will depend on how their interest rate compares with the distribution of rates offered by the other banks. Denoting the probability that

[12]Since all savers are identical, the market share equals the probability a saver chooses that bank.

a saver chooses bank n given a rate distribution as $\Pr(n|\hat{i}_t^n, i_t^{-n})$, the bank problem is:

$$i_t^n = \arg\max_{\hat{i}_t^n} \Pr(n|\hat{i}_t^n, i_t^{-n}) \cdot (i_t^{CB} - \hat{i}_t^n - \chi_t^n) \tag{1.1}$$

This gives the first order condition[13]:

$$\frac{d}{di_t^n} \Pr(n|i_t^n, i_t^{-n}) \cdot (i_t^{CB} - i_t^n - \chi_t^n) = \Pr(n|i_t^n, i_t^{-n}) \tag{1.2}$$

Interest rates are dispersed if costs χ_t^n are dispersed. A bank with higher costs will choose lower interest rates, accepting a lower market share to prevent a larger fall in their markup.

1.2.2 Households

In this simple model, households only choose two things each period: consumption, and how much attention to pay to choosing between the different banks. Attention increases the interest rate the household achieves relative to the distribution of rates on offer.

Specifically, I assume that there is a large representative household composed of many individuals. Each period the household decides how much each individual will consume and save, and how much attention they will pay to the choice of savings products, to maximise expected lifetime utility. As in the rational inattention literature, 'attention' in this model refers to the amount of information that each individual processes about the banks before deciding which bank to choose for their portion of the household's saving. All asset income is redistributed among individuals each period, so there is no inequality within the household. The household problem is therefore:

$$\max_{c_t, b_t, i_t^e} \sum_t \beta^t \big(u(c_t) - \mu \mathcal{I}(i_t^e)\big) \tag{1.3}$$

subject to

$$c_t + b_t = b_{t-1}(1 + i_{t-1}^e) + y_t \tag{1.4}$$

$$\mathcal{I}'(i_t^e) > 0, \mathcal{I}''(i_t^e) > 0 \tag{1.5}$$

Here c_t is consumption, b_t is real bond holdings, i_t^e is the effective interest rate faced by the household (the average over individuals), and y_t is exogenous income.

The novel element of this problem is the term $\mathcal{I}(i_t^e)$, the amount of attention required for the household to earn an effective interest rate i_t^e on assets bought in period t (which

[13]The market share function derived in section 1.2.2 is smooth as long as savers have less than full information, and is such that equation 1.2 is sufficient for profit maximisation.

pay off in $t+1$), which will be derived below[14]. The key properties of this function are expressed in condition 1.5: if the household pays more attention they will earn a higher rate of interest, but the interest rate gain from more attention diminishes as attention grows. Households choose how much attention to pay by balancing the expected future marginal utility of higher interest income with the costs of attention. I have modelled the costs of attention as a simple additively separable utility cost, with a constant marginal cost μ, as is common in the rational inattention literature (see for example Maćkowiak and Wiederholt (2009)). This can be thought of as costly cognitive effort. In appendix A.1.1 I show that monetary costs lead to the same qualitative conclusions.

In the maximisation I allow the household to directly choose the effective interest rate they face, rather than choosing the amount of attention to pay. These two specifications are equivalent, but the first order condition on the effective interest rate has a more readily interpretable form than a first order condition on attention.

The first order conditions comprise an Euler equation and a first order condition on the effective interest rate:

$$u'(c_t) = \beta \mathbb{E}_t (1 + i_t^e) u'(c_{t+1}) \tag{1.6}$$

$$\beta b_t \mathbb{E}_t u'(c_{t+1}) = \mu \mathcal{I}'(i_t^e) \tag{1.7}$$

The first order condition on effective interest rates (equation 1.7) is crucial in unpicking this model. It shows that households choose attention to equalise the marginal utility of higher asset returns next period with the marginal cost of the attention required to achieve it. The marginal utility of higher asset returns is simply the marginal utility of income in the next period multiplied by the amount saved. Attention therefore rises when consumption is expected to be low, as then the marginal utility of future income rises. It is marginal utility in the following period that matters because the bank choices made in period t only change income when the bonds bought pay out in period $t+1$.

Attention also rises when the marginal information needed to increase effective interest rates ($\mathcal{I}'(i_t^e)$) falls, as this reduces the marginal cost of increasing asset income. After deriving the relationship between attention and effective interest rates below I show that this marginal cost falls when interest rate dispersion rises, as a small increase in attention leads to a small improvement in the probability of an individual choosing a high interest rate bank. With more dispersed interest rates, this improvement has a larger impact on effective interest rates, and so the extra attention needed to marginally increase i_t^e is low.

The first order condition on effective interest rates also implies that a wealthier household will choose to process more information, and so will experience a higher interest rate.

[14]This function will depend on the distribution of interest rates offered by banks, but I have dropped this dependence from the written function to save on notation.

This encourages further saving through the Euler equation[15] (equation 1.6), but the non-concavity this implies is small enough at plausible parameter values that the first order conditions remain sufficient for utility maximisation(see appendix A.2.1). I assume that households are net savers (government bonds are in positive net supply), so $b_t > 0$ and the household always chooses to process some information.

I now turn to the derivation of $\mathcal{I}(i_t^e)$, from the decisions of individuals, who face a discrete choice rational inattention problem as studied in Matějka and McKay (2015).

Individuals start the period with a prior belief about the probabilities of different states of the world, i.e. interest rate distributions and the positions of banks within that. I assume that individuals share information on returning to the household at the end of the period, so all individuals have the same priors. I denote the probability that an individual chooses bank n if they process no information and rely only on their priors as $\mathcal{P}_n$. Following Matějka, Steiner and Stewart (2017) I refer to this unconditional choice probability as the 'predisposition' towards bank n.

In general, though, individuals will process some information before choosing a bank. They have access to an infinite set of information about banks. If an individual processed enough of that information before making their bank choice - if they paid enough attention - they would be able to precisely identify the best interest rate in the market and choose it with probability 1. However, because attention is costly, the household chooses to limit the amount of information each individual can process before choosing their bank. Intuitively, each individual could visit every bank in the market and observe their interest rate, and so correctly identify the best product in the market, but doing so requires a great deal of effort and so is prohibitively costly. I further assume that individuals cannot share information within the period.

There are therefore two challenges facing an individual. Using terminology from Matějka and McKay (2015), an individual must decide on an *information strategy* (what kinds of information to process given their limited attention capacity) and an *action strategy* (how to translate that information into a bank choice). Formally, we can write this as the individual choosing the covariance of a noisy signal and the true distribution of banks, subject to a constraint on the amount of information about the bank distribution the signal can contain. The individual then observes a realisation from that noisy signal, updates their beliefs and chooses a bank. The quantity of information embodied in a particular signal structure is defined (following Sims (2003)) as the expected reduction in

[15]The interaction between attention and wealth implies that the model actually has two steady states, one in which all households are identical and another in which some households are wealthy and attentive, while others remain at the borrowing constraint paying no attention. As the data in section 1.4 is only informative about average household choices, I study the model with identical households. I study the two-agent steady state in a related model in chapter 2.

entropy between the prior and the posterior formed after observing the signal realisation.

Using lemma 1 from Matějka and McKay (2015), we can leave the belief distributions and signal structures in the background, and rewrite the individual's problem in terms of conditional choice probabilities[16]. The individual's maximisation problem becomes:

$$\max_{\Pr(n|s_t)} \mathbb{E}_{s_t}\Big(\sum_{n=1}^{N} i_{s_t}^{n} \Pr(n|s_t)\Big) \text{ subject to} \tag{1.8}$$

$$\mathcal{I}_t = -\sum_{n=1}^{N} \mathcal{P}_n \log(\mathcal{P}_n) + \mathbb{E}_{s_t} \sum_{n=1}^{N} \Pr(n|s_t) \log(\Pr(n|s_t)) \tag{1.9}$$

The choice variable $\Pr(n|s_t)$ is the probability that the individual chooses bank n given the state of the world is s_t, where s_t summarises the interest rate distribution and the order of banks within it. The individual chooses a decision rule (a set of conditional choice probabilities for each possible ranking of banks s_t) in order to maximise their expected interest rate, as the redistribution of asset income across individuals each period renders them risk neutral over interest rates. They maximise subject to the constraint that $\Pr(n|s_t)$ cannot deviate too far from the predisposition $\mathcal{P}_n$, the choice probability that would be observed for bank n if the individual had access to no more information than their prior. The more attention the household allows individuals to pay, the more their conditional choice probabilities can deviate from these predispositions, towards the unconstrained choice rule in which $\Pr(n|s_t) = 1$ if bank n offers the highest interest rate in state s_t, and $\Pr(n|s_t) = 0$ otherwise.

Solving the individual's rational inattention problem gives a familiar multinomial logit choice rule:

$$\Pr(n|i_t^n, i_t^{-n}) = \frac{\mathcal{P}_n \exp(\frac{i_t^n}{\lambda_t})}{\sum_{k=1}^{N} \mathcal{P}_k \exp(\frac{i_t^k}{\lambda_t})} \tag{1.10}$$

Here I have replaced the notation for a state of the world s_t with the interest rate distribution in time t, made up of the rate offered by bank n and the rates at all of their competitors. The variable λ_t is the Lagrange multiplier on the attention constraint 1.9 in the individual problem, or the shadow value of information. As the household increases attention, holding all else equal the constraint becomes less binding and the shadow value of information falls[17].

[16] See Matějka and McKay (2015). Intuitively, it is never optimal to use information processing capacity on two distinct signal realisations that imply the same action, so there is a one-to-one mapping from signal realisations to actions. We can therefore solve the problem by looking at actions (i.e. choice probabilities) rather than explicitly solving for signals.

[17] Attention could rise without any change in λ_t if the dispersion of interest rates rises. Then the marginal benefit of information for the individual (in terms of increasing their expected interest rate) rises, so if the shadow price of information stays constant attention will rise.

The household decides how much each individual will save before knowing whether they have chosen a bank offering a high or low interest rate. Combined with the income sharing around the household, this means that all individuals save the same amount, and the interest rate the household faces across all of their saving is the expected interest rate achieved by each individual's bank choice. It is this average rate that I refer to as the effective interest rate i_t^e:

$$i_t^e = \sum_{k=1}^{N} i_t^k \Pr(k|i_t^k, i_t^{-k}) \tag{1.11}$$

Substituting out for the optimal conditional choice probabilities using equation 1.10, this becomes:

$$i_t^e = \frac{\sum_{k=1}^{N} i_t^k \mathcal{P}_k \exp(\frac{i_t^k}{\lambda_t})}{\sum_{k=1}^{N} \mathcal{P}_k \exp(\frac{i_t^k}{\lambda_t})} \tag{1.12}$$

As attention increases (λ_t falls), individuals successfully choose higher interest rate banks with a greater probability, and so the effective rate experienced by the household rises. The effective interest rate is therefore an increasing function of the probability of successfully choosing high interest savings products, and information processing $\mathcal{I}$ increases when individuals are more discriminating between banks. Therefore $\mathcal{I}'(i_t^e) > 0$. Diminishing returns to attention ensure that $\mathcal{I}''(i_t^e) > 0$.

The predispositions are where this rational inattention model allows for more flexibility than the search-based models explored in appendix A.1.1. If there is some reason, aside from the current interest rate, for individuals to be more likely to choose one bank than another, that can simply be incorporated into $\mathcal{P}_n$. The model can therefore incorporate some banks having more 'brand recognition' than others, and so attracting inattentive individuals with a higher probability. It can also allow for persistence in the ordering of banks, in which case knowledge of past states of the world is informative about the current state, and so affects the prior probability of choosing particular banks. While it is possible to construct a search model with bank-specific variation in the probability of individuals meeting each bank, as would be necessary to account for these situations, models of this type quickly become intractable (see e.g. Menzio and Trachter (2015) for a setting with one large seller and a continuum of identical small sellers).

I study the case of persistence in bank costs, and so in the positions of banks within the interest rate distribution, in appendix A.3.1. For the modelling in the main body of the paper however I assume for simplicity that bank costs, and so interest rate rankings, have no persistence[18], and no banks have more brand power than others. The predispositions

[18]There is in fact very little persistence in bank interest rate rankings for the products studied in section 1.4 (see appendix A.3.2), though this may not be true of all assets. The Burdett-Judd models common in this literature also have no persistence, as all price-setters follow identical mixed strategies.

then all equal $1/N$, and the conditional choice probabilities and effective interest rate become:

$$\Pr(n|i_t^n, i_t^{-n}) = \frac{\exp(\frac{i_t^n}{\lambda_t})}{\sum_{k=1}^{N} \exp(\frac{i_t^k}{\lambda_t})} \tag{1.13}$$

$$i_t^e = \frac{\sum_{k=1}^{N} i_t^k \exp(\frac{i_t^k}{\lambda_t})}{\sum_{k=1}^{N} \exp(\frac{i_t^k}{\lambda_t})} \tag{1.14}$$

1.2.3 Implications

I now analyse the novel channel in this model. A shock that causes consumption to fall leads to higher attention, and so households face higher interest rates relative to the distribution of offers. That distribution shifts up as the deposit market gets more competitive. Through both channels household effective interest rates rise, which through the Euler equation amplifies the consumption fall.

The key equations are the consumption Euler equation (equation 1.6), the first order condition on attention (equation 1.7), and the bank profit maximisation condition (equation 1.2). The bank condition was left in section 1.2.1 in terms of the probability of savers choosing each bank. Substituting in the conditional choice probabilities from equation 1.13 this becomes:

$$\big(1 - \Pr(n|s_t)\big) \cdot (i_t^{CB} - i_t^n - \chi_t^n) = \lambda_t \tag{1.15}$$

Similarly, using these choice probabilities we have that $\mathcal{I}'(i_t^e) = \lambda_t^{-1}$. The household first order condition on attention therefore becomes:

$$\beta b_t \mathbb{E}_t u'(c_{t+1}) = \mu \lambda_t^{-1} \tag{1.16}$$

This shows formally that attention rises with interest rate dispersion through a lower $\mathcal{I}'(i_t^e)$, as claimed in section 1.2.2. As interest rates become more dispersed, the shadow value of information λ rises. If consumption and saving are constant, equation 1.16 implies that the household responds by increasing attention, partially relaxing the information constraint in the individual problem, and so bringing λ back down. As I abstract from movements in asset supply[19] (b_t), the only other cause of attention variation is the expected marginal utility of income in the following period.

If attention rises, then the bank first order condition (equation 1.15) implies that the distribution of interest rates will shift up (proof in appendix A.2.2). Intuitively, higher

[19] I treat them as constant here, but in the quantitative model the shocks to μ are isomorphic to shocks to b_t, so it can be thought of as being subject to exogenous shocks in section 1.5.

attention means that the demand facing an individual bank becomes more elastic to changes in that bank's interest rate relative to their competitors, as choice probabilities can depend more on specific realisations of interest rates. With more elastic demand, markups decrease, and so the interest rates offered to households rise relative to the policy rate. Furthermore, each bank wants to increase their interest rates to keep pace with rate rises at their competitors, because interest rates are strategic complements in this market[20].

Therefore if a shock causes consumption to fall, households increase attention, which means they experience higher interest rates than if attention had remained constant through two channels. First, the probability of an individual choosing high interest rate banks rises, increasing the effective interest rate relative to the distribution of rates on offer. Second, the increased competition in the deposit market causes banks to increase the interest rates they offer, so the rate distribution shifts up. Through the consumption Euler equation, this encourages households to delay consumption, and so consumption falls by even more than it would have done without an attention change. Variable attention therefore amplifies shocks to consumption, unless the shock also reduces interest rate dispersion so much that attention actually falls. In section 1.5 I find that this is rare, so on average variable attention amplifies the consumption effect of shocks.

1.3 Data

To provide evidence on cyclical attention to savings, I combine data from two sources. To observe the choice set facing households, I digitise 14 years (1996-2009) of monthly editions of Moneyfacts, a magazine for UK financial advisers. To observe household choices within that set, I combine this with data on average interest rates earned on newly opened savings products each month from the Bank of England. In this section I explain the nature of these datasets, and provide some institutional background on the specific savings market I study.

1.3.1 Data Sources

Each month Moneyfacts magazine publishes tables of the interest rates and product characteristics of the vast majority[21] of saving and credit products on offer from retail

[20]In appendix A.2.2 I also show that interest rate dispersion falls when attention rises. Qualitatively, a rise in attention therefore has the same effect as a rise in search effort in Burdett and Judd (1983).

[21]The publishers aim to cover the universe of products, but acknowledge that they may occasionally miss a niche product from a small provider. As I focus on average household choices in a common product category (fixed interest rate saving bonds), the data should contain all relevant products.

financial institutions in the UK. A key advantage of this data is that it reports all observable dimensions of product heterogeneity which are relevant for savers, which means that the interest rate dispersion I find after controlling for all of these characteristics cannot be explained by observable product differentiation. The magazine reports the full set of relevant characteristics because it is designed for household financial advisers: if savers care about a product characteristic then financial advisers need to know about it.

Of all of the saving (and borrowing) products available in the data, I focus on the specific subset of fixed interest rate savings products, for which the product characteristics are simple and easily quantifiable. This enables me to account for product heterogeneity. In contrast, mortgages and other loans, as well as other more complicated savings products, have many more dimensions of product heterogeneity, and many products have their own idiosyncratic features, made evident by the paragraph of notes accompanying each observation in the data. Such idiosyncrasies would make accounting for product differentiation in interest rate dispersion extremely difficult. Furthermore, Moneyfacts only reports the advertised interest rates on products: for savings products the vast majority of households receive this rate, but for loans there is the potential for adjustments based on the risk of the individual borrower, so it is not possible to cleanly identify the choice set for loans in the data[22]. Further details on fixed interest rate savings products are given in section 1.3.2.

Household choices within this market are reported in the Quoted Household Interest Rate published by the Bank of England. This gives the average interest rate earned by households each month on a subset of fixed interest rate savings products which are identical along all the dimensions of product heterogeneity identified in Moneyfacts. It therefore directly relates to a set of products which are identical except for the interest rate, and which can be easily identified in the Moneyfacts data. Importantly, the average interest rate reported is for accounts opened in that month only, not the stock of all active accounts, which would include accounts opened in previous months when interest rates were different.

There are several Quoted Household Interest Rate series available for fixed rate savings products with different combinations of product characteristics. I focus on the series for products with a term of one year, an investment of £5000, and where interest is paid annually, because the Quoted Household Interest Rate series goes back to 1996 for these products, whereas the series for other combinations of features have only been published since 2009. This is one of the most common combinations of product features in the market.

[22]A further advantage of studying fixed interest rate savings products is that they are mainstream products throughout the sample. In contrast, ISAs (another commonly used savings product in the UK) were only introduced in 1999.

1.3.2 Institutional Background

Retail savings products are provided in the UK by conventional banks and building societies, which offer deposit products to fund mortgage lending[23]. Deposits at all of the institutions in the data were covered by deposit insurance up to £35,000 throughout the period studied, substantially above the £5,000 investment size of the products considered (I return to the issue of deposit insurance and bank risk in section 1.4.1). The largest four institutions had 74% of the market for current accounts in 2000 (Vickers, 2011)[24], and the largest branch networks. The market for savings accounts is much less concentrated, with a Herfindahl-Hirschman Index between 20% and 30% lower than the current account market between 2000 and 2008 (Vickers, 2011).

Fixed interest rate savings products are commonly used in the UK. The FCA (2015) found that 12% of households held these products[25], and they accounted for 20% of all cash savings balances in the UK. Indeed, the Moneyfacts data confirms that the market is large, as there are an average of 200 such products available each month in the sample. Even in the subset of products with the characteristics that match the Quoted Household Interest Rate there are an average of 34 products each month, and all of the major banks are present.

In addition to the advantages of studying fixed interest rate bonds described above, there are two other factors which aid analysis of choices in this particular market. Firstly, product bundling is uncommon in this market. In its 2015 report, the FCA found that 76% of savers using fixed rate bonds use an institution which is not their 'main provider of financial services'. I can observe instances of explicit bundling: if, for example, a particular fixed rate product can only be purchased by someone with a current account at that bank, that is noted in the data. I do not remove the few products for which this is the case before analysing the data because they are not removed in the Quoted Rate data, but removing them does not substantially change the distribution of offered rates.

[23]The main differences between building societies and banks are that building societies are owned by their customers, and are more limited than banks in how much of their funding can come from wholesale money markets. I will not distinguish between the two types of provider as industry experts suggest it is not important for consumer choices (Maundrell, 2017). For a review of the differences see BBC (2005). As the degree of wholesale funding could be related to bank risk, I discuss this in section 1.4.1.

[24]This market share fell gradually to 64% in 2008, then rose to 77% in 2010 due to mergers in which large banks bought failing rivals (Vickers, 2011). Cruickshank (2000) reports that in 1998 the 'big four' banks had just 19% of savings accounts by number of accounts, but this does not account for deposit size so may not accurately reflect their market share.

[25]In the 2006 Wealth and Assets Survey 5% of UK households reported holding these products, and for those that did they constituted 31% of household asset portfolios on average. The discrepancy between this figure and that of the FCA (2015) could stem from many products in this market being advertised as 'fixed rate savings accounts' rather than 'fixed-term investment bonds with fixed interest rates' as they are called in the WAS. Some households holding a fixed-rate bond may have therefore mistakenly described it as a 'savings account' in the WAS.

Second, the interest rate is the key product feature that matters for households in this market. In a survey of holders of fixed interest rate bonds the FCA (2015) found that the interest rate was the most important product feature for the large majority of savers in this market. Savers hold fixed rate savings bonds as assets, not for transactions or any other purposes. This is important for my analysis, as customer service and the convenience of a large branch network are unobservable product features that I cannot easily control for. That these do not matter much to savers means that this is unlikely to explain much of the interest rate dispersion I find in section 1.4.1. The presence of a local branch is less important for these products than others because they are of a fixed maturity, so the saver does not need to interact with the bank on as regular a basis, as is the case for products with the potential for continual adjustment (FCA, 2015). In addition, in the model in section 1.2 I assumed that the only reason households may not choose the highest interest rate product is because of limited information, and in section 1.4.2 I use the interest rate that households achieve relative to the distribution of offers as a measure of the 'success' of their decision making, and argue that variation in this could be driven by variable attention. These exercises assume that savers always prefer a higher interest rate product, so unobservable product features do not weigh heavily on the value households get from their saving products.

1.4 Empirical Results

In this section I explore household choice using the datasets described in section 1.3. First, I show that there is substantial heterogeneity in interest rates offered by retail banks which cannot be explained by product heterogeneity. Without interest rate dispersion, the choice of one savings product over another would have no impact on the interest rate that households experience. I then construct a summary statistic for the 'success' of household choice, which measures the interest rate households actually achieved relative to the distribution of rates on offer that month and is closely related to attention in the model in section 1.2. I show that on average, households more reliably choose higher interest rate products when the unemployment rate is high and the average level of interest rates is low, which is consistent with the model.

1.4.1 Interest Rate Dispersion

Each month in the sample, households could achieve a wide range of interest rates by choosing different products from different providers. The median within-period standard deviation and interquartile range of interest rates are 60 and 75 basis points respectively.

The median within-period average interest rate is 521 basis points.

Clearly, some of this dispersion is due to the fact that these products are not all close substitutes. They differ in the length of the bond, the minimum investment required, and the frequency with which interest is paid. In section 1.4.1 I account for this product heterogeneity, and show that in all months observable product differentiation explains no more than 53% of interest rate dispersion, and in the majority of periods it explains much less than that. I also argue that dimensions of unobservable product heterogeneity, such as perceived bank risk, are unlikely to explain much of the remaining dispersion. I then provide evidence that limited attention is a likely cause of the remaining interest rate dispersion in section 1.4.1. This means that many savers could increase their interest income without changing any other characteristics of their savings product by switching to other providers. Increased attention to the choice of savings products would lead to this kind of switching, which is how attention affects the interest rate households experience.

Interest Rate Dispersion is not explained by Product Differentiation

I show that product differentiation cannot fully explain the substantial interest rate dispersion in the market for fixed interest savings products in two ways. First, for each period I regress the available interest rates on all product characteristics reported by Moneyfacts. The adjusted R^2 of this regression never exceeds 0.53, so in every month a maximum of just over half of the variation in interest rates is explained by observable characteristics. The median adjusted R^2 across the sample is just 0.18. Second, I consider the group of products in the Quoted Rate data, which are identical across all product characteristics except interest rates, and I show that even among these products the mean within-period standard deviation of interest rates is 43 basis points (on a mean interest rate of 520 basis points). Importantly for these exercises, Moneyfacts supplies all relevant observable dimensions of product differentiation, so the remaining interest rate dispersion must be driven by some other factor, such as limited attention.

For the first exercise, I run the following regression each month, where X_i contains all of the product characteristics listed by Moneyfacts for product i in that month:

$$ir_i = \alpha + \beta X_i + \varepsilon_i \tag{1.17}$$

Across all periods, the median adjusted R^2 for this regression is 0.18, and the maximum adjusted R^2 is 0.53. At most, just over half of the variation in interest rates can therefore be explained by observable product characteristics.

The second exercise considers only a set of products which are identical along all dimensions of product heterogeneity recorded by Moneyfacts. If the market is perfectly

competitive, and unobserved product heterogeneity is negligible, the products considered should have the same interest rate. This is not what is observed. Within the set of products which qualify for the Quoted Household Interest Rate data (see section 1.3 for details) the mean within-month standard deviation of interest rates is 43 basis points, on an average interest rate of 520 basis points. In October 2000, as an example, savers could earn annual rates of return between 450 and 680 basis points at different banks on a product with identical characteristics (the standard deviation of rates that month is 44 basis points, which is the median across the sample). The histogram of these rates is plotted in figure 1.1. There is substantial interest rate dispersion which cannot be explained by observable product differentiation.

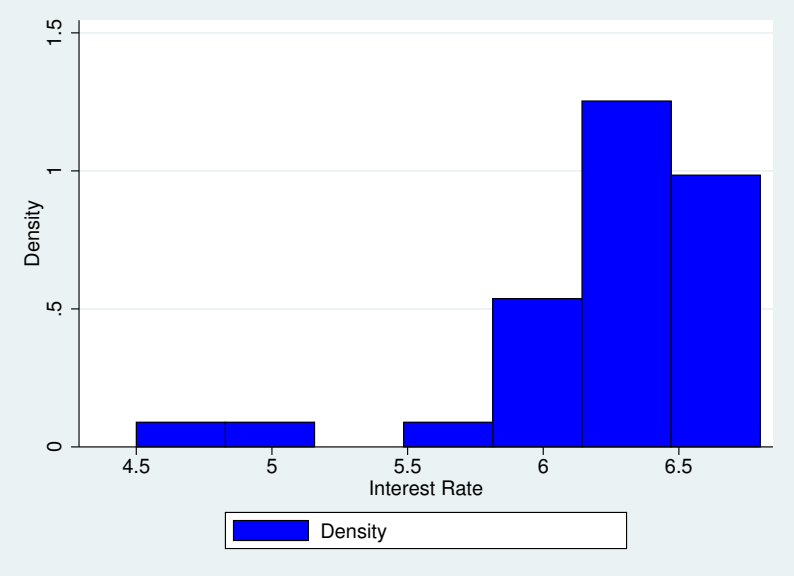


Figure 1.1: Histogram of annual interest rates on fixed interest rate bonds and term accounts on offer in October 2000

These two exercises, however, only control for observable product heterogeneity. The dispersion could be due to unobserved heterogeneity of the products or providers that is known to households. One possible source of such unobserved product differentiation is 'implicit bundling': if households have a preference for saving with the same institution they use for their current account, mortgage, and other financial services, then providers with a smaller range of offerings in other product areas may have to pay higher interest rates to compensate savers for the lack of this convenience. The evidence collected by the regulator on this market suggests that this is not a substantial driver of interest rate heterogeneity (see section 1.3.2)[26].

[26]In fact, the mechanism studied in sections 1.2 and 1.5 could be reinterpreted as households choosing how much convenience to give up in order to achieve higher interest rates, rather than how much costly attention to pay to achieve those higher rates. The intuition and qualitative results would be identical in such a convenience-driven model.

Another potential driver of interest rate dispersion is bank risk. A bank that is more likely to fail might have to offer savers higher interest rates to compensate them for the risk that the savers will lose their deposits. This is unlikely, however, to be a significant driver of rate dispersion in this market. Throughout the sample deposits in the UK are insured up to £35,000 (£50,000 after October 2008) per depositor per provider, which is far above the £5,000 investments I study. This removes the majority of risk to savers of bank failure, though if the insurance was not perfectly credible, or the insurer was expected to be slow in paying back deposits in failed banks, then it does not completely eliminate risk as an explanation for rate dispersion. In addition, Chavaz and Slutzky (2018) find that deposit rates in the UK are on average uncorrelated with a variety of measures of bank risk. As interest rate dispersion is substantial in every month of my sample, this suggests that risk is not the main driver of the dispersion. Chavaz and Slutzky do find that riskier banks offer higher interest rates when they face spikes in household attention (measured by Google searches), primarily during the 2008 financial crisis. This suggests that risk may explain why the dispersion of interest rates among similar products rises during the financial crisis in my data, but their results on the relationship between risk and deposit rates on average imply that other factors must also be at play. This is supported by the fact that including bank fixed effects in regression 1.17 still leaves the mean and median unexplained within-month standard deviation of interest rates at 39 and 37 basis points respectively[27].

There could, of course, still be other sources of unobserved product differentiation which explain the dispersion of interest rates that I have not considered here. I therefore proceed by arguing from the other side, giving evidence that there are substantial costs of information/search in this market, and therefore that limited attention could explain why interest rate dispersion persists in equilibrium.

Limited Attention is a plausible explanation of Interest Rate Dispersion

The presence of costly search, information, or attention has been proposed as an explanation of equilibrium price dispersion in a large number of papers, both theoretical and empirical, starting with Stigler (1961) (see Baye, Morgan and Scholten (2006) for a review). The existence of interest rate dispersion not accounted for by observed product

[27]This is an inferior way of capturing risk than that of Chavaz and Slutzky (2018), who use proprietary time-varying measures of bank risk from the Bank of England. Adding bank fixed effects ignores changes in bank risk over time, and also removes all variation which causes a bank to offer persistently high or low rates, whether that is driven by risk or not. The model implies that banks offer persistently higher (lower) interest rates if they have a disproportionately small (large) weight in household prior beliefs (see appendix A.3.1). The regression with bank fixed effects should therefore be taken as further suggestive evidence that the Chavaz and Slutzky results apply to the fixed-rate market specifically, as well as to retail deposits in general.

differentiation is not, however, evidence in itself that households are less than fully informed about the savings products available to them. I therefore provide evidence that information costs, which lead to inattention, are in fact important in this market.

The clearest piece of evidence for the role of information costs, which would make households inattentive, comes from the FCA (and their predecessor the FSA), who regulate the market for savings products in the UK. In a study of retail financial services for the regulator, Cook et al (2002) concluded that:

> "Shopping around is not cost free since consumers have to spend time and effort. The extent to which consumers shop around the market will depend on the benefits they think they can get and the costs of them doing it."

Other reports by the regulator (FSA (2000), FCA (2015)) on this market have similarly concluded that households could benefit if they searched harder for their financial products, but that such search is costly.

In addition to the remarks of the regulator, the founding of Moneyfacts, the magazine from which I obtain the savings product data, is itself evidence that information costs are substantial in retail financial markets. Moneyfacts was founded to provide data on savings and credit products to financial advisers, because until that point obtaining this information for product comparisons had been difficult (Moneyfacts (2019))[28]. This suggests that it is costly (in time, effort or money) for households to obtain this information from elsewhere: the magazine would not have been founded, and would not keep selling subscriptions, if data on the full set of available savings products was easy to find. Since less than 8% of UK households employ financial advisers (Aegon, 2017) the existence of the magazine has not itself removed the information friction behind saver inattention.

The rapid spread of comparison websites covering savings products in the early 2000s supports this evidence. The largest such comparison site in the UK, MoneySupermarket, brought in £105 million of revenue in 2006 because very large numbers of people visited the site[29] each month to compare a variety of products, including savings products (Connon, 2007). Savers would not need to visit a comparison website if they were already fully informed about the products on offer. However, as with the founding of Moneyfacts, these websites did not reduce the cost of information to zero. It still takes time and effort to use the websites, to process the information and translate it to choices. Indeed, in 2019 the Financial Times ran an article about one bank's strategy for attracting depositors

[28]Similarly, a rival comparison service MoneySupermarket began in 1987 when the founder realised that it was very difficult for brokers to compare available mortgage deals (Hohler, 2007).

[29]By the end of 2006 they had 4 million users per year (Hohler, 2007). MoneySupermarket earns revenue by charging firms each time a consumer clicks through from the site to that firm's product (Connon, 2007). High revenues therefore reflect large numbers of households using the comparison site.

titled "How Monzo is banking on customer apathy" (Kelly, 2019), indicating that savers are not fully attentive to their choices despite the availability of comparison websites.

Other authors have also concluded that inattention plays an important role in retail financial product markets. Martín-Oliver, Salas-Fumás, and Saurina (2009) find evidence that there is less interest rate dispersion among Spanish banks in markets where households have a greater incentive to pay attention. Branzoli (2016) finds that fewer consumers make the mistake of choosing a product which is strictly dominated by another product at the same bank when they have a greater incentive to pay attention to their choices. For the UK, Adams et al. (2019) find evidence of substantial inattention to savings product choices in a large randomised controlled trial using savers at five retail financial institutions.

Finally, I will discuss below how the endogenous attention decisions studied in the model in section 1.2 can explain the time series variation in how households choose from among the set of offered rates.

1.4.2 Constructing φ: a summary statistic for household choice

In this section I use the Moneyfacts and Bank of England data to study how successful households are at choosing the highest interest rate product in the market each month. To do this I compute for each month the difference between the average interest rate earned by households opening new accounts and a benchmark rate, the average interest rate on offer at the four largest banks. I argue that a saver paying no attention would face this benchmark rate on average, and any increase in the rates savers face above this can be seen as an improvement in their choices. Normalising this difference by the standard deviation of interest rates on offer that month ensures that the measure is not mechanically affected by changes in the dispersion of interest rates, and gives a statistic that while model-free in construction is closely related to attention in the model.

I construct the 'no-attention' benchmark interest rate to reflect a probable predisposition (to use the language of section 1.2) towards larger market players: small 'challenger' banks are likely to be discovered only if the saver does some careful research, as they do not have large numbers of physical branches or large advertising budgets. Specifically, I construct the benchmark rate by taking the average interest rate on offer from the 'big four' banks[30]. Throughout the sample period these four banks hold most of the market share in many retail banking markets, including current accounts and mortgages, and have many more branches than other banks (Office of Fair Trading, 2008). Using this as the benchmark interest rate assumes that households paying no attention to their choice

[30]These are Barclays, HSBC, Lloyds, and Royal Bank of Scotland.

of savings product are likely to go to their closest bank branch, or the bank where they hold a current account. Alternative benchmarks, such as weighting banks by their number of branches or the size of their balance sheets, would be strongly correlated with this simple benchmark because the big four consistently dominate others on these metrics.

Figure 1.2 shows the histogram of interest rates available in October 2000 on the subset of fixed interest rate savings products which appear in the Quoted Household Interest Data, with the benchmark interest rate shown in red and the quoted rate (the average interest rate achieved on products bought that month) shown in green. The benchmark rate is 106 basis points below the maximum rate that households could achieve[31]. While they do not all get that rate, on average savers do somewhat better than they would have if they paid no attention to their choice, earning an average of 6.24% interest, 50 basis points above the benchmark rate.

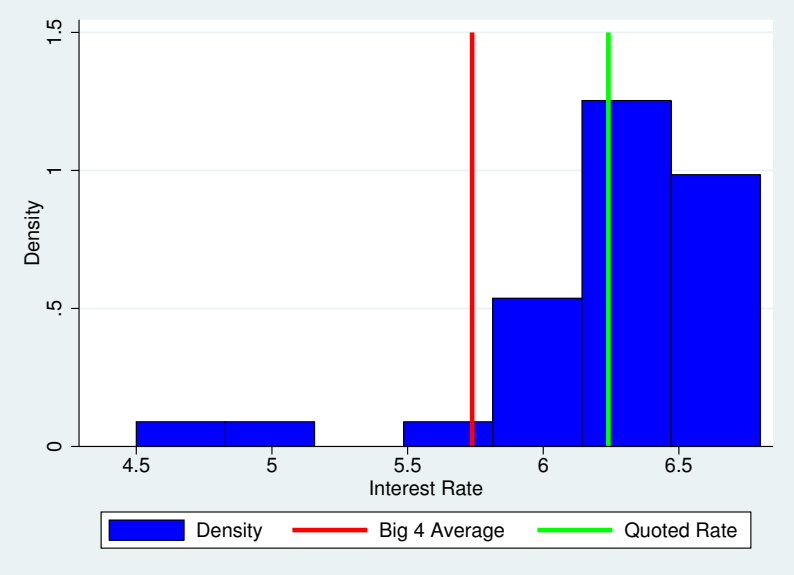


Figure 1.2: Histogram of annual interest rates on fixed interest rate bonds and term accounts on offer in October 2000

The statistic on household choice which I will study is the distance between the household mean and the benchmark rate, normalised by the standard deviation of the interest rate distribution that month. I denote the resulting statistic by φ:

$$\varphi_t := \frac{\mathbb{E}_h i_t - i_t^b}{\sigma(i_t)} \tag{1.18}$$

φ is a summary statistic on how households chose from among a distribution of interest rates in a given month. Note that φ is homogeneous of degree zero in interest rates, so market-wide trends in the level of nominal interest rates do not mechanically affect φ.

[31] The highest rate on offer in October 2000 from a big four bank was 6.1%.

If household decisions are driven by real interest rates rather than nominal rates, φ is unaffected by changes in inflation expectations for the same reason.

Although this statistic is not derived using any particular model, it is closely related[32] to attention in the model in sections 1.2 and 1.5. In section 1.2.2 I showed that when a household pays more attention the effective interest rate they experience rises relative to what they would have achieved if they processed no information and simply followed their predispositions. This corresponds to a rise in the average rate achieved by households relative to the benchmark rate, and so a rise in φ. I also showed that attention is only a function of conditional choice probabilities, so if interest rates all move further apart but choice probabilities stay the same attention has not changed. Normalising the gap between the average achieved rate and the benchmark rate by the standard deviation of interest rates ensures that changes in rate dispersion do not mechanically alter φ.

1.4.3 Properties of φ

The key innovation of φ relative to existing estimates of information processing in savings markets, aside from not having to rely on any specific structural modelling assumptions, is that it can be measured each month. I can therefore study how choice behaviour changes over time, at a high enough frequency to observe co-movements with aggregate variables over the business cycle. In the graphs below I plot the time series of φ, and show that it is positively correlated with the unemployment rate and negatively correlated with the level of interest rates over the business cycle.



Figure 1.3: φ over time, 6 month moving averages.

[32] In appendix A.4 I show that there is an exact correspondence between φ and attention in the model with two banks, and that attention and φ remain closely related with more banks in the market.

A substantial portion of the variation occurs at business cycle frequencies[33]. The largest falls in φ occur during the growth periods of 2004-2005 and 2006-mid 2008. Shortly after the beginning of the Great Recession in the UK in mid-2008, φ began to rise sharply. There was also a substantial rise in φ from July 2001 - April 2002. Although the UK avoided recession during this period, it was a time of slowing growth, and the unemployment rate rose relative to trend.

These observations suggest a countercyclical pattern in φ, which is confirmed in figure 1.4. These plots show the (HP-filtered) cyclical component of φ against the cyclical components of the average interest rate in the savings market studied[34] and in unemployment. Lower interest rates and higher unemployment are associated with higher φ. These relationships are strongly statistically significant: the slope coefficients on interest rates and unemployment both have p-values below 0.01%.

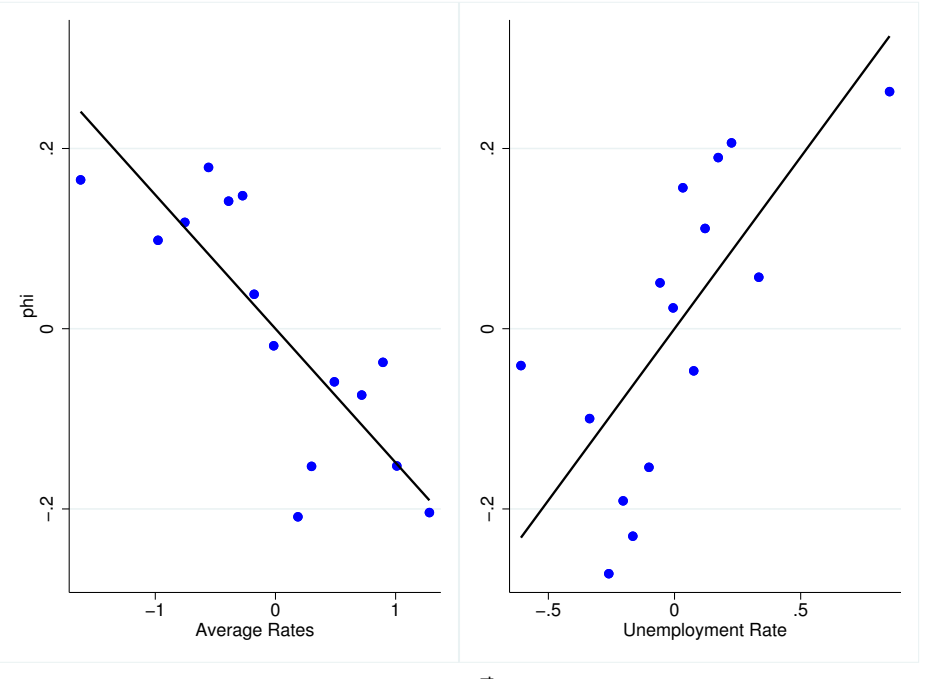


Figure 1.4: φ against (unweighted) average interest rates among products considered in the Quoted Household Interest Rate data and unemployment. All series are cyclical components after HP filtering. Black solid lines are from linear regressions, which give $\hat{\varphi} = -0.149\hat{\bar{i}} + 2.7\text{e-}9$, $R^2 = 0.146$ and $\hat{\varphi} = 0.381\hat{u} + 9.3\text{e-}10$, $R^2 = 0.170$ respectively. Blue circles are averages of φ and the regressor of interest within groups of observations, grouped by their position within the distribution of the regressor.

[33]Figure 1.3 plots a moving average of φ to aid visualization. 20% of the sample spectral mass of the unsmoothed series lies between 6 and 50 quarters, the business cycle frequency domain suggested by Beaudry, Galizia and Portier (2020).

[34]I use the unweighted mean of the interest rates in the market, but all rates tend to move together here, so using the benchmark rate, the Quoted Household Interest Rate, or the interest rate on one year UK treasury bills makes little quantitative difference, and no difference to the qualitative conclusions.

When interest rates are high and unemployment is low, savers choose products with low interest rates, close to those offered by the big four banks. As rates fall and unemployment rises, households move up through the distribution of offered rates, more reliably choosing the higher interest rate products in the market, and so achieving higher interest rates relative to the distribution of offers than they did when average rates were high and unemployment was low[35]. In appendix A.5.1 I obtain the same result using alternative versions of φ. In particular, I show that in contractions the average interest rate achieved by households moves closer to the highest interest rate on offer in the market, as well as increasing away from the benchmark rate. I keep to this measure of φ in the main body of the paper, however, because of its close correspondence to attention in the model, which enables it to discipline the role of attention in the model.

These cyclical patterns can be explained by the household attention decisions studied in section 1.2. In recessions, consumption tends to be low, so the marginal utility of interest income is high, increasing the incentives to pay attention[36]. In addition, when average rates are low in this market the dispersion of interest rates tends to be high, increasing the benefits of attention[37]. Finally, if there is a 'search for yield' motive, i.e. if there is something about low levels of interest rates that make households want to work harder to increase their returns, this would also encourage greater attention, and so higher φ, when average rates are low[38]. In the model in sections 1.2 and 1.5 I allow for the first two channels to operate, leaving examination of the search for yield mechanism for future work.

I have argued above that bank risk does not play a large role in this market. In appendix A.5.2 I show that changes in the size and composition of the fixed-rate bond market are also unable to explain the cyclical patterns in φ_t. While other explanations of the data are in principle possible, variable attention is therefore the leading candidate.

These results contrast with those of Yankov (2018), who finds that attention is high

[35]These relationships are not driven by extreme events. The linear relationship between interest rates and φ is not significantly different before and after the run on Northern Rock in September 2007. The relationship between φ and unemployment is significantly steeper before the crisis than after, but it remains positive and strongly significant in both sub-samples.

[36]Similarly, when unemployment is high the opportunity cost of time spent shopping around is low. This does not feature explicitly in the model as the cost of attention is a simple additively separable utility cost. However, for most shocks consumption and output co-move, and so labour supply rises with consumption, in which case this opportunity cost of time channel is qualitatively the same as the marginal utility of income channel.

[37]$Corr(\bar{i}, \sigma(i)) = -0.3$ and is significant at the 0.1% level. This correlation is partly driven by the substantial increase in interest rate dispersion during the crisis, which may be partly due to heightened awareness of bank risk (see section 1.4.1). However, this correlation remains negative and significant if I exclude the crisis periods.

[38]Search for yield often refers to financial institutions taking on more risk to increase their returns when yields are low (e.g. Martinez-Miera and Repullo, 2017). This is somewhat different from the search for yield mentioned here, as in this setting there is no change in the riskiness of household investments.

when rates are *high*. There are two reasons for this. First, the interest rate dispersion he documents in US markets is positively correlated with the level of interest rates, which is not the case in my UK data[39]. This means that incentives to search rise with interest rates in his data, but not in mine. Second, to keep the Burdett-Judd equilibrium tractable enough to estimate he studies a two-period model in which a saver's only source of income is interest on their assets. This means that consumption does not vary much in the model, so the main driver of search is the dispersion of interest rates. To the extent that interest rates and consumption are positively correlated, marginal utility will rise when interest rates fall, implying the benefits of search will rise. In section 1.5 I find that for the UK data the marginal utility of income is indeed a key driver of information choices.

1.5 Quantitative Assessment

In this section I study the quantitative significance of cyclical attention to saving in an estimated DSGE model for the UK. Cyclical attention amplifies the consumption response to most shocks, as the marginal utility of income channel described in section 1.2 is estimated to be powerful. This amplification is substantial, increasing the consumption response to government spending and TFP shocks (the two shocks explaining the largest fraction of consumption variance) by 43% and 28% respectively. Overall, the variance of consumption is 17% higher than if attention is held fixed at steady state.

Changes in attention affect the model in the same way as risk premium shocks, and can in fact explain a substantial portion of the business cycle fluctuations otherwise attributed to the risk premium. The important difference is that attention is an endogenous response to other shocks, and so can be influenced by policy. In particular, the majority of the stabilisation effects of holding attention constant can be achieved by reducing the cost of information μ: reducing μ by 50% reduces the variance of consumption by 10%.

1.5.1 Model

Since the data in sections 1.3 and 1.4 concerns savings markets in the UK, I base the model on the medium-scale DSGE model for the UK of Harrison and Oomen (2010).

[39]This may be because he looks at a range of CD's with different sizes of investment. Estimates of equation 1.17 show that the size of the investment is strongly related to the rate on offer. If the differences between rates on low and high balances grow when the level of rates rise, this could explain some of the discrepancy.

Full Information Block

The model is a medium-scale small open economy New Keynesian model, with many of the frictions that have become standard in the quantitative macroeconomics literature. Households consume domestic goods and imports, monopolistically supply differentiated labour varieties, and save through risk-free domestic and foreign bonds, money, and by investing in capital which they rent to firms. They can vary capital utilisation at a cost. They face external consumption habits, capital adjustment costs, nominal wage adjustment costs (with partial indexation to past wages), and portfolio adjustment costs that introduce a friction in holdings of foreign bonds.

Domestic firms hire utilisation-adjusted capital services and labour to monopolistically produce intermediate goods, which are aggregated by perfectly competitive final goods firms who supply home and export markets. Intermediate goods firms face price adjustment costs with partial indexation to past prices, with different adjustment costs for the home and export markets.

A monetary authority sets the interest rate on domestic government bonds following a Taylor Rule with interest rate persistence. The fiscal authority issues a positive amount of bonds, engages in wasteful government spending, and collects lump sum taxes. With full information the model features Ricardian Equivalence. With rational inattention a debt increase only affects consumption because it increases the incentives to pay attention to savings. Changes in debt are therefore isomorphic to changes in the cost of attention μ (see equation 1.16), and so without loss of generality I fix the supply of (real) bonds at 1, and allow for shocks to μ. I refer to these as 'attention shocks' below, but they could equally be interpreted as shocks to government debt.

Foreign variables (inflation, export demand, relative export prices, interest rates) are assumed to follow a VAR process estimated outside of the model, as in Adolfson et al (2007). Details of this are in appendix A.6.

There are 11 shocks outside of the information problem: to TFP, government spending, the disutility of labour, the capital adjustment cost, the consumption Euler equation (risk premium shock), the price markup on domestic goods, the nominal interest rate (monetary policy shock), and to each of the four international variables.

The only changes I make to the Harrison and Oomen (2010) model, aside from the introduction of inattention to savings as set out below, are that I use a risk premium shock rather than a discount factor shock, and I change the frequency of the model to monthly to make best use of my monthly attention data[40]. For further details on the

[40]Changing the frequency of the model does not affect the equations, but does mean I adjust some of the calibrated parameters and priors used in the estimation of the model from those used by Harrison and Oomen.

model setup please therefore see Harrison and Oomen (2010).

Attention Problem

As in section 1.2, I assume that the household is made up of many individuals, who each purchase their domestic bonds from one of a finite number of banks. The banks' problem is as in section 1.2.1. To keep the estimation simple I set the number of banks to 2. The information problem only affects the market for domestic bonds, not foreign bonds or capital.

The only difference this makes to the existing household FOCs is that I replace the nominal policy rate in the consumption Euler equation, the FOC on capital and the money demand equation with the nominal effective interest rate, averaged over individuals in the household. Foreign exchange market participants can buy bonds directly from governments, so the interest rate that matters for UIP is the policy rate.

I assume that each period a ranking of banks is drawn. One bank, which I will refer to as the 'good' bank and index by the superscript g, draws a low cost $\chi_t^g = 0$. The other bank draws a high cost, and so I will refer to them as the 'bad' bank (superscript b). They face $\chi_t^b = \tau_1 + \tau_2(i_t^{CB} - \bar{i}^{CB}) + e^{\zeta_{\tau,t}} > 0$. I allow for this cost to depend on the policy rate as a reduced-form way for the model to capture the observed correlation of interest rate dispersion with the level of policy rates. The mean-zero AR(1) shock $\zeta_{\tau,t}$ causes exogenous fluctuations in interest rate dispersion. There is no persistence in the bank cost rankings: each bank has a 50% probability of drawing the low costs each period.

As in section 1.2, households choose how much attention individuals pay to choosing between banks. More attention increases the effective interest rate by improving the probability that an individual will choose a high-rate bank, but it comes at an additively-separable utility cost with a constant marginal cost.

The first order condition on attention therefore takes a similar form to equation 1.16, with the only differences being that μ is now subject to a mean-zero AR(1) shock process $\zeta_{\mu,t}$, the stock of saving is set to 1, and the future marginal utility of income is affected by inflation:

$$\beta \mathbb{E}_t \frac{U'(c_{t+1})}{\Pi_{t+1}} = \mu e^{\zeta_{\mu,t}} \lambda_t^{-1} \tag{1.19}$$

Each individual faces a discrete choice rational inattention problem over the two banks. Since there is no persistence in the rankings of costs faced by banks, and so in the positions of each bank in the interest rate distribution, individuals have uninformative priors. Solving the rational inattention problem, we therefore have that the probability

of choosing bank n given that bank n is the good bank that period is p_t^g:

$$p_t^g = \frac{\exp(\frac{i_t^g}{\lambda_t})}{\exp(\frac{i_t^g}{\lambda_t}) + \exp(\frac{i_t^b}{\lambda_t})} \tag{1.20}$$

The effective interest rate faced by the household is the average over individuals:

$$i_t^e = p_t^g i_t^g + (1 - p_t^g) i_t^b \tag{1.21}$$

Banks choose interest rates to maximise expected profits. Their first order condition is the same as equation 1.15 derived in section 1.2, which for the good and bad bank respectively reduces to:

$$(1 - p_t^g) \cdot (i_t^{CB} - i_t^g) = \lambda_t \tag{1.22}$$

$$p_t^g \cdot (i_t^{CB}(1 - \tau_2) - i_t^b - (\tau_1 - \tau_2 \bar{i}^{CB}) - e^{\zeta_{\tau,t}}) = \lambda_t \tag{1.23}$$

Bank profits and transaction costs are redistributed back to the representative household as a lump sum.

There are therefore 5 new variables not in the Harrison and Oomen (2010) model: i_t^e, λ_t, p_t^g, i_t^g, i_t^b. The new equations are the first order condition on attention (equation 1.19), the choice probability rule (equation 1.20), the definition of i_t^e (equation 1.21), and the two bank first order conditions (equations 1.22 and 1.23). There are two new shocks, to attention ($\zeta_{\mu,t}$) and rate dispersion ($\zeta_{\tau,t}$).

1.5.2 Estimation

I conduct a Bayesian Maximum Likelihood estimation of the model solved to a log-linear approximation. There are 11 standard observable variables: GDP, consumption, inflation, the 3-month treasury bill rate, investment, real wages, hours worked, and foreign inflation, industrial production, interest rates, and relative export prices. The foreign variables are trade-weighted averages of the other G7 countries. On top of these I add 3 observables from the Moneyfacts data: the mean and standard deviation of deposit rates, and the choice statistic φ. I use data from 1993-2009.

The three extra observables require three extra shocks: I have already introduced AR(1) shocks to the cost of attention and the dispersion of bank costs, and for the final shock I allow for i.i.d. measurement error for the mean rate[41].

I follow Harrison and Oomen in setting some parameters to match standard values

[41]This is preferred by the estimation to including a structural shock to mean interest rates relative to the policy rate. The measurement error does not explain a substantial fraction of the mean rate series in the estimation.

or long-run features of UK data (details in appendix A.6). I do the same for τ_1, the constant in the bank cost function, choosing it to match the steady state dispersion of interest rates in the Moneyfacts data.

For the priors on each variable to be estimated I again follow Harrison and Oomen, adjusting them as required for the change in the frequency of the model. The only new parameters to estimate are the cost of attention μ, the cyclicality of bank costs τ_2, and the persistence and volatility of the new shocks. μ must be greater than 0, but there are no such restrictions on τ_2. I choose relatively weak priors for both in the absence of strong evidence for the values they should take[42].

1.5.3 Results: Amplification from attention

The key novel parameters in the estimation are the cost of information μ and the cyclicality of bank cost dispersion τ_2, which are estimated to be 0.0124 and -0.1394, both significantly different from zero, with much tighter estimates than the prior distributions. To interpret these estimates I compare the estimated model to an alternative with the same equations and parameters, but where attention is held at its steady state each period. Switching off cyclical variation in attention in this way substantially weakens the transmission of shocks through the economy, so variable attention amplifies shocks.

In table 1.1 I report the magnitude of the cumulative response of consumption to each shock over a year in the static attention alternative, relative to the baseline estimated model. A value below 1 implies that consumption responds by less to the shock in the fixed attention model than with variable attention. The shocks are ordered according to the share of consumption volatility they explain, with the most important shock first[43].

For most shocks, consumption is substantially less responsive when attention is held at its steady state. For government spending and TFP shocks, which together explain 67% of consumption volatility in the baseline estimated model, attention variation amplifies the consumption response by 43% and 28% respectively. Overall, the variance of consumption is 17% larger with variable attention than if attention is held at steady state.

[42]Specifically, I choose a gamma prior for μ, with mean 0.01 and standard deviation 0.5, and a uniform prior for τ_2 with mean 0 and standard deviation 0.5.

[43]Attention and dispersion shocks are excluded because attention shocks have no effect with fixed attention, and dispersion shocks become isomorphic to risk premium shocks.

Table 1.1: Cumulative consumption response to shocks relative to variable attention baseline.

Shock	Fixed Attention
Govt spending	0.699
TFP	0.783
Markup	1.042
Risk premium	0.949
Foreign demand	0.744
Monetary policy	1.015
Labour disutility	1.046
K adj. cost	0.982
Foreign ir	0.869
Export price	0.447
Foreign inflation	0.495

The intuition is as in section 1.2: when a shock causes consumption to fall, the marginal utility of income rises, so attention rises. More attention increases the effective interest rate within the distribution of offers, and causes that distribution to shift up. The household experiences higher interest rates, and reduces consumption even further.

This is also amplified by a further general equilibrium effect not seen in section 1.2. After a contractionary shock, variable attention reduces output and inflation relative to where they would be with fixed attention. The monetary authority therefore sets a lower policy rate than with fixed attention. Since τ_2 is estimated to be negative, this lower policy rate leads to greater interest rate dispersion, encouraging even more attention[44].

Amplification from variable attention remains substantial even though the information problem only applies to a subset of the household portfolio, due to a set of no-arbitrage conditions. For households to hold all types of assets the expected benefits of holding them must all be equal. If the household pays more attention to domestic bonds and so increases their interest rate there, the rate on other assets must adjust to match, and so it does not matter that the information problem does not apply to the whole portfolio. In fact, capital provides an extra channel through which attention amplifies fluctuations: when attention rises the interest rate on domestic bonds exceeds the expected return on capital, so investment drops until the returns are equalised, adding to the contraction.

For three shocks, however, shutting off variable attention leads to larger consumption responses, though only the price markup shock plays a substantial role in consumption

[44]The implication that attention is higher when the level of interest rates is lower fits with the empirical results in section 1.4. It also offers a potential alternative explanation for the finding in Drechsler, Savov and Schnabl (2017) that deposit rates rise relative to treasury bill rates when the level of interest rates falls. They attribute this to market power in deposit markets, but it could also be that households pay more attention when rates fall, causing banks to increase rates on average.

fluctuations[45]. This is because interest rate dispersion falls when policy rates rise. If there is a shock that causes a small consumption fall but a large rise in the policy rate, then this dispersion effect will dominate and attention will fall. In this case the interest rates households experience will fall relative to the fixed attention case, mitigating the initial fall in consumption. This dispersion effect is small enough that for most shocks causing consumption and interest rates to move in opposite directions (e.g. G, TFP) the marginal utility of income effect dominates and attention amplifies the shock. However, for markup, monetary policy, and labour disutility shocks there is a large change in inflation, and so in policy rates. Attention therefore comoves positively with consumption, dampening the shock, though only by a small amount as the marginal utility effect is strong enough to prevent large changes in attention.

Variable attention therefore amplifies the response of consumption to most shocks. For shocks that cause consumption and output to co-move, such as TFP shocks, then this also amplifies the output response. For other shocks, however, output and consumption move in opposite directions (e.g. government spending shocks), and in those cases the amplification of the consumption effect mitigates the output response to the shock.

1.5.4 Discussion and Policy Implications

An alternative way to understand the effect of variable attention on consumption is to compare it with a risk premium shock. This shock disturbs a wedge between the interest rate experienced by households and the policy rate, which is precisely the effect of a change in attention to savings[46]. The key difference between attention and risk premium shocks is that attention is an endogenous household choice, so is influenced by policy.

In fact, the correspondence between attention and risk premium shocks means that variable attention can provide a structural explanation of risk premium shocks, which is often absent in DSGE models despite the prominent role for these shocks in fitting such models to the data (see Fisher (2015) for an alternative interpretation). To see the quantitative ability of variable attention to explain risk premium shocks, I compare the baseline estimated model with an otherwise identical model without information frictions. The full information model is estimated in the same way as the baseline, with the same data except for the variables associated with the attention problem.

[45]Price markup shocks account for 17.2% of consumption variance in the baseline model, where monetary policy and labour disutility shocks account for 1.2% and 0.9% respectively.

[46]This is as long as the profits and transaction costs of banks are transferred back to the household lump sum. If instead the transaction costs are treated as waste they would enter the goods market clearing condition and so the resource constraint would be affected by changes in attention, while it isn't by risk premium shocks. Since the quantitative exercise finds that transaction costs at banks are very small relative to output this effect is at most small.

With no information friction, the risk premium shock explains 25% of the variance of consumption, and 19% of the variance of output[47]. Moving to the baseline model with inattention the risk premium shock becomes substantially less important, explaining 10% of both consumption and output.

Cyclical attention can therefore plausibly explain approximately half of the business cycle volatility otherwise attributed to risk premium shocks in the UK. Very little of the fall in the importance of risk premium shocks is made up for by shocks to attention, which explain just 1% of consumption and output variance in the baseline model. This half of the risk premium shock is therefore mostly explained by an endogenous response of attention to other shocks. In particular, the share of consumption and output variance explained by TFP and price markup shocks increases when adding the information friction. Government spending also explains a greater share of consumption variation. Full results are in appendix A.6.

Importantly, endogenous attention choices can be affected by policy, where exogenous risk premium shocks cannot. One policy that has an intuitive effect on attention is to reduce the cost of information, for example through financial education programmes or regulation to ensure clearer disclosure and presentation of bank pricing policies.

After a permanent fall in the cost of information μ, households pay more attention to savings in steady state. This reduces the amplification from variable attention through two channels. First, attention becomes more sharply convex in effective interest rates at higher levels of attention ($\mathcal{I}''(i_t^e)$ increases), and so fluctuations in the marginal utility of income produce smaller fluctuations in attention. Second, greater attention reduces the equilibrium dispersion of interest rates, which reduces the impact of attention fluctuations on effective interest rates. For these reasons, reducing μ by 50% (and keeping all other parameters as in the estimated model) reduces the variance of consumption by 10%.

1.6 Conclusion

I have presented a novel channel through which aggregate shocks affect consumption. In theory and in data, households are more successful at choosing higher interest rate savings products in contractions, because they pay more attention to their choice when the marginal utility of income is high. An improvement in these savings choices increases the interest rate households face, and so causes current consumption to fall as households postpone more consumption to the future. Countercyclical variation in attention therefore amplifies the consumption response to the shocks that drive the business cycle.

[47]It explains the second largest share of consumption variance (after government spending shocks), and the third largest share of output variance (after TFP and foreign demand shocks).

In an estimated model of the UK economy, variable attention amplifies the effect of aggregate shocks on consumption: the variance of consumption is 17% higher than it would be if attention remained constant, and the effect of cyclical attention on some specific shocks is substantially larger than that. Variable attention also explains approximately half of the business cycle fluctuations attributed to risk premium shocks in a full information version of the model.

Since attention, unlike the risk premium shock, is an endogenous choice made by households, it can be affected by policy. In particular, policies aimed at making it easier for households to 'shop around' for financial products could reduce business cycle volatility, providing another argument in favour of policies such as financial education and clear disclosure of bank pricing policies.

Chapter 2

The Attention Trap: Rational Inattention, Inequality, and Fiscal Policy

Abstract

I show that when households face costs of acquiring information about asset returns a simple New Keynesian model with ex-ante identical households features a steady state with persistent inequality. The representative agent steady state is unstable with respect to idiosyncratic shocks. The unequal steady state has a two-agent (TANK) structure, where inequality is maintained by endogenous household attention decisions. Asset-poor households gain little from paying attention to finding savings products with high returns, so don't pay attention to the choice between products and earn lower interest rates than wealthier households. This interest rate heterogeneity gives rise to steady state wealth inequality between households with identical preferences and market access. Importantly, attention decisions and interest rate heterogeneity are endogenous objects. I show that after a fiscal expansion which is sufficiently large to temporarily lift poor households slightly off their borrowing constraints, aggregate consumption rises by less and is more persistent in this 'attention trap' model than in a comparable model where wealth inequality is driven by heterogeneous discount factors.

2.1 Introduction

A recent literature has demonstrated that the effects of fiscal stimulus packages are crucially dependent on the distributional consequences of the policy. However, models in this literature usually assume that wealth inequality is driven by exogenous factors which are unaffected by fiscal policy: poor households are at their borrowing constraint because they do not have access to financial markets (Bilbiie, 2019a; Debortoli and Galí, 2018) or because they are less patient than 'saver' households (Bilbiie, Monacelli, and Perotti 2013). In contrast, an empirical literature has highlighted a variety of factors contributing to wealth inequality which are not so clearly fixed in the face of policy changes, such as heterogeneity in asset returns (Benhabib, Bisin, and Luo (2018), Fagereng et al (2019)). An established literature starting with Arrow (1987) suggests that some of this return heterogeneity exists because wealthier households have more to gain from higher returns, so invest more in information about assets before deciding where to invest.

In this paper I ask how the effects of fiscal policy vary when wealth inequality is assumed to be driven by heterogeneous asset returns, linked to wealth through optimal information choices. Crucially, the information a household chooses to obtain about their potential investment choices is not exogenous to fiscal policy. Compared to a model in which inequality is driven by discount factor heterogeneity, both intratemporal redistribution and debt-funded tax cuts have smaller effects on aggregate consumption on impact, but the effects are more persistent.

To study the endogenous response of information choices to fiscal policy, and how that impacts aggregate consumption, I construct a two-agent New Keynesian (TANK) model which differs from most TANK models in that households all have identical preferences and access to asset markets. Instead, households must choose how much information to process before choosing between different savings products, some of which offer higher interest rates than others. I refer to information processing as attention following the literature on rational inattention (e.g. Sims (2003)). Attention is costly, but with more attention the household is more likely to identify which savings products are offering the best interest rates.

This attention decision leads endogenously to a two-agent structure because attention choices depend on wealth in an interaction I term the 'attention trap'. Wealthier households have more to gain from increasing their interest rate because that higher rate operates over a larger pool of assets, and so they pay more attention and earn higher interest rates than poorer households: wealth gives a household more 'skin in the game'. This is the mechanism proposed by Arrow (1987). If two households are not on exactly equal wealth, the less wealthy household will experience lower interest rates, and so will

save less, than the wealthier household. The difference in interest rates therefore acts just like the heterogeneous discount factors in Bilbiie, Monacelli and Perotti (2013) (hereafter BMP), and as in their model it leads to a steady state with two types of households, wealthy households on their Euler equation and poor households against a borrowing constraint. The poor households do not save, even in steady state, because they optimally pay no attention to choosing between saving products and so face very low interest rates.

The model also has a representative agent steady state, in which all households save the same amount and therefore pay the same amount of attention as each other. The attention trap, however, renders this steady state unstable with respect to idiosyncratic shocks. An arbitrarily small shock that affects a subset of households will cause the marginally poorer households to pay less attention, earn lower interest rates, and save less, until they reach their borrowing constraint[1]. Which households end up constrained is therefore random, whereas in many TANK models it is driven by ex-ante differences in preferences or access to asset markets.

I use this model to study the effect of fiscal policy shocks, and compare it to a TANK model in which the inequality is driven by differences in household discount factors. Fiscal policy is a particularly interesting policy to study in part because of its increasing relevance while interest rates remain close to their effective lower bounds in many economies, but also because the effects of fiscal policy in theoretical models have been shown to depend substantially on the inclusion of household heterogeneity (see e.g. BMP, Kaplan and Violante (2018)). In this paper I ask if those effects also depend on the cause of that heterogeneity.

If the fiscal policy shocks are small, I find that they have the same impact in the attention trap and heterogeneous discount factor models. Poor households remain at their borrowing constraints, and do not therefore vary their attention with policy.

If, however, the policy shock is large enough that it temporarily lifts poor households slightly off their borrowing constraint then the response of aggregate consumption to the policy is considerably smaller on impact and more persistent in the attention trap model. This is because the interest rate heterogeneity that returns the wealth distribution to its steady state after the policy shock in the attention trap model is endogenous to the policy: the fiscal expansion leads to poor households saving a small amount of assets, which in turn means they increase their attention to choosing between savings products,

[1]This same mechanism would operate even if the representative agent steady state has households holding zero assets, as long as households are allowed to borrow up to some positive limit. A small idiosyncratic shock pushes the adversely affected households to borrow, and so to pay attention to finding lower interest rates, while the wealthier households pay attention to finding higher interest rates. The interest rate heterogeneity and resulting unequal steady state therefore arise just as in the model with positive net saving from the representative household.

and so the interest rates they face rise further than in the heterogeneous discount factor model. This means that poor households save more of their tax windfall in the attention trap model, and engage in more consumption smoothing, which reduces the impact of the policy on consumption, and makes the consumption response more persistent.

It is important to study larger policy shocks for two reasons. First, very large fiscal stimuli have been used recently to combat macroeconomic crises. The American Recovery and Reinvestment Act of 2009 expanded US government debt by $787 billion, playing a key role in the 18.6 percentage point rise in the US government debt-to-GDP ratio between 2009 and 2012 (Dupor, 2017). In March 2020 the U.S. House of Representatives passed an even larger stimulus to combat the economic effects of the Covid-19 pandemic, which amounted to more that 10% of US GDP and involved sending $1,200 directly to every adult and $500 to each child, mirroring the lump sum tax reductions studied in this paper. These stimuli are substantially larger than any increase in debt-to-GDP considered in this paper. If real interest rates remain persistently low, such large fiscal expansions may become more common as monetary policy struggles to stabilise demand in recessions (Blanchard and Summers, 2019). Second, while the model only contains two types of households, in reality there is a rich distribution of wealth, with some households close to lifting off their borrowing constraints each period. Even small fiscal expansions will therefore lead to some households starting to save, and the analysis in this paper suggests that they play an important role in the effects of fiscal policy on consumption.

A further implication of the attention trap model is that when interest rate dispersion among saving opportunities falls the power of fiscal policy to influence consumption also falls. This is because interest rate heterogeneity is what keeps poor households at their borrowing constraint. If the range of interest rates on offer shrinks then poor households paying no attention will face an interest rate closer to that faced by wealthy households, and so they will start to save after a smaller fiscal stimulus. For a given expansionary shock, poor households save more of their tax windfall, and smooth out their consumption more, if interest rate dispersion is lower. This is relevant for policymakers because Eggertsson et al (2019) find evidence that banks are unwilling to pass on negative interest rates to households even when policy rates are below zero, which means we should expect to see lower interest rate dispersion when average interest rates are close to zero because the zero lower bound prevents banks from offering rates too far below the average. Fiscal policy may not be as able to fill the gap left by monetary policy at the zero lower bound as Blanchard and Summers (2019) and others have suggested.

Related Literature: This paper contributes principally to the literature on the role of household heterogeneity in macroeconomic policy. Kaplan, Moll, and Violante (2018) show that wealth heterogeneity substantially increases the power of monetary pol-

icy in a rich heterogeneous-agent New Keynesian (HANK) model[2], and Debortoli and Galí (2018) show that a simpler two-agent (TANK) model, with households divided into standard intertemporal optimisers and hand-to-mouth consumers, can match most of this effect. The model I build has a similar two-agent structure, but in contrast to Debortoli and Galí (2018) the hand-to-mouth households in my model are at their borrowing constraint because of endogenous attention decisions. I also focus on fiscal policy rather than monetary policy, as in Bilbiie, Monacelli and Perotti (2013). I extend their work by studying a two-agent model where household heterogeneity is endogenous, and so responds to fiscal policies. The attention trap model could also be used to analyse how monetary policy affects aggregate demand when wealth inequality is due to endogenous attention choices, but that exercise would require a careful examination of how attention responds to average interest rate changes[3].

The link between attention to saving choices and wealth at the heart of the attention trap was first proposed by Arrow (1987). Since then, a number of papers have assessed the quantitative importance of this mechanism for wealth inequality in rich heterogeneous agent models (Campanale, 2007, Lusardi, Michaud, and Mitchell, 2017). McKay (2013) also uses a model of this kind to quantitatively study the welfare effects of social security privatisation. This paper extends this literature by showing that the Arrow (1987) mechanism delivers a two-agent structure in an otherwise standard New Keynesian model, and that fiscal policy in such a model is less powerful than in an equivalent two-agent model with exogenous wealth heterogeneity. Empirical support for this link between attention and wealth has recently been provided by Fagereng et al (2019) and Briggs et al (2020).

A closely related paper is Best and Dogra (2018). To my knowledge, theirs is the only other paper in which endogenous returns heterogeneity leads to ex ante identical households being divided into wealthy and poor in steady state. Their focus, however, is on the stability of different equilibria, and so they build a Ramsey model which is sufficiently tractable to solve analytically. Since my focus is on policy, I extend their work by showing that a two-agent steady state exists in an otherwise standard New Keynesian model, and I explore how the implied wealth inequality responds to and affects the transmission of fiscal policies.

In modelling the Arrow (1987) mechanism with households choosing how much costly information to process about saving products, I also extend the literature on rational

[2]Bilbiie (2019b) shows that this amplification is conditional on countercyclical inequality in an analytic-HANK model which, like this paper, features two types of household. However unlike the models labelled as TANK here Bilbiie's analytic HANK model features idiosyncratic shocks causing households to change types, bringing in the precautionary saving motives present in richer HANK models.

[3]See chapter 1 for data on this, and the theoretical implications of attention responding to the level of interest rates. In that paper I focus on the attention of a representative household, not on the two-agent steady state of this kind of model that I analyse here.

inattention in macroeconomics. Beginning with Sims (2003), several papers have explored how costly attention affects macroeconomic models when agents process information about normally distributed random variables (e.g. Adam (2009), Maćkowiak and Wiederholt (2009, 2015)). While this work has been instructive and valuable, Sims (2006) notes that rational inattention has potential applications in many situations in which the assumption of normality is not appropriate. This paper develops one such application, using the discrete choice framework of Matějka and McKay (2015) to embed rational inattention to savings product choices into a New Keynesian model. This discrete choice framework has previously been applied to importing (Dasgupta and Mondria, 2018) and hiring decisions (Acharya and Wee, 2019)[4].

Finally, it is often argued that fiscal policy must have a greater role in stabilising shocks as the natural rate of interest falls and policy rates become more likely to hit their lower bound (e.g. Blanchard and Summers, 2019). My model highlights a potential limiting force on the power of fiscal policy in such an environment if low real interest rates are associated with low interest rate heterogeneity. Low rate heterogeneity implies that a given fiscal policy will raise poor households further away from their borrowing constraint, which means they smooth the gains from the expansion more over time, reducing the impact of the policy. Fiscal policy will continue to have an effect at the zero lower bound, and as long as there is some interest rate dispersion[5] this effect will be stronger than in a representative agent model, but the model does suggest that fiscal policy may not be as powerful in a low interest rate environment as implied by models with exogenous wealth heterogeneity.

I set out the attention trap model in section 2.2. I show how this endogenously creates a two-agent model in section 2.3, then compare the effects of fiscal policy in this model with a model with heterogeneous discount factors in section 2.4. I show that lower interest rate dispersion reduces the power of fiscal policy in section 2.5. Section 2.6 concludes.

2.2 Model

The model is a standard New Keynesian model in many respects, except that there are two savings products available to households each period, with different interest rates. Households must process costly information to discover which offers the better deal. In this way attention (i.e. information processing) will lead households to achieve higher

[4]Lei (2019) has investors processing costly information about the mean return of risky investments, but in contrast to my work attention has no impact on the interest rate investors earn.

[5]Writing for the Financial Times, Kelly (2019) suggests that there is still some dispersion even with current rates at extremely low levels, showing that one bank is offering a deposit account identical to a competitor, but with a considerably lower interest rate.

interest rates on their assets, as hypothesised by Arrow (1987).

I restrict households to be net savers, so they always prefer higher (rather than lower) interest rates. This simplifies the household information problem, but is not necessary for the results. I discuss this further in section 2.3.

2.2.1 Savings Products

Each period, many financial institutions buy one-period government bonds at a price $(1+i_t)^{-1}$, then sell them to households at marginal cost. Every period, the institutions draw their marginal costs from an i.i.d. distribution[6]. With probability g, they have no marginal cost beyond the price paid to the government. I shall refer to these as the 'good' savings products. With probability $1-g$, the financial institution faces an extra marginal cost of $\chi > 0$. The 'bad' products from these institutions are therefore sold to individuals at price:

$$\frac{1}{1+i_t^b} = \frac{1}{1+i_t} + \chi \tag{2.1}$$

The 'bad' products are sold at a higher price, which implies a lower interest rate ($i_t^b < i_t$). Households will face a no-borrowing constraint, and government bonds will be in positive net supply, so all households in the model are net savers. They therefore strictly prefer the higher interest rates of the 'good' products to the 'bad'.

2.2.2 Households

Households are composed of many individuals. The household decides the consumption and labour supply of each individual, and how much attention each individual will pay to the choice of savings products, subject to a no-borrowing constraint. More attention will imply that the household experiences a higher interest rate.

The household sets the quantity of bonds b_t that each individual will buy before knowing which of the two available bond prices they face, and there is perfect consumption sharing around the household. Each individual chooses one savings product[7]. They choose a 'good' product with probability q_t, which is a function of information gathered

[6]Assuming i.i.d. costs simplifies the rational inattention problem facing households, as it means that information processed about dealers in one period has no value in future periods. This ensures that the model is tractable enough to be used for policy analysis. See Matějka, Steiner, and Stewart (2017) for results on information processing problems in dynamic settings.

[7]Redistributing consumption around the household means that individuals are risk neutral with respect to the bond price, seeking only to minimise their expected bond price. After processing information, they have a posterior belief about the probabilities of dealers being good or bad. Their risk-neutrality means there is nothing to gain from spreading their savings across multiple banks, so they choose a single bank for their portion of the household's saving.

$\mathcal{I}$ (which I refer to as attention, as is common in the rational inattention literature). There is a constant marginal cost μ to increasing $\mathcal{I}$.

These assumptions imply that the household faces the average, or 'effective', bond price:

$$\frac{1}{1+i_t^e} = \frac{q_t}{1+i_t} + \frac{1-q_t}{1+i_t^b} \tag{2.2}$$

Each individual faces a discrete choice rational inattention problem[8]. As they always prefer the 'good' savings product, they seek to maximise the success probability q. Using the entropy-based measure of information processing common in the rational inattention literature, we therefore have[9] for $q \in [g, 1)$:

$$\mathcal{I}_t = q_t \log\left(\frac{q_t}{g}\right) + (1-q_t)\log\left(\frac{1-q_t}{1-g}\right) \tag{2.3}$$

With no information processing, the individuals choose savings products at random and so choose a 'good' product with probability $q = g$. As attention increases, the individual processes more information before deciding on a product, and so chooses successfully with a greater probability. The probability of choosing a good product is concave in information processing, so there are diminishing returns to attention[10].

The household problem is:

$$\max_{c_t, n_t, b_t, i_t^e} \sum_t \beta^t \left(\frac{c_t^{1-\gamma}}{1-\gamma} - \frac{n_t^{1+\nu}}{1+\nu} - \mu \mathcal{I}(i_t^e) \right) \tag{2.4}$$

subject to

$$c_t + \frac{b_t}{1+i_t^e} = \frac{b_{t-1}}{\Pi_t} + w_t n_t + \mathcal{D}_t - T_t \tag{2.5}$$

$$b_t \geq 0 \tag{2.6}$$

Equation 2.5 is the budget constraint in real terms. b_t denotes real bond purchases in period t, Π_t is gross inflation and w_t is the real wage. I have written the effective interest rate i_t^e as a choice variable, which reflects attention choice. The effective interest rate is an increasing function of the probability of successfully choosing a good savings product q_t, and information processing $\mathcal{I}$ is an increasing function of q_t. Therefore $\mathcal{I}'(i_t^e) > 0$.

[8] These are studied in detail in Matějka and McKay (2015). In this simple case with only two possible outcomes we do not need their results linking this problem to the multinomial logit model.

[9] This function is derived in appendix B.1.1, where it is shown that the number of products available makes no difference to the relationship between information processing and the success probability q. The proportion of products which are 'good' (g) is all that matters.

[10] The attention required to choose a good product with a probability arbitrarily close to 1 is finite (proof in appendix B.1.1). No individual will approach that point as $\frac{dq_t}{d\mathcal{I}_t}$ approaches zero as q tends to 1, and so the returns to attention tend to zero.

Diminishing returns to attention ensure that $\mathcal{I}''(i_t^e) > 0$.

The first order conditions comprise an Euler equation, a labour supply condition, and a first order condition on attention:

$$c_t^{-\gamma} = \beta(1 + i_t^e)\mathbb{E}_t\left(\frac{c_{t+1}^{-\gamma}}{\Pi_{t+1}}\right) + (1 + i_t^e)\theta_t \tag{2.7}$$

$$c_t^{\gamma} n_t^{\nu} = w_t \tag{2.8}$$

$$b_t c_t^{-\gamma} = \mu(1 + i_t^e)^2 \mathcal{I}'(i_t^e) \tag{2.9}$$

θ_t is the Lagrange multiplier on the borrowing constraint, which is positive when the constraint binds and zero otherwise. The first order condition on attention (equation 2.9) implies that a wealthier household will choose to process more information, and so will experience a higher interest rate. This encourages further saving through the Euler equation (equation 2.7), but the non-concavity this implies is small enough at plausible parameter values that the first order conditions remain sufficient for utility maximisation (proof in appendix B.1.2). This interaction between attention and wealth is the reason this model has a steady state with two types of households, some wealthy and others at the borrowing constraint, despite those households having identical preferences and access to asset markets. In the remaining sections I denote a variable pertaining to wealthy or poor households only with a superscript w or p respectively.

2.2.3 Firms

There is a perfectly competitive final goods producer. They combine varieties of intermediate goods $Y_t(i)$ to make the final good using a standard CES aggregator:

$$Y_t = \left(\int_0^1 Y_t(i)^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{\epsilon}{\epsilon-1}} \tag{2.10}$$

Cost minimisation gives the demand curve facing an intermediate goods producer i:

$$Y_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\epsilon} Y_t \tag{2.11}$$

The price index P_t is also the final good producer's marginal cost, and so the price offered to consumers:

$$P_t = \left(\int_0^1 P_t(i)^{1-\epsilon}\right)^{\frac{1}{1-\epsilon}} \tag{2.12}$$

Intermediate goods firms are monopolistically competitive, face Rotemberg (1982) quadratic adjustment costs in prices, and have the production function:

$$Y_t(i) = A_t N_t(i)^{1-\alpha} \tag{2.13}$$

Here $N_t(i)$ is labour employed by the intermediate firms. Labour supplied from one household is a perfect substitute for labour from any other household.

The firm chooses their price to maximise the expected discounted sum of profits, subject to this production function and demand:

$$\max_{P_t(i)} \mathbb{E}_0 \sum_{t=0}^{\infty} \Lambda_{0,t} \left[P_t(i) Y_t(i) - W_t N_t(i) - \frac{\psi}{2} \left(\frac{P_t(i)}{P_{t-1}(i)} - 1 \right)^2 P_t Y_t \right] \tag{2.14}$$

subject to

$$Y_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\epsilon} Y_t \tag{2.15}$$

$$Y_t(i) = A_t N_t(i)^{1-\alpha} \tag{2.16}$$

There is some difficulty in heterogeneous agent models in deciding who owns the firms, which is important for the stochastic discount factor $\Lambda_{s,t}$. I assume that firm ownership is equally distributed among households, but they contract out firm operation to an agent with a stochastic discount factor of $\Lambda_{t,t+1} = \frac{\beta}{\Pi_{t+1}} \left(\frac{\bar{c}_t}{\bar{c}_{t+1}} \right)^{\gamma}$, where $\bar{c}$ is aggregate consumption across the two households[11]. All profits are distributed back to households equally.

I obtain the Phillips Curve from the first order condition for an intermediate goods producer, substituting in that all firms here are identical and so set the same prices and produce the same quantity of output:

$$Y_t(1-\epsilon) + \frac{\epsilon}{1-\alpha} w_t \left(\frac{Y_t}{A_t} \right)^{\frac{1}{1-\alpha}} - \psi(\Pi_t - 1)\Pi_t Y_t + \psi\beta \mathbb{E}_t (\Pi_{t+1} - 1)\Pi_{t+1} Y_{t+1} \left(\frac{\bar{c}_t}{\bar{c}_{t+1}} \right)^{\gamma} = 0 \tag{2.17}$$

The production function of intermediate producers gives us an equation for aggregate production, since all firms make identical choices. We have:

$$Y_t = A_t (\omega n_t^w + (1-\omega) n_t^p)^{1-\alpha} \tag{2.18}$$

[11] This is similar to the assumption of endogenous discount factors without internalization in Schmitt-Grohe and Uribe (2003).

2.2.4 Fiscal and Monetary Policy

Government issues a positive amount of short term bonds each period and funds its interest expenditures with lump sum taxes. This could be (for example) because the government wants to supply some bonds to ensure that households have access to a risk-free asset if they want it. This assumption ensures that the wealthy household is a net saver, so they always prefer higher interest rates. This keeps the information problem simple, but it is not required for the results. The qualitative results are the same if we assume that government debt has zero supply but households are allowed to borrow up to some positive limit. This is discussed in section 2.3.

The government budget constraint is:

$$\mathcal{B}_{t-1} = \frac{\mathcal{B}_t}{1+i_t} + P_t(\omega T_t^w + (1-\omega)T_t^p) \tag{2.19}$$

Here $\mathcal{B}_t$ is the supply of bonds issued in period t. Bond market clearing implies:

$$\mathcal{B}_t = \omega B_t^w + (1-\omega)B_t^p \tag{2.20}$$

The government is subject to the interest rate set by the central bank, which is that offered by the good savings product, not the average interest rate received by households. This is because 'good' financial institutions have no transaction costs, buy bonds from the government and then sell to individuals at marginal cost.

In real terms, the government budget constraint is:

$$\frac{\bar{b}_{t-1}}{\Pi_t} = \frac{\bar{b}_t}{1+i_t} + (\omega T_t^w + (1-\omega)T_t^p) \tag{2.21}$$

The lump sum taxes on the two types of households are always equal, except in the event of a 'Robin Hood' i.i.d. redistributive tax shock z_t:

$$T_t^w = T_t^p + z_t \tag{2.22}$$

Monetary policy is conducted according to a Taylor rule subject to a monetary policy shock M_t:

$$1 + i_t = \bar{i} + \left(\frac{\Pi_t}{\bar{\Pi}}\right)^{\delta_\Pi} \left(\frac{Y_t}{\bar{Y}}\right)^{\delta_Y} + M_t \tag{2.23}$$

2.2.5 Market Clearing

All output Y_t must be consumed in period t by households, be used to pay price adjustment costs in intermediate firms, or be lost to transaction costs in bad financial

institutions. The market clearing condition is therefore:

$$Y_t = \omega c_t^w + (1-\omega)c_t^p + \frac{\psi}{2}(\Pi_t - 1)^2 Y_t + \left(\frac{1}{1+i_t^b} - \frac{1}{1+i_t}\right)\left(\omega(1-q_t^w)b_t^w + (1-\omega)(1-q_t^p)b_t^p\right) \quad (2.24)$$

This completes the model. In all simulations in sections 2.4, 2.5 and the appendices the model is solved nonlinearly using the perfect-foresight LMMCP solver for Dynare. A shock is realised in period 1, and from that point on all households and firms know the future transition paths of the economy.

2.3 Steady States

There are two steady states of this model. First, there is a standard representative agent steady state. In this section I show, however, that there is also a steady state with households separated into two types, wealthy (w) households on their Euler equation and poor (p) agents up against their borrowing constraint. This is the same form as the two-agent models in BMP, Debortoli and Galí (2018), Bilbiie (2008) and others, even though (unlike those existing models) the households have identical preferences and access to asset markets. Poor households do not converge to the positive wealth level of the rich households because their lack of wealth means they do not have an incentive to pay attention to choosing good savings products, and so they face low interest rates, which encourages them not to save. The less attention wealthy households pay to choosing good savings products in steady state, the closer poor households are to lifting off their borrowing constraint. The representative agent steady state is unstable with respect to idiosyncratic shocks. It is the two-agent steady state which I will use to analyse the effects of fiscal policy.

2.3.1 Existence of two steady states

If there was no information problem, or if attention was costless, all households would earn the same interest rate on any savings. The Euler equations for the wealthy and poor households respectively in steady state in this scenario are:

$$1 + \bar{i} = \frac{1}{\beta} \quad (2.25)$$

$$1 + \bar{i} = \frac{1}{\beta + \theta \bar{c}^{p\gamma}} \quad (2.26)$$

These two can only both be true if $\theta = 0$, which implies that the poor household is not against the borrowing constraint, and therefore that the households are identical. Existing two agent New Keynesian models obtain steady state inequality either by assuming that the Euler equation doesn't hold for poor households because they are barred from asset markets (e.g. Debortoli and Galí, 2018) or by assuming that the two households have different β parameters (e.g. BMP, Monacelli and Perotti (2011)). In this model, it is instead attention that maintains wealth inequality in steady state.

With costly attention, the steady state Euler equations become:

$$\frac{1}{1+\bar{i}^{e,w}} = \frac{\bar{q}^w}{1+\bar{i}} + \frac{1-\bar{q}^w}{1+\bar{i}^b} = \beta \tag{2.27}$$

$$\frac{1}{1+\bar{i}^{e,p}} = \frac{\bar{q}^p}{1+\bar{i}} + \frac{1-\bar{q}^p}{1+\bar{i}^b} = \beta + \theta \bar{c}^{p^\gamma} \tag{2.28}$$

Substituting out for the bad interest rate in terms of the good (using equation 2.1) and combining we obtain:

$$\chi(\bar{q}^w - \bar{q}^p) = \theta \bar{c}^{p^\gamma} \tag{2.29}$$

If wealthy households pay more attention to choosing between bond dealers than poor households (i.e. if $\bar{q}^w > \bar{q}^p$) then $\theta > 0$ in steady state, which implies that the poor household is at their borrowing constraint. Steady state attention is given by the first order condition on interest rates (equation 2.9), remembering that for given interest rates on good and bad savings products (given i and i^b) an increase in the effective interest rate i^e must come from an increase in the success probability q:

$$\bar{b}^h \bar{c}^{h^{-\gamma}} = \mu(1+\bar{i}^{e,h})^2 \mathcal{I}'(\bar{i}^{e,h}), \quad h \in \{w, p\} \tag{2.30}$$

The marginal increase in information processing from an increase in the effective interest rate, $\mathcal{I}'(\bar{i}^{e,h})$, is positive and convex. Furthermore, the marginal information is zero when no information is processed, i.e. when the household members choose randomly from among the available bond dealers:

$$\mathcal{I}'(i^e) = 0 \iff \mathcal{I}(i^e) = 0 \iff q = g \tag{2.31}$$

A poor household at the borrowing constraint has zero assets, so $\bar{b}^p = 0$, which in equation 2.30 implies they optimally pay no attention to their choice of bond dealer. As household wealth $\bar{b}^h$ rises the left hand side of equation 2.30 rises, and so therefore does attention[12].

[12]This is true despite the fact that rising wealth leads to an increase in consumption and so to a fall in the marginal utility of consumption $\bar{c}^{h^{-\gamma}}$. This fall is smaller than the rise in wealth for plausible calibrations. The marginal utility of consumption appears because that gives the value of extra interest income in future periods.

This is because the costs of information do not depend on wealth, but the benefits of being more informed are higher interest rates on the stock of wealth saved by the household - so if more is saved, the benefits of information are larger. This mechanism was first suggested as an explanation of large persistent inequality by Arrow (1987), and it has been studied by many others since[13].

There is therefore a steady state in which wealthy households pay more attention to choosing between bond dealers than poor households who are against their borrowing constraint. Poor households have no 'skin in the game', and so pay no attention - which means they face lower interest rates on average, which reduces their incentive to save. This is the 'Attention Trap', and it gives a two-agent steady state without the assumptions made by others in the TANK literature that hand-to-mouth households are against the borrowing constraint because they are excluded from asset markets, or they are less patient than wealthy households.

A further implication of equation 2.29 is that if wealthy households pay less attention to their savings product choices (for example, if the cost μ rises), or if χ falls, then $\theta\bar{c}^{p^\gamma}$ must fall. θ is the marginal utility of a loosening of the borrowing constraint to a poor household, so a fall in this implies that poor households must be closer to lifting off that constraint, which will be important in section 2.5. As long as wealthy households process some positive amount of information (as long as μ remains finite and consumption is positive), and there remains some interest rate heterogeneity among bond dealers ($\chi > 0$), then the poor households will remain at the borrowing constraint in steady state ($\theta > 0$).

Equation 2.29 also highlights that the two-agent steady state does not rely on the borrowing constraint being precisely at zero, or on there being a positive supply of government bonds. The key requirement is that the probability of choosing low-price bonds is higher for wealthy households than for poor households against their borrowing constraint. If we instead built the model with a zero net supply of bonds and a negative minimum asset limit (i.e. we allowed some borrowing) we would find the same results. Poor households would borrow by selling bonds to bond dealers, and wealthy households would save by buying them. Borrowers would process information to try and find the dealers offering the highest price (the lowest interest rate), while wealthy households would aim for a low bond price as in the model above. Just as in this model, poor households would therefore face lower interest rates than wealthy households, which would sustain the wealth inequality.

[13]Recently, Fagereng et al (2019) have found evidence that wealthier households do earn greater interest rates, even within an asset class. Briggs et al (2020) find that after winning lottery money, Swedish households are more likely to participate in the stock market. The magnitudes suggest that there is a cost to investing in the stock market in excess of the monetary cost, and that wealthier households are more willing to pay that cost. The cost of paying attention is an obvious candidate.

The effects of policy would also be qualitatively the same with these alternative assumptions. As poor households lift off their borrowing constraint they have less skin in the borrowing game, so they pay less attention to finding the highest possible bond price, so the interest rate they face rises, just as it does in the model with a positive supply of government debt and no borrowing.

2.3.2 Instability with respect to idiosyncratic shocks

Although there are two steady states of this model, the representative agent steady state is unstable. An arbitrarily small shock to the incomes of some, but not all, households is sufficient to push the economy permanently away from the representative agent steady state. This highlights a conceptual difference between this model and existing TANK models. In such models, the identity of the poor households is pre-ordained: they are the households with less patience (BMP) or who are exogenously barred from asset market participation (Debortoli and Galí 2018). Here, which households end up at the borrowing constraint, achieving a lower level of lifetime utility[14], is random. The households are all identical, but when some households are hit with a small negative shock the attention trap leads to them becoming permanently poorer.

I prove that the representative agent steady state is unstable with respect to idiosyncratic shocks by contradiction. I first assume that after some idiosyncratic shock the poorer households expect to return to the representative agent steady state in the long run (i.e. I assume the steady state is stable). I then show that with those expectations returning to the representative agent steady state is inconsistent with equilibrium.

The first step of this argument consists of proposition 1.

Proposition 1 *Assume that the income of household 1 is strictly less than the income of household 2 in period t, and that both households expect their incomes to converge in the long run. If there are no further shocks, the income difference between the households will be larger in period $t+1$ than in period t if:*

$$\frac{1+i_t^{e,p}}{1+i_t^{e,w}} > \left(\frac{(\gamma-1)\nu}{\gamma(1+\nu)}\right)^{\frac{\gamma\nu}{\gamma+\nu}} {w_t}^{-\frac{\gamma(1+\nu)}{\gamma+\nu}} {c_t^w}^{\gamma} \tag{2.32}$$

Proof. See appendix B.1.3. In that appendix I also show that condition 2.32 will be satisfied for plausible calibrations. In particular, note that the condition is always satisfied

[14]Constrained households have a lower consumption in steady state than wealthy households in this model. If this is not the case (as in Debortoli and Galí (2018)) constrained households have a lower level of utility on average as long as asset supply is away from zero in at least one period. In that case constrained households have lower utility because they are unable to smooth consumption intertemporally, and their behaviour is different to those who can.

in the case of log utility ($\gamma = 1$). ■

Given this, we have that the representative agent steady state is unstable with respect to small idiosyncratic shocks that reduce the income of some, but not all, households. If the economy is to return to the representative agent steady state after such a shock, the gap between the incomes of the households must at some point begin to shrink. If households expect this, proposition 1 tells us that the income differential can never shrink, and so there is no equilibrium path back to the representative agent steady state that is consistent with rational expectations.

To see the intuition for this, consider a shock which makes the taxes on one household greater than taxes on the other household by a small amount in period 1. This tax separation is unanticipated, and exists for period 1 only. The standard effect of such a tax change is that the household that is hit with the higher taxes reduces consumption and saving, and increases their labour supply. The low tax household does the opposite.

The novel effect in this model is that the high tax household, who will from now on be referred to as the poor household, also chooses to reallocate their attention. With lower savings, they have less to gain from choosing the good savings product and getting a higher interest rate, so they process less information before making their choice. They therefore see a lower effective interest rate. This encourages an increase in consumption which outweighs the negative wealth effect of the tax change. This implies an even greater fall in savings. Conversely, the rich (low tax) household increases savings and attention, and decreases consumption.

This process continues, with the savings and attention levels of the two households diverging even though taxes are aligned after period 1. From period 2 onwards, there are no differences between households except their wealth, and that is sufficient to cause further divergence in wealth. The poor household's consumption starts to fall after a while because they are losing interest income as their wealth declines. Similarly, the wealthy household increases consumption after a while. The new steady state therefore has the wealthy household consuming more, and working less, than the poor household, because of their extra asset income.

There is nothing distinguishing the households from each other in the representative agent steady state. It is a random idiosyncratic shock that makes some households poorer than others, and from that point on the feedback between attention and wealth ensures that those households stay poorer permanently. This feedback is the 'Attention Trap'. Graphs of a simulation showing the transition from the representative agent to the two-agent steady state after such a small idiosyncratic shock are provided in appendix B.2.

2.4 Fiscal Policy

In this section I show that the model where wealth inequality is driven by the attention trap implies that large fiscal shocks - both intratemporal redistribution and debt-funded tax cuts - have smaller and more persistent effects on aggregate consumption than is the case in a model with no inattention where inequality is due to heterogeneous discount factors. This is because large tax cuts cause poor households to temporarily lift off the borrowing constraint. This happens in both the attention trap and heterogeneous discount factor models, but in the attention trap model those poor households also start to pay more attention to their choice of savings product. They therefore earn higher interest rates, and so they save even more of the windfall from the tax cuts, reducing the impact of the policy on aggregate consumption. This also prolongs the period over which the poor households spend the windfall and return to the borrowing constraint, which makes the effects of the policy more persistent.

If we consider fiscal expansions which are sufficiently small that poor households remain at the borrowing constraint, as BMP do, then the only difference between the attention trap model and the heterogeneous discount factor model comes from the way the attention of wealthy households varies with policy. As this is less clear than the attention response of poor households[15] I abstract from this issue by considering the case of sufficiently low information costs that wealthy households pay enough attention to choosing bond dealers that q^w is smaller than 1 by a negligible amount in all periods[16]. This implies that for small fiscal expansions the attention trap and heterogeneous β models are identical. They are not the same, however, for larger fiscal expansions that lift poor households temporarily off their borrowing constraints.

Studying larger fiscal stimuli is useful for two reasons. First, large fiscal stimuli are not theoretical curiosities. The ratio of U.S. government debt to GDP increased by an average of 6.2 percentage points each year between 2009 and 2012, largely due to the stimulus in the American Recovery and Reinvestment Act passed to mitigate the effects of the financial crisis (Dupor, 2017), and the U.S. stimulus package passed in March 2020 to combat the economic effects of the Covid-19 pandemic amounted to 10% of U.S. GDP. The debt-funded tax cut I consider raises the debt-to-GDP ratio by 3.83 percentage points[17] Such large expansions may well be seen again in the event of another crisis,

[15]The difficulty arises because for wealthy households, any small changes in wealth after a fiscal policy shock are outweighed in the attention decision by changes in the interest rate distribution for most calibrations of the model. In this simplified model, the dispersion of interest rates mechanically rises with the policy rate. However, in chapter 1 I show that empirically attention is negatively correlated with interest rates. I therefore leave discussion of that result to that paper.

[16]In the calibration used for simulations q^w remains within 10^{-40} of 1 in all periods of the simulations.

[17]When I consider a larger debt-funded tax cut at the end of section 2.4.2, the debt-to-GDP ratio rises by 8.66 percentage points. This is still not far off the 8 percentage points a year increase that was the

since with interest rates still close to zero monetary policy may have limited ability to stabilise demand. Second, this model is extremely simplified, and in reality there is a rich distribution of wealth. Even small policy changes will cause some households to lift off (or hit) borrowing constraints, and the exercises in this section show that the behaviour of these households has a substantial impact on the effects of fiscal policies.

Of course, the exercise I study here is an expansionary fiscal policy shock at the steady state, and in reality the two large fiscal expansions I have mentioned were intended to mitigate large recessions. This is not, however, inconsistent with the idea that the fiscal expansion will lift some people off their borrowing constraint, which is key in differentiating the attention trap model from the heterogeneous β model. This is because the large fiscal expansions mentioned were not targeted exclusively at those at borrowing constraints before the recessions began: some of those receiving the stimulus money had only just been pushed to their constraints by the recession, so it is plausible that the stimulus relaxed their constraints.

2.4.1 The Comparison Model

In the sections studying fiscal policy below I compare the effect of policies on the TANK model from section 2.2 with a two-agent New Keynesian model in which there is no attention problem, and steady state wealth inequality is driven by discount rate heterogeneity, as in BMP.

The model is identical to that in section 2.2, except that the cost of information $\mu = 0$, and a proportion $1-\omega$ of households have a lower discount factor than other households, $\beta^p < \beta$. The Euler equations of the two household types in steady state together imply:

$$\beta = \beta^p + \theta \bar{c}^{p^\gamma} \tag{2.33}$$

Since poor households have a lower discount factor, in steady state it must be that they are at their borrowing constraint ($\theta > 0$). At any given interest rate, poor households save less than wealthy (high β) households. The only steady state therefore has a two-agent structure similar to that in the model in section 2.2. The key difference is that the force maintaining wealth inequality here comes from preferences, which are exogenous to fiscal policy, unlike in the attention trap model, where that driving force comes from endogenous attention choices. Whereas who ends up at the borrowing constraint is random in the attention trap model, depending only on initial endowments, here the poor households are poor because they are less patient than others. The full model equations are given in appendix B.3.

US average from 2007-2009, and is below the size of the Covid-19 stimulus.

2.4.2 Debt funded tax cuts

Tax cuts to all households funded by increases in government debt offer a clear view of the differences between the attention trap and heterogeneous β models, as well as being an important part of the way the US government responded to both the Great Recession and the Covid-19 pandemic[18]. Such tax cuts are expansionary in both models, but if they are sufficiently large that poor households lift off their borrowing constraint the aggregate consumption response is smaller on impact and more persistent in the attention trap model. Similar results are obtained for tax cuts to poor households funded by contemporaneous tax hikes on wealthy households (see appendix B.5).

A small debt funded tax cut that does not lead to poor households temporarily leaving the borrowing constraint is expansionary in the period of impact for two reasons (discussed in detail in BMP). First, the tax windfall is all consumed by poor households, who have a MPC of 1. Second, in the future periods after the initial shock taxes will be high, and so poor household consumption will be below its steady state. This reduces inflation through the Phillips curve, and households therefore expect lower future inflation once they see that there has been a debt-funded tax cut. As a result, households expect lower future real interest rates, because they know the central bank will respond to the deflation by reducing nominal interest rates more than proportionally to the deflation (they satisfy the Taylor principle). This increases wealthy household consumption.

However, the elasticity of aggregate consumption to a debt-funded tax cut is smaller (in both models) for shocks large enough to lift poor households slightly off their borrowing constraint. Poor households spend less than 100% of their tax windfall, which also means wages don't rise by as much as if poor households remained hand-to-mouth in all periods. Furthermore, if the tax cut is large enough that poor households still hold some wealth when taxes first rise to pay off the debt, this supports aggregate consumption in that future period because poor household consumption won't fall as much when taxes rise: the MPC is less than 1 because they can adjust their saving. This mitigates the expected future fall in the real interest rate, further reducing the initial aggregate consumption response to the policy.

In addition to this effect, in the attention trap model poor households saving a small amount increase their attention to choosing between savings products (from zero). Increasing attention means that poor households face higher effective interest rates than if they kept attention constant. They therefore save even more, reducing their MPC out of the tax windfall and so reducing the aggregate consumption response to the policy below

[18] A lump sum tax cut in the model is a close approximation to the mailing of checks carried out in both circumstances, though the real policies did taper the amount of transfer at very high levels of income.

that in the heterogeneous β model. With less of an increase in consumption, wages also do not rise by as much, which further reduces the effect of the policy on consumption. Poor household wealth is more persistent, and so therefore is the consumption increase, which further reduces the initial consumption rise as it increases expected future real interest rates relative to the heterogeneous β model[19].

These effects are shown in figure 2.1, which shows the results of a debt-funded tax cut in the two models. Details of the calibration are given in appendix B.4: the two models share all common parameters (so have the same steady state), and the novel parameters in each model have been chosen such that the poor households are equally close to lifting off the borrowing constraint before the policy hits in each model[20]. The size of the policy shock is such that on impact the 35% of households who were at the borrowing constraint go from having no assets to holding approximately 1% of the stock of government debt in the attention trap model. The debt-to-GDP ratio rises by 3.83 percentage points to achieve this. Higher future taxes on all households cause government debt to return to steady state following an AR(1) process with persistence 0.5.

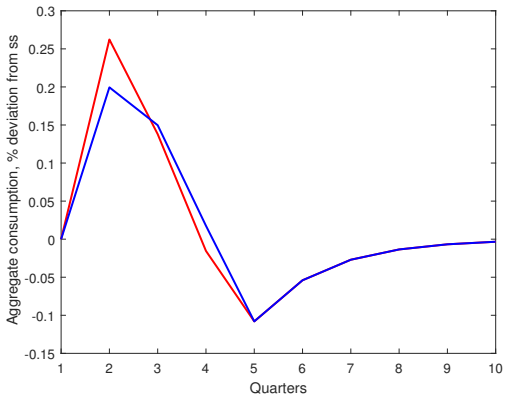
The response of aggregate consumption (figure 2.1a) is substantially lower on impact in the attention trap model (in blue) than in the heterogeneous β model (in red): with this calibration the rise in aggregate consumption on impact is 24% lower in the attention trap model. Aggregate consumption in the attention trap model is also more persistent, which means that the present value of the aggregate consumption deviations is only 11% lower in the attention trap model.

The mechanism can be seen in figure 2.1b, which shows the effective interest rates earned by wealthy (black dashed) and poor (blue solid) households in the attention trap model. As poor households lift off their borrowing constraint their increase in attention to choosing between savings products causes their effective interest rate to rise substantially relative to the rate experienced by the wealthy households, which is the policy rate. The gap between the effective interest rates of the two households widens as debt is brought back to steady state and poor households reduce their saving back to the borrowing constraint, causing them to reduce their attention. This would happen even without further tax changes through the logic of the attention trap, but it is accelerated here by the higher taxes used to repay the government debt, which reduce the income, and so saving, of all households.

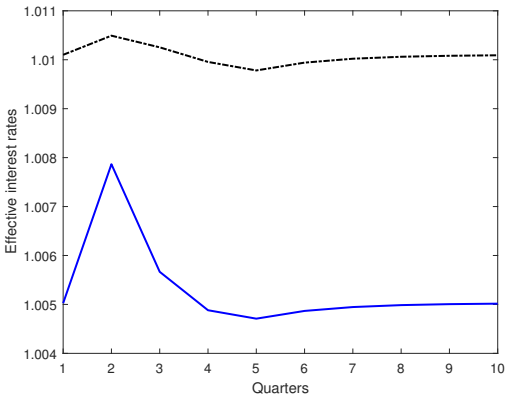
[19]In appendix B.1.4 I show analytically that these mechanisms imply that the consumption response to the policy in the attention trap model is smaller on impact than in the heterogeneous β model.

[20]That is θ^p, the shadow value of relaxing the borrowing constraint for poor households, is equal in steady state in the two models.

Figure 2.1: Effects of a debt-funded aggregate tax shock



(a) Aggregate consumption in the attention trap model (blue) and the heterogeneous β model (red).



(b) Effective interest rates facing poor (blue solid) and wealthy (black dashed) households in the attention trap model.

This effect is even more pronounced for bigger policy shocks. The larger the shock, the more saving the poor households do, and so the more attention they pay to choosing between bond dealers, which causes the two models to diverge further. After a larger debt-funded tax shock which causes the poor households to collectively amass 2.5% of the stock of government debt (a 8.7 percentage point increase in debt-to-GDP), aggregate consumption responds 52% less to the policy on impact in the attention trap model than in the heterogeneous β model, and the present value of consumption deviations is 25% lower in the attention trap model.

Even though the tax cut is distributed evenly across the population, the heterogeneity in effective interest rates falls when the shock hits. This is because wealthy households are already paying enough attention to identify good bond dealers with probability $q^w \approx 1$, so even as their wealth rises with the expansion of asset supply they cannot meaningfully increase their interest rate by more than the increase in the policy rate. In contrast, poor households start out paying no attention, and as they start to save they increase this, and so their interest rate rises even relative to the policy rate[21]. However, from the first order condition on attention (equation 2.9) we can see that the effective interest rates earned by the two households will only equalise if they save the same amount (in marginal utility units).

This kind of fiscal policy cannot therefore remove interest rate heterogeneity entirely. To absorb the bulk of the extra government debt, wealthy households reduce consumption,

[21]We would in fact see this effect even if wealthy households were not near full attention, as the diminishing marginal returns to attention ($\mathcal{I}''(i^e) > 0$) imply convex costs of effective interest rates. Wealthy households will therefore increase their effective rates by less after an aggregate expansionary shock than poor households will, as the marginal cost of doing so is less sharply increasing for poor households paying little attention. However, the effective interest rates of the two groups will never equalise exactly unless attention is the same, which is not possible with a uniform tax cut.

which along with their large rise in saving implies that the difference in saving in marginal utility units between households actually grows. Redistributive taxation, in contrast, can permanently remove interest rate and wealth heterogeneity, if the transfers are such that all households have the same income in one period. This, however, is a knife-edge result, as even if the incomes of the households differ by a very small amount the attention trap will kick in as discussed in section 2.3.2, taking the economy back to the two-agent steady state.

To assess the welfare consequences of this policy, note that along with the standard New Keynesian welfare losses from inflation and output gaps, there are two further factors affecting welfare in this model. Firstly, the diminishing marginal utility of consumption, and increasing marginal disutility of labour, imply that consumption and labour supply heterogeneity are detrimental to welfare (as shown for other TANK models in Bilbiie (2008) and Debortoli and Galí (2018)). The second effect is specific to the attention trap model: transaction costs paid by inefficient bond dealers are waste, so increasing the share of bonds sold through bad dealers is also detrimental to welfare.

These effects go in opposite directions when the tax shock hits. Since wealthy households receive asset income that poor households do not, in steady state they consume more and work less than poor households. The shock reduces this consumption and labour inequality because while poor households consume some of the tax windfall, wealthy households reduce consumption in response to the large rise in policy rates necessary to ensure households absorb the extra asset supply. However, the shock also increases the share of bonds purchased from inefficient dealers, because poor households increase their saving (from zero), and they pay less attention to their choice of bond dealer than wealthy households. Which effect dominates, and so whether the debt increase is positive or negative for welfare, depends on how strongly these two factors are weighted in the social welfare function, which will be a function of model parameters and normative decisions over how strongly to weight the utility of poor households in the social welfare function.

The welfare loss from inattentive savers also shows why the government may not wish to remove inequality by taxing asset income, the most direct way to remove the source of the heterogeneity. The only way to eliminate the inequality completely with this policy would be to tax away all interest income and redistribute it equally, but that would completely remove the incentive for wealthy households to pay attention to their choice of bond dealer, thereby generating welfare losses from inefficient saving.

2.5 Low Interest Rate Dispersion

As natural rates of interest have fallen in many developed economies (Holston, Laubach and Williams, 2017), several papers have suggested that fiscal policy should play a greater role in stabilising demand in recessions as monetary policy is increasingly restricted by an effective lower bound (Blanchard and Summers, 2019). Low interest rates, however, may well coincide with low interest rate dispersion, since banks are unwilling to pass on negative interest rates to households even when policy rates are negative (Eggertsson et al 2019)[22]. In this section I show that low interest rate dispersion reduces the power of both intratemporal redistributive tax shocks and debt-funded tax cuts to influence aggregate consumption, because when wealthy and poor households face interest rates which are closer together the force keeping poor households at the borrowing constraint is reduced. Smaller fiscal policy shocks are needed to raise poor households off the borrowing constraint, and a given shock increases poor household wealth and attention by a greater amount.

The stimulus required to lift poor households temporarily away from the borrowing constraint is heavily influenced by θ^p, the steady state shadow value of relaxing the borrowing constraint for poor households. It is given in equation 2.29:

$$\chi(\bar{q}^w - \bar{q}^p) = \theta^p \bar{c}^{p^\gamma}$$

In the attention trap model the dispersion of interest rates is set exogenously by χ. A lower dispersion in the model means a lower χ, and so a lower θ^p. That is, when interest rate dispersion is lower, the shadow value of relaxing the borrowing constraint on poor households falls. If the value of relaxing the borrowing constraint falls, this must mean that the savings poor households would have chosen without that constraint have risen closer to zero, as nothing else has changed in the model to alter that value[23].

Intuitively, this is because the force maintaining the wealth inequality in steady state is the difference in effective interest rates experienced by the two types of household. With a smaller spread of interest rates in the market, this difference in interest rates shrinks, and so therefore does the distance of poor households from being able to start saving.

This means that when interest rate dispersion is lower poor households lift off the

[22]Writing in the Financial Times in 2018 Barrett (2018) noted that "as interest rates continue to bump along the bottom, the difference between the best and the worst easy-access savings rates are not that great".

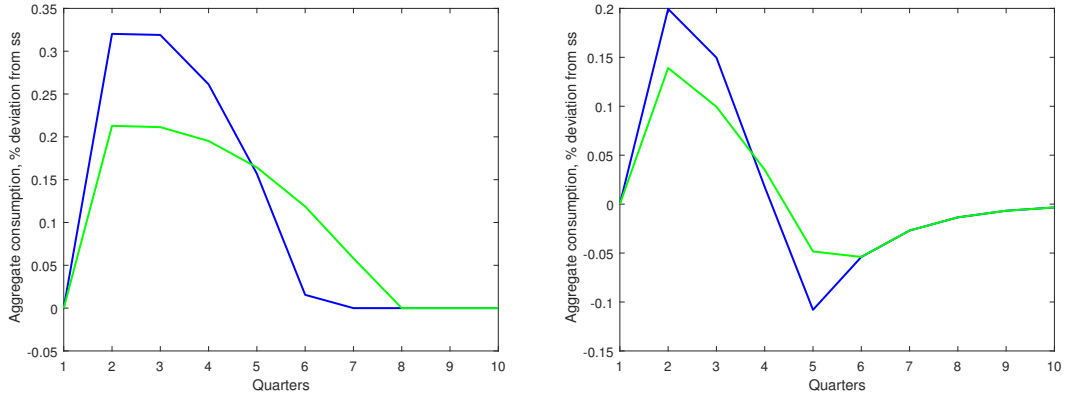
[23]This is true here because I study calibrations in which $q^w \approx 1$, so the steady state policy rate stays the same and no wealthy household is affected by the change in i^b, which means that in steady state nothing changes except θ^p.

borrowing constraint after smaller tax cuts, and for a given tax cut they increase their saving further above zero. This reduces the effect of the tax cut on aggregate consumption. It is not the case, however, that poor household attention rises by more in the low dispersion environment: the benefits of attention are lower when χ is low because good and bad interest rates are closer together. In this calibration this more than offsets the effect of the extra wealth on poor household attention.

The graphs below show the response of aggregate consumption to the uniform debt-funded tax cut studied in section 2.4, and to a one-period tax cut for poor households funded by a contemporaneous tax rise for wealthy households. Both shocks lead to poor households accumulating 1% of all government debt in the baseline (high χ) calibration. The blue lines show aggregate consumption with that baseline calibration, and the green lines show aggregate consumption with a low interest rate dispersion, in which χ has been reduced by half.

For both policies a lower interest rate dispersion implies a smaller and more persistent response of aggregate consumption. The change in aggregate consumption on impact is smaller with low rate dispersion for both policies, by 30% for the debt-funded tax cut and 34% for the redistribution. The expansion of aggregate consumption is also more persistent with low dispersion, so while the present value of consumption deviations is lower with low dispersion, the difference is not as great as the impact effect (22% lower for the debt-funded tax cut, 11% lower for the redistribution).

Figure 2.2: Aggregate consumption responses to expansionary fiscal policy vary with interest rate dispersion in the attention trap model



(a) Redistributive tax shock with high (blue) and low (green) interest rate dispersion.

(b) Debt-funded tax shock with high (blue) and low (green)interest rate dispersion.

The argument that fiscal policy should play a larger role in demand stabilisation in a world of low interest rates is therefore somewhat dependent on why we believe there is persistent wealth inequality. If it is because of differences in labour productivity, or preferences, which are exogenous to fiscal policy (as in the heterogeneous β model) then

fiscal policy can indeed be powerful in this environment. If inequality is in part driven by how much information households choose to gather about their savings choices, however, as Arrow (1987), Fagereng et al (2019), and others have suggested, then fiscal policy may also be weakened in a low interest rate environment if low interest rates coincide with a low interest rate dispersion. Of course, if the interest rate hits the zero lower bound then fiscal policy may still be needed to stimulate consumption, as monetary policy will be unable to help and fiscal policy will still have a small effect, as long as interest rate dispersion does not also hit zero.

2.6 Conclusion

Bilbiie, Monacelli and Perotti (2013) show that the effects of fiscal stimulus packages in a simple New Keynesian model with persistent wealth inequality are heavily dependent on the redistributive aspects of the package. In this paper I have shown that for fiscal stimuli large enough to raise poor households temporarily away from their borrowing constraints, the cause of steady state inequality also plays a large role in determining the response of consumption to the policy. In Bilbiie, Monacelli and Perotti, the steady state wealth inequality comes from differences in household preferences, which are exogenous to fiscal policy. I have proposed an otherwise similar model in which the wealth inequality is driven by an 'attention trap', an interaction between attention to choosing between savings products and wealth first proposed by Arrow (1987). Crucially, attention is not exogenous to fiscal policy.

A fiscal stimulus that raises poor households above their borrowing constraint encourages those households to pay more attention to their choice of savings product, which means they experience a rise in the effective interest rate they face relative to a model with fixed attention. They therefore choose to save more of their tax windfall, reducing their consumption and making the effects of the policy more persistent as poor households engage in more consumption smoothing.

Attention also depends on the interest rate environment. If banks do not pass on negative interest rates to households (Eggertsson et al, 2019), then the range of interest rates on offer to households will be lower in a world where average interest rates are low, as the zero lower bound prevents banks offering rates much below the average rate. With less interest rate dispersion, attention to choosing good investments makes less difference to the household's return, and so the force keeping poor households at their borrowing constraint is weaker. This further reduces the effect of fiscal stimulus on consumption, because poor households are lifted further from their borrowing constraints and so smooth consumption more than if interest rate dispersion was high.

Chapter 3

A Rational Inattention Unemployment Trap*

Abstract

We show that introducing rational inattention into a model with uninsurable unemployment risk can generate multiple steady states, when the model with full information has a unique steady state. The model features persistent, heterogeneous labour market expectations, consistent with survey evidence. In a heterogeneous agent New Keynesian model, rational inattention to the future hiring rate generates three steady states: an unemployment trap with mild deflation and a low (but positive) job hiring rate, a middle steady state with moderate employment and inflation, and an 'employment trap' with high employment and inflation. Large mutations in the distribution of household beliefs can shift the economy between steady states.

*This chapter is co-authored with Martin Ellison.

3.1 Introduction

There is a long history of papers suggesting that self-fulfilling expectations might allow an economy to become stuck in a bad steady state (see Diamond 1982 for an early example). One source of these fluctuations is the interaction of labour market expectations and precautionary saving. If households believe that their future employment prospects are bleak, they will increase precautionary savings today. The fall in aggregate demand that results causes employment to fall, confirming the pessimistic beliefs. This feedback loop is empirically important: Heathcote and Perri (2018) provide evidence that precautionary savings were a key driver of consumption around the onset of the Great Recession. Existing models of this mechanism have households precisely co-ordinating their labour market expectations on a particular equilibrium, as is common in models with multiple equilibria (Morris and Shin, 2000).

In this paper, we show that if it is costly for households to process information about future labour market conditions, their optimal information choices can generate multiple self-fulfilling steady states in a model which would have a unique steady state if households were fully informed. The unemployment trap generated in this way relies on the interaction of labour market expectations and precautionary saving, but it does not feature the strong co-ordination of household beliefs present in existing models of self-fulfilling labour market expectations.

There are two important assumptions that drive this result. Firstly, we assume that the hiring rate out of unemployment, which is crucial for precautionary saving in models with frictional labour markets (e.g. Ravn and Sterk, 2020), is not observed directly by households. Households can obtain signals about the hiring rate, but this information processing is costly. Following Sims (2003) and others in the rational inattention literature, the cost is proportional to the informativeness of the signals, and so households choose to process somewhat noisy signals before deciding on their consumption. The hiring rate is naturally bounded between 0 and 1, so the commonly used result that rationally inattentive agents choose linear signals with Gaussian noise does not apply. The bounded support of the hiring rate implies that households choose signals with a discrete number of possible realisations even though the underlying variable is continuous (Matějka 2017). This nonlinear signal structure is what drives the possibility of multiple steady states.

The information processing cost implies that households have limited information about *realisations* of the hiring rate. Our second key assumption is that households also have limited information about the *structure* of their environment: they do not precisely know the true equilibrium marginal distribution of the hiring rate. This is related to

the 'internal rationality' studied by Adam and Marcet (2011). To our knowledge, we are the first paper to examine this combination of rational inattention and imprecise prior beliefs. This is motivated by an observation from the Survey of Consumer Expectations: expectations have a much greater variance than the underlying hiring rate. This suggests that households do not have a good understanding of the range of values usually observed for the hiring rate.

The information choices made by households in this environment imply labour market expectations which are heterogeneous and persistent. There is also only weak co-ordination of beliefs: when the hiring rate falls, households on average receive signals that indicate it has fallen, and so average hiring rate expectations fall. These signals, however, are very noisy, so at the same time as the average expectation is falling a little, some households will be shifting their expectations up, and others will be expecting a large collapse in the hiring rate.

Survey labour market expectations also display these properties. In the Survey of Consumer Expectations, households are asked to predict the probability that, should they lose their main job today, they would find another suitable job within three months. This is precisely the hiring rate which drives precautionary saving in our model and in Ravn and Sterk (2020). Even after controlling for a wide range of personal characteristics, there is a great deal of disagreement about this rate each month, and expectations are highly persistent at the household level. We also find evidence that co-ordination of beliefs is weak, as in our model: if households all simultaneously agree on shifts in the hiring rate, then changes in average beliefs should account for a large amount of the variance of changes in household-level expectations over time. In fact, changes in average expectations account for just 0.1% of the variation in household belief changes. That is, most changes in household beliefs come from idiosyncratic factors. The noise in household signals in our model is a source of this kind of idiosyncratic variation.

The combination of information processing costs and imprecise prior beliefs is central to our results. This is because households facing noisy signals bias their expectations towards their prior expectation, as can be seen in an argument by contradiction. In the models we study, the unemployment trap is not a steady state if households have full information about the realisations of the hiring rate. Suppose, then, that an unemployment trap is a steady state under rational inattention, and households know the true equilibrium distribution of the hiring rate before they process any information about its particular realisation that period. Around the unemployment trap steady state, household prior expectations equal the steady state hiring rate associated with the trap. Households bias their beliefs towards this, and so precautionary saving behaviour is very close to the optimal action for that hiring rate under full information. That full information

precautionary saving behaviour at that low hiring rate did not produce a steady state, and therefore such a trap cannot be a steady state under rational inattention, *unless* households do not precisely know the true equilibrium distribution of the hiring rate, in which case their beliefs are not necessarily biased towards the full information action.

We find that multiplicity does not disappear over time even if households observe a long history of signals and update their prior beliefs. In order to update their beliefs about the distribution of the hiring rate, households must process new information, which comes with a cost. Households process information until the marginal benefit equals this marginal cost. If prior beliefs already contain some information from previously observed signals, the optimal amount of new information processing will be very small. This logic is explored in detail by Matějka, Steiner, and Stewart (2017). Crucially, this prevents beliefs about the distribution of the hiring rate collapsing to the truth, which means that multiple steady states survive even when beliefs can update over time[1]. We find the unemployment and employment trap steady states are locally stable, while the middle steady state is unstable, with respect to 'mutation' shocks which alter the distribution of prior beliefs in the population. Large mutations can cause the economy to transition between steady states.

Related Literature: This paper relates to several strands of existing literature. Firstly, there is a vast literature studying fluctuations and traps driven by self-fulfilling expectations (see Cooper and John (1988) for a review of the early literature). Specifically, in our model changes in labour market expectations affect precautionary saving decisions, which in turn affect aggregate demand and so the labour market. Challe and Ragot (2016) show that a tractable model featuring this mechanism fits US data on aggregate consumption significantly better than benchmark models, and Challe et al (2017) find that the feedback loop between unemployment risk and precautionary saving played a significant role in the Great Recession. Beaudry, Galizia, and Portier (2020) show that it can lead to a 'Hayekian' recession after an over-accumulation of durable goods. Closely related to our paper are Heathcote and Perri (2018) and Ravn and Sterk (2020), who show that self-fulfilling labour market expectations can lead to the existence of multiple steady states: an economy can get stuck in a bad steady state where pessimistic beliefs persist indefinitely. We contribute to this literature by showing that these unemployment traps can be generated purely through the optimal information choices of rationally inattentive households, even in a model which is linear under full information. Moreover, our model does not require the co-ordination of beliefs which Morris and Shin (2000) argue is often necessary in models with multiple equilibria.

[1]In Adam and Marcet (2011), beliefs are 'near-rational', that is they cannot be empirically distinguished from the truth. The difference to our paper is that for Adam and Marcet, the only constraint on learning is the availability of data. For us, learning requires information processing, which is costly.

We also contribute to the literature on rational inattention (RI). Most existing models with RI have agents with quadratic objective functions collecting costly information about a random variable with a known Gaussian distribution[2] (Sims, 2003, Maćkowiak and Wiederholt, 2009). This has proved useful in explaining price stickiness (Maćkowiak and Wiederholt, 2009, Matějka, 2015), consumption patterns (Sims, 2003, Luo, 2008, Luo et al, 2017), business cycle patterns (Maćkowiak and Wiederholt, 2015), and other macroeconomic phenomena. Matějka (2017) and Jung et al (2019) show, however, that assuming a bounded prior belief leads to very different results to the quadratic-Gaussian formulation. Specifically, they show that the optimal decision rule of an agent facing a rational inattention problem with a bounded prior entails the agent choosing to limit themselves to a discrete number of options, even when the optimal choice under perfect information is continuous. As the probability of finding a job is naturally bounded by 0 and 1, our model displays these features. This paper is therefore a response to the call in Sims (2006) to explore the implications of RI away from the quadratic-Gaussian case. To our knowledge we are the first to incorporate rationally inattentive households with bounded prior beliefs into a general equilibrium setting[3].

Our framework also relates closely to the literature on internal rationality. Adam and Marcet (2011) show that allowing for internally rational agents, who optimise given their beliefs but do not precisely know the equilibrium distributions of state variables, has important effects on asset pricing models. Adam, Kuang, and Marcet (2012) use this to explain movements in house prices and the current account. We extend this literature by showing that the interaction of internal rationality and information processing costs creates multiplicity, where neither assumption generates this by itself. We believe that we are the first to combine these two information restrictions. We deviate from existing literature on internal rationality in that we do not assume that agent beliefs are close to the true equilibrium distribution of the state variable. The typical logic for 'near-rationality' is that agents learn over time, so would discard any beliefs which are very dissimilar to the truth. This does not happen when households learn in our model, because to learn they must process costly information, which means households stop learning before they reach near-rational beliefs.

Finally, our work contributes to the literature on heterogeneous expectations in macroeconomics. Armantier et al (2015) and Meeks and Monti (2019) show that inflation expectations are heterogeneous across households; in section 3.4 we document that the same is

[2]This is convenient as the optimal posterior belief about the shock, after processing information, is also Gaussian.

[3]Stevens (2019) explores a model in which the opportunity for firms to pay a fixed cost to exactly reveal the state of the economy places a bound on the range of shock realisations under which a given signal structure is expected to be in operation, which leads to the same kind of discrete decision rules for her firms that we obtain for households.

true for hiring rate expectations. The theoretical implications of heterogeneity in inflation expectations have been studied by Andrade et al (2019), Wiederholt (2017), among others. In contrast to this literature, we study heterogeneous labour market expectations.

Section 3.2 illustrates the potential for an unemployment trap in a simple static model. We then introduce our mechanism into a version of the HANK model from Ravn and Sterk (2020), in which the future hiring rate is very important in household decisions, in section 3.3. We show that the combination of rational inattention and imprecise prior beliefs about the hiring rate generates three possible steady states: an unemployment trap with a low hiring rate and mild deflation, a middle steady state with higher employment and moderate inflation, and an 'employment trap' with very high employment and high inflation. In section 3.4 we show that several key features of our model are found in survey data on hiring rate expectations. Section 3.5 concludes.

3.2 Static Model

This section illustrates our mechanism in a simple static model. There are two subperiods in this setup, which we refer to as morning and afternoon. The consumer problem is related to Ravn and Sterk (2020) and to our HANK model in section 3.3, but is substantially simplified by the static nature of the problem. The firm side is also kept very simple, to illustrate the key forces driving the unemployment trap. In section 3.2.8 we construct a dynamic model by repeating this static model, allowing households to update their prior beliefs over time using information processed in previous periods. This demonstrates that the unemployment trap does not disappear even when households update their prior beliefs over time.

3.2.1 Households

There is a unit mass of households. All households are employed in the morning, and they receive a wage of 1. With probability ω a household loses their job at the end of the morning. There is then a round of hiring by firms, so newly separated workers find employment for the afternoon with probability η. If they are employed in the afternoon, the household receives a wage w. If they are unemployed they receive the value of home production $\theta < w$.

Households can save or borrow in the morning at interest rate R. The budget constraints for household i in the morning, afternoon if employed, and afternoon if unemployed (respectively) are:

$$c_{mi} + \frac{b_{mi}}{R} = 1 \tag{3.1}$$

$$c^e_{ai} = w + b_{mi} \tag{3.2}$$

$$c^u_{ai} = \theta + b_{mi} \tag{3.3}$$

Combining these we have:

$$c^e_{ai} = w + R(1 - c_{mi}) \tag{3.4}$$

$$c^u_{ai} = \theta + R(1 - c_{mi}) \tag{3.5}$$

Households have quadratic consumption utility: $U(c_{ti}) = -\frac{1}{2}(c - \bar{c})^2$.

The household problem is therefore to choose morning consumption c_{mi} to maximise:

$$V_{mi} = -\frac{1}{2}(c_{mi} - \bar{c})^2 - \frac{1}{2}\beta\mathbb{E}_{mi}\omega(1-\eta)(c^u_{ai} - \bar{c})^2 - \frac{1}{2}\beta\mathbb{E}_{mi}(1 - \omega(1-\eta))(c^e_{ai} - \bar{c})^2 \tag{3.6}$$

If the households observe the hiring rate η all households make the same choice of morning consumption c_{mi}, to satisfy the FOC (dropping the i subscripts):

$$c_m = \beta R\omega(1-\eta)\big(\theta + R(1 - c_m)\big) + \beta R\big(1 - \omega(1-\eta)\big)\big(w + R(1 - c_m)\big) \tag{3.7}$$

In this case, morning consumption is an increasing linear function of the afternoon hiring rate, due to a simple consumption smoothing motive[4]. Solving out for morning consumption we have:

$$c_m = \frac{\beta R}{1 + \beta R^2}\big(R + w - \omega(1-\eta)(w - \theta)\big) \tag{3.8}$$

3.2.2 Rational Inattention Problem

We now relax the assumption that households can precisely observe the afternoon hiring rate η. Instead, we assume that they can collect information about η from a variety of noisy signals, but doing so is costly. This cost is increasing in the informativeness of the signal chosen. This is formalised in equation 3.12 below. As well as the amount of information in the signal (which we will denote κ), the agent must also choose how this information is to be structured: they could choose a signal which is very accurate in some ranges of η but not in others, for example.

The payoff function being maximised is as in equation 3.6. The agent views the hiring rate as exogenous[5]. In maximising their expected utility the agent must decide on an

[4]When the hiring rate rises expected afternoon consumption rises as the probability of being employed rises, and so optimal morning consumption rises. Throughout the paper we refer to the saving households engage in to insure themselves against future unemployment as 'precautionary saving', but we do not mean by this that the certainty equivalent level of afternoon consumption is below the expected level of c_a, which would require a utility function with a positive third derivative. We simply mean that households choose to save in this model as a precaution against future unemployment.

[5]The hiring rate will in fact be endogenous to aggregate agent choices, but the agent does not take this into account. This will be explored in more detail in section 3.2.7. Notice that this assumption

optimal decision rule to map the signals they are able to process to consumption.

The solution to this problem therefore takes the form of an *information strategy* and an *action strategy*. The information strategy gives the amount of information the agent should process, and what form the signals should take. The action strategy maps from signal realisations to consumption choices.

This household problem is closely related to the firm profit maximisation problem in Matějka (2017). We therefore proceed by following the steps used to solve that firm problem. First, we simplify the household problem by noting that the value function in equation 3.6 is strictly concave in morning consumption c_{mi}, so there is a unique function mapping the optimal morning consumption choice c^*_{mi} to the expectation of the afternoon hiring rate formed by the household in the morning, $\mathbb{E}_{mi}\eta$. Furthermore, the optimal morning consumption choice is a continuous and strictly increasing function of the expected hiring rate, so there is a one-to-one mapping between the expected afternoon hiring rate and optimal morning consumption. A household will never choose a signal structure with two distinct possible realisations that imply the same expected hiring rate, because distinguishing between the two realisations is a waste of information processing. There will therefore be a one-to-one mapping between signal realisations and the optimal morning consumption choice. We can therefore leave the signal choice in the background of the problem, and study the optimal decision rule linking morning consumption to the hiring rate, subject to the information costs of implementing such a rule.

Specifically, we express the household's decision rule as $f_i(\eta, c_{mi})$, the joint probability density function over the hiring rate and morning consumption. That is, given a particular afternoon hiring rate η, the household chooses how often they will choose each different possible level of morning consumption. They are aware that signals contain noise, so they are deciding how often, and by how much, they are willing to choose the wrong c_{mi} for each level of η, given the information costs of reducing those mistakes.

The problem of household i is therefore:

$$f_i = \arg\max_{\hat{f}_i} \mathbb{E}_{mi}[V(\eta, c_{mi})] - \psi\kappa(\hat{f}_i, g_i) = \arg\max_{\hat{f}_i} \int_\eta \int_{c_{mi}} V(\eta, c_{mi})\hat{f}_i(\eta, c_{mi}) d\eta dc_{mi} - \psi\kappa(\hat{f}_i, g_i) \tag{3.9}$$

subject to

$$\int_{c_{mi}} \hat{f}_i(\eta, c_{mi}) dc_{mi} = g_i(\eta) \qquad \forall\eta \tag{3.10}$$

$$\hat{f}_i(\eta, c_{mi}) \geq 0 \qquad \forall\eta, c_{mi} \tag{3.11}$$

$$\kappa(\hat{f}_i, g_i) = H[g(\eta)] - \mathbb{E}_{c_{mi}} H[\hat{f}_i(\eta|c_{mi})] \tag{3.12}$$

means that higher order beliefs do not affect the household problem.

The function $H[.]$ is the entropy of the distribution over which it operates. That is:

$$H[g_i(\eta)] = -\int g_i(\eta) \log g_i(\eta) d\eta \tag{3.13}$$

The first constraint (equation 3.10) ensures that the marginal distribution of the hiring rate η obtained from the optimal joint pdf is consistent with $g_i(\eta)$, household i's prior belief about the distribution of the hiring rate [6].

The second constraint (equation 3.11) is that the solution must be positive everywhere, as required for the decision rule to be a joint pdf.

The final constraint (equation 3.12) is the information processing constraint. Entropy $H[.]$ is a measure of the dispersion of a distribution. The first term of constraint 3.12 is the entropy of the prior. The prior reflects the information held by the household about the distribution of the hiring rate before receiving any signals. We will assume for now that this prior is the same for all households and is uniformly distributed (this is relaxed in section 3.2.8), so the prior is rather dispersed and entropy is high. The second term is the expected entropy of $f_i(\eta|c_{mi})$, the updated distribution over the hiring rate believed by the household after taking in the available signals[7]. A precise posterior knowledge of η would give a very low posterior entropy, so the entropy difference from the prior would be large. Information costs in this model are proportional to this difference, that is how much the agent can learn from the signals. Note that with identical prior beliefs, each household now faces exactly the same problem, and so each household chooses the same decision rule f_i. We therefore drop the i subscripts from the decision rule and the prior belief g. Households will still choose heterogeneous values for morning consumption, because the optimal signals contain idiosyncratic noise.

Given a particular level of information processing κ, the household must decide how to allocate that 'information processing budget'. They could, for example, ensure they make no mistakes at all when the hiring rate is above a certain threshold, but in doing so they must accept that they will make larger mistakes with higher probability when η is below that level, or pay for more processing capacity. This means that the household faces a trade-off: for a particular κ they can distinguish between a several values of η which are close together, but that reduces the entropy of the posterior a great deal, so

[6] In Matějka (2017), this marginal distribution of the variable subject to rational inattention is the true distribution of that variable. This will not be the case here, as η will be determined endogenously in the model. Instead, the marginal distribution obtained by integrating the joint pdf over consumption should be interpreted here as the distribution of the hiring rate the household is expecting to see from their 'ignorance prior' (see section 3.2.7).

[7] The 'information content' of the signals is incorporated into the choice of c_{mi}, so the conditional distribution of η given the choice of c_{mi} tells us what the agent believes about η. This is because of the one-to-one mapping between signals and actions discussed above.

outside of that small range of η their posterior $f(\eta|c_{mi})$ must remain dispersed. When η is in that small range, the household making that decision will be very accurate in choosing optimal c_{mi}, but when η is outside of that range they will make large mistakes with a high probability. Alternatively, they can choose to allocate their information processing capacity to distinguishing between a small number of cases which are far apart. They are then never precise in predicting the hiring rate, but they make large mistakes less often. This is what drives the result in Matějka (2017) that the agent optimally restricts themselves to a small number of discrete levels of the choice variable, even though a continuous range of that variable is available, when the optimal κ is sufficiently small that the information constraint binds.

3.2.3 Rational Inattention Solution

The hiring rate is naturally bounded by 0 and 1. Assume that prior beliefs are uniform over the whole of this range: $g(\eta) \sim [0, 1]$. The optimal decision rule for a marginal cost of information of $\psi = 0.002$ is plotted in figure 3.1. This cost is sufficiently high that the optimal information strategy is to collect signals which are less than perfectly informative about η. The information processed at this cost is such that the household optimally restricts themselves to two levels of consumption, even though under perfect information the optimal consumption choice is continuous in the hiring rate. The logic behind this is discussed in section 3.2.2 above, and in detail in Jung et al (2019) and Matějka (2017). As the hiring rate increases (and so the optimal choice of morning consumption under perfect information increases), the probability a household chooses the higher level of morning consumption in their 'menu' increases.

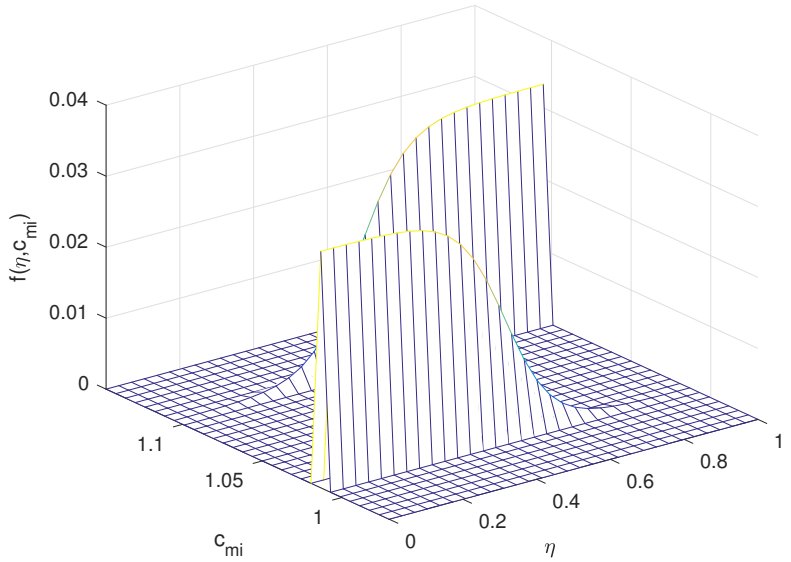


Figure 3.1: Optimal decision rule for $\psi = 0.002$

3.2.4 Aggregate Consumption

There is a unit mass of households making this decision. They all face the same labour market conditions, but we assume that they receive different noisy signals and/or they interpret those signals differently. Therefore for each level of the hiring rate some agents choose each of the morning consumption levels in the optimal 'menu' which arises from the RI problem with uniform priors (equations 3.9 - 3.12). The proportions on each level of consumption are determined by the probabilities in the optimal joint pdf obtained as the decision rule from the household problem.

Therefore for each level of η we obtain aggregate morning consumption c_m using:

$$c_m(\eta) = \int_{-\infty}^{\infty} c_{mi} f(c_{mi}|\eta) dc_{mi} \tag{3.14}$$

With no information processing[8], households choose consumption to maximise $\mathbb{E}_{mi}V$ given their prior belief (where $\mathbb{E}_{mi}\eta = 0.5$ for all households), as they cannot update their beliefs beyond that. They do not therefore change their consumption choice at all as the hiring rate varies. As the optimal information processing capacity κ increases (as ψ decreases), some information about η is processed, so households begin to choose different levels of c_{mi} for different underlying η. For low values of κ, households optimally restrict themselves to two values of morning consumption. Importantly, the aggregate morning consumption function has a wave-like shape around its perfect information equivalent.

As shown in Matějka (2017), as information processing κ rises further, more choices of c_{mi} are introduced into the optimal menu. As this occurs the aggregate response of morning consumption to the hiring rate approaches the perfect information first order condition. For ease of exposition, the graphs in this paper are all drawn for information costs that imply households choose an optimal menu with two levels of consumption, but this is not important for the results. An example with a lower information cost in this static model is shown in appendix C.1.

The aggregate consumption function for an example calibration is shown in figure 3.2. Consider first the case where information processing is constrained but non-zero (the blue curve). This meets the full information consumption function (in red) at $c_m = 1.06, \eta = 0.5$. However, at this hiring rate in the full information model, every household consumes the same amount. In contrast, in the rational inattention model, half of the households get a signal that the hiring rate is 'high' and choose the high level of $c_{mi} = 1.105$. The other half receive a signal that η is 'low', and so consume the lower level $c_{mi} = 1.013$.

[8]Households cannot process negative amounts of information - they cannot choose to forget information in their prior belief. That means that for any information cost ψ above a certain threshold, households hit this no-forgetting constraint and process no information ($\kappa = 0$).

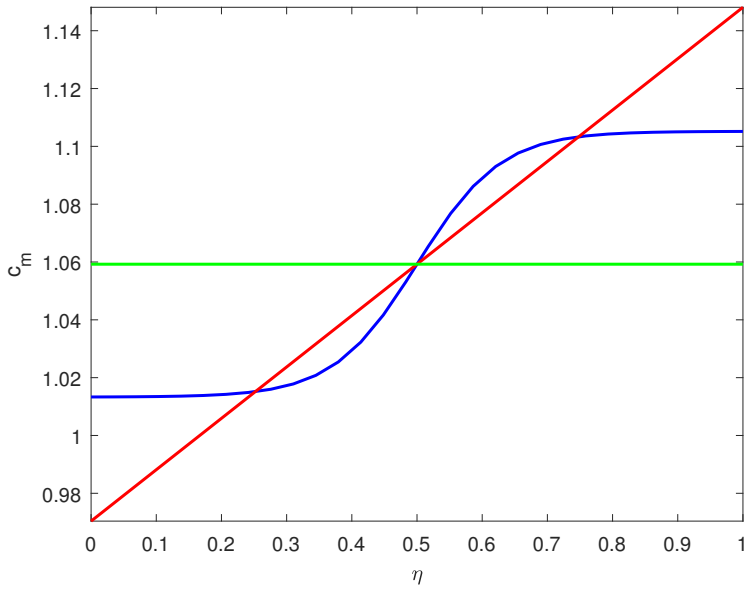


Figure 3.2: Aggregate morning consumption function with $\kappa = 0$ (green), $\kappa = 0.5$ (blue) and in the unconstrained case (red)

The flat sections of the aggregate consumption function under rational inattention occur where changes in the hiring rate do not lead to much change in the proportions of agents choosing each level of consumption in their menus. In figure 3.1 above, it can be seen that this is the case for extreme high and low values of η, and $c_m(\eta)$ is flat in these regions accordingly. In contrast, as η moves from 0.4 to 0.6, large numbers of households switch from choosing the low level of consumption to the high level, and this corresponds to the steep section of the corresponding aggregate morning consumption function in figure 3.2.

The shape of the aggregate consumption function is therefore driven by the shape of the curves in the optimal decision rule: if the probability of choosing the low level of consumption in figure 3.1 fell linearly as η increased the aggregate consumption function would be linear. In fact, the distribution of $c_{mi}|\eta$ for the values of c_{mi} in the optimal menu is logistic in shape, which is what gives rise to the wave-like shape of the aggregate response curve[9]. Beliefs, and so choices, are therefore endogenously sticky in certain regions of the support of η, as a result of the optimal signal structure that comes out of the entropy-based cost function. It is this non-linearity which leads to multiple equilibria in our model.

3.2.5 Firms

Usually, rational inattention models specify that agents collect costly information about exogenous variables, often shock processes. In contrast, we assume that the hiring rate

[9]Matějka and McKay (2015) study in detail how RI leads to the logit model.

is determined in equilibrium by firm hiring decisions[10]. For the purposes of this simple static example, it will be sufficient to say that firms hire more workers when aggregate demand is high, so:

$$\eta = H(c_m), \text{ with } H'(c_m) > 0 \tag{3.15}$$

The focus of this static model is household information choices, and the aggregate consumption function these imply, so for simplicity we use a linear H function throughout this section, though this is not necessary for our results. Appendix C.2 microfounds such a process for the hiring rate using a model with working capital. In section 3.3 the firm side of the model is standard, as in Ravn and Sterk (2020).

3.2.6 Equilibrium

It is important here that the hiring rate η is not necessarily uniformly distributed like the prior beliefs of the households. In Matějka (2017) and many other models in the rational inattention literature, agents' prior beliefs about the distribution of the unknown variable are correct. We extend the RI literature by considering prior beliefs which do not precisely match the true equilibrium distribution of the relevant variables. This may be a plausible situation for information processing about variables which are difficult to learn about, perhaps because they are not easily understood, or because the data is not reported at the front of central bank communications and other news sources, so the variable's history is not easy to observe. This assumption is related to the 'internal rationality' in Adam and Marcet (2011). The survey data in section 3.4 suggests that for the hiring rate, households do not know the true equilibrium distribution.

In particular, we begin by assuming that households have a uniform 'ignorance' prior, which would be justified if the households do not understand how their decisions (and those of other agents) affect the hiring rate[11]. This is discussed in section 3.2.7. We study equilibrium in the static model with these uniform priors. In section 3.2.8 we take this as a starting point and repeat the static problem many times, allowing households to update their prior beliefs each period. We show that the multiplicity does not disappear when agents update their priors away from the uniform starting point.

The separation of the prior belief and the true equilibrium distribution of η means

[10] In a standard model where agents perfectly understand how endogenous variables are determined this distinction is irrelevant, as agents know the mapping from shocks to endogenous variables. The distinction is relevant here because we assume that households do not know how the hiring rate is determined in equilibrium.

[11] The initial prior does not need to be necessarily uniform for our results. We require a prior belief which is significantly dispersed, hence the term 'ignorance prior', but the precise distribution is unimportant. The prior belief must also be bounded, which is ensured in this case as the hiring rate must be between 0 and 1 by definition.

that the aggregate consumption function $c_m(\eta)$ remains as in figure 3.2, and the hiring rate is then determined endogenously by the interaction of this aggregate consumption and the firm hiring function 3.15.

The graph below shows this equilibrium interaction. The blue and red aggregate consumption functions are as in figure 3.2. The equilibrium condition 3.15 is added in black. Under full information there is one equilibrium, but under RI there are three: an unemployment trap with low morning consumption, and a low hiring rate, so low employment, a middle equilibrium and an 'employment trap' with very a high hiring rate and consumption.

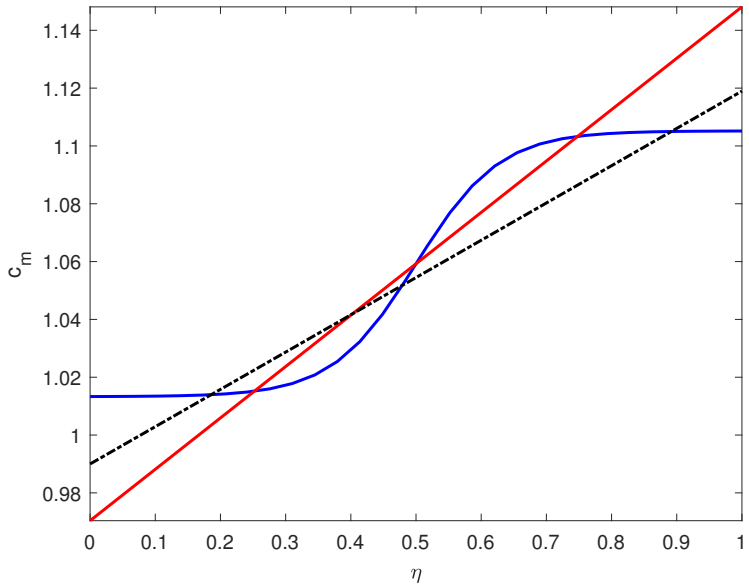


Figure 3.3: Aggregate consumption response to changes in the hiring rate for full information (red) and rational inattention with $\psi = 0.002$ (blue), with the firm condition 3.15 in black

Consider first the middle equilibrium under rational inattention, at $\eta = 0.48$. As explained in section 3.2.4, at this hiring rate, close to the middle of the uniform prior beliefs, aggregate consumption is close to that under full information, but there is dispersion in household choices underlying this which is not present with full information. The key difference between aggregate consumption in these two models comes when η falls a little from this central equilibrium. Under perfect information, all agents respond the same way, by reducing their consumption a little. With rational inattention, however, the fall in the hiring rate leads to a large number of agents switching from the high level of morning consumption $c_{mi} = 1.105$ to the low level $c_{mi} = 1.013$. This means that aggregate consumption falls a great deal, so much so that there is another equilibrium at $(\eta = 0.19, c_m = 1.014)$ in which almost all households choose the low level of consumption. This is the unemployment trap. Similarly, there is a high employment steady state at $(\eta = 0.89, c_m = 1.105)$ which is also not present under full information.

In effect, households are using their limited information processing capacity to decide if they face a 'high' or 'low' hiring rate. As η moves a little below 0.5, the majority of agents decide on 'low', and consume accordingly, whereas if they knew the hiring rate more precisely they would choose consumption based on only a 'somewhat low' η. However that in itself would not be sufficient for multiplicity. It is crucial that, as the hiring rate falls even further, households do not decide on even lower consumption. They continue to believe that the hiring rate is 'low' even when it become 'extremely low'. That is why the aggregate consumption function flattens out at very low values of the hiring rate, which is why there is an equilibrium with very little labour market activity.

3.2.7 Requirements for multiplicity

The multiplicity of equilibria is driven by the non-linearity in the aggregate consumption function, which arises endogenously from the optimal information processing decisions of households. The first key requirement for our result is that households process costly information about the hiring rate, with an 'ignorance prior'. That is, households have only imprecise knowledge of the true equilibrium distribution of the hiring rate.

To see why, consider the equilibrium in figure 3.3. The true equilibrium distribution of η contains just three discrete points. If household prior beliefs matched this distribution, the optimal level of consumption in their decision rule when they received a signal that η was low would be closer to 1.06 (the optimal value when $\eta = 0.5$) than it is in the figure, as they do not need to worry about the possibility of an extremely low η. This, in turn, would mean that the equilibrium value of η in the unemployment trap would be higher. Iterating this logic and also applying it to the high employment trap, we can see that if beliefs were indeed close to the true distribution of η, the only equilibrium to survive would be the full information equilibrium.

The households in most standard rational inattention papers (e.g. Maćkowiak and Wiederholt, 2015) do not have ignorance priors. In such models households understand all of the direct and general equilibrium effects mapping shocks into endogenous variables, and they know the distribution of shocks, so they can deduce the equilibrium distributions of all endogenous variables. That kind of model therefore assumes that households have a large amount of information about mechanisms at work in the economy (including how other households make decisions), but then it places limits on the information that those very well-informed households can obtain about the *realisations* of the shocks.

Our assumption of 'ignorance priors' builds on these existing rational inattention models by assuming that households do not perfectly understand the links from shocks to aggregate household choices to the hiring rate[12]. This assumption is related to the

[12]This is significantly easier to solve than a general equilibrium rational inattention model without

removal of 'external rationality' in Adam and Marcet (2011). Like them, we believe that it is plausible that agents do not precisely know the true stochastic processes and mechanisms underlying the determination of the endogenous variables they face, especially when those mechanisms are complicated and involve the choice behaviour of many other agents[13]. The survey data on hiring rate expectations in section 3.4 supports this view: the variance in survey expectations (even after accounting for a wide variety of household characteristics) is an order of magnitude larger than the long run variance of the true hiring rate. It appears that, just as in our model, households do not know the appropriate range of values for the hiring rate.

However, even if households do not understand the mechanisms leading to a particular equilibrium distribution of the hiring rate, they could obtain accurate prior beliefs if they had observed the hiring rate over many periods, and had learned its equilibrium distribution over time. This is related to the concept of 'near-rationality' in Adam and Marcet (2011): in their model subjective beliefs are sufficiently close to the true equilibrium process that agents cannot distinguish between the two given the data they observe. In contrast, we start from a uniform prior, which is extremely dispersed and far from the true equilibrium distribution. In fact, we do not require that the prior is uniform, but it does need to be more dispersed than the true equilibrium distribution and be bounded. In section 3.2.8 we show that as long as agents start at some point in time with such dispersed priors, and they can change their information processing over time, our results hold even when households update their prior beliefs each period. This is because the initial reduction in dispersion of prior beliefs after one period of information processing leads households to process less information in the following periods, as households then rely more on their more informative priors and less on new signals (as in Matějka, Steiner, and Stewart, 2017). Prior beliefs therefore do not approach the true distribution of the endogenous variable[14].

The other important condition for multiplicity is that the firm hiring function $H(c_m)$ must be upward sloping. The aggregate consumption function under rational inattention is non-linear, but is always upward sloping. If it was downward sloping at any point, then an increase in the hiring rate must be leading to a rise in the probability that households get a signal that the hiring rate is low. An information strategy that leads to expected

ignorance priors, because higher order beliefs cannot matter for household choices when households do not understand how the actions of others feed into the hiring rate

[13]We also follow Adam and Marcet in assuming that the agents in our model make choices that rationally maximise expected utility given their beliefs.

[14]The idea that agents might start with a uniform prior when they first attempt to learn about something, then update, has been suggested as an explanation for experimental decision making results (e.g. Fox and Clemen, 2005). We only differ from that account in that our agents update rationally given an information processing constraint, rather than through a behavioural heuristic.

consumption falling when labour market prospects improve cannot maximise expected utility when the full information consumption function implies $\frac{dc^*_{mi}}{d\eta} > 0$. The rational inattention aggregate consumption function cannot therefore be downward sloping. This means that if $H'(c_m) < 0$, there will be just one equilibrium.

We therefore need a degree of strategic complementarity between households and firms for our multiplicity results. This is very plausible in models of precautionary saving based on labour market expectations: Ravn and Sterk (2020) show that this complementarity exists as long as labour income risk is countercyclical, which is satisfied if real wages are approximately acyclical. Strategic complementarity often increases the volatility of consumption and labour market variables in response to shocks, but it only leads to multiple equilibria if one or more model equations is sufficiently non-linear. In Ravn and Sterk (2020), there is an unemployment trap steady state at a hiring rate of zero because the Phillips Curve is kinked at that point by the requirement that vacancy posting cannot be negative[15]. Our model is different because the non-linearity in the consumption function that generates the multiplicity does not come from such an imposed cutoff (though we agree that imposing that vacancies must be positive is sensible), or from some exogenously imposed non-linearity in another part of the model. The non-linearity arises endogenously from optimal household information choices. For this reason, in section 3.3 we find an unemployment trap with low, but positive, labour market activity, whereas the unemployment trap in Ravn and Sterk has zero employment. We also find a high employment trap.

3.2.8 Dynamic Solution

Our multiplicity result relies on household prior beliefs being more dispersed than the true equilibrium distribution of the hiring rate. In a model where households process information about the hiring rate each period, they might update their prior beliefs using information processed in previous periods. Our multiple equilibria will only survive in the long run if this prior belief updating does not lead to beliefs collapsing to the true distribution of the hiring rate over time. The model in section 3.2.1 is static, so there is no room for agents to learn in this way. In this section, we repeat the static problem from that simple model to give households the opportunity to learn about the distribution of the hiring rate. For this section, a 'period' is composed of a morning and afternoon. As in section 3.2.1, all households are employed each morning, and some lose their job at the

[15]In appendix A.1.2 they generate an unemployment trap steady state with a positive hiring rate through a different non-linearity that enters the household Euler equation when the assets held by households have a positive net supply. In section 3.3 we focus on the case where the net supply of assets is zero, so this additional nonlinearity is not present.

end of the morning. Firms then hire new workers for the afternoon in a frictional labour market. As in the static model, households consume all of their income and savings each afternoon, so the income inequality created by the frictional afternoon labour market does not lead to wealth inequality, which would otherwise need to be tracked between periods. Households start from the uniform prior belief in the first period, but they can use information from one period to update their prior beliefs about the distribution of η for the next period. Household prior beliefs are the only link between periods. We show that in this simple model, beliefs do not converge to the true distribution of η, even in the long run. Multiple steady states therefore persist even when households update their prior beliefs over time.

We will consider a simple rule for updating beliefs. If household i reaches the posterior belief $f(\eta_t|c_{mit})$ after processing signals in the morning of period t, we will suppose that their prior belief before information processing in period $(t+1)$ is given by:

$$g_i(\eta_{t+1}|c_{mit}) = \rho f_i(\eta_t|c_{mit}) + (1-\rho), \quad \rho \in [0,1] \tag{3.16}$$

Intuitively, we take a weighted average of last period's posterior and the uniform $(0,1)$ initial prior, with the weight ρ interpreted as a measure of the (perceived) persistence of the hiring rate η. The particular updating rule is not important, what matters is that households take some of the information processed to arrive at their posterior belief in one period and use it to inform their prior belief in the next period. True Bayesian updating is not possible because that requires that households know the process generating the hiring rate, which is ruled out by our use of ignorance priors.

A further implication of ignorance priors is that households cannot precisely compute the value that information processed in one period will have in future periods. We therefore follow Kreps (1998) in assuming households maximise 'anticipated utility'. That is, when households make their choices in period t they do not take into account the informational value of those choices for future periods. This is a standard way to avoid the intractability that comes from optimisation problems where current choices generate information which may be useful in future periods, and it is particularly relevant here since households have no way of precisely valuing future information due to their ignorance of the process determining the hiring rate. Cogley, Colacito and Sargent (2007) showed that in a simple monetary model the solution to the anticipated utility problem is a good approximation to the fully rational model where future information values are taken into account.

With this assumption, the first period optimal decision rule $f(\eta_1, c_{mi1})$ (from the uniform prior) is as in figure 3.1. This is because the anticipated utility assumption means that the agent acts as if they face a series of unconnected static problems, since

nothing else connects the periods in this simple example. If the prior is uniform in period 1, the first static problem is exactly the same as in section 3.2.2. There are two possible posterior beliefs that the households could finish period 1 with, corresponding to the two levels of morning consumption in the optimal menu. If a household chooses the high level of consumption, it is because they have received a signal that the hiring rate is high, and similarly the low value of consumption corresponds to a low signal. There are therefore two possible signal realisations, and so two possible posterior beliefs. Formally, these posteriors are given by the conditional distribution $f(\eta_1|c_{mi1})$.

The period 2 priors will be weighted averages of these posteriors and the initial uniform prior. As long as $\rho > 0$, the households will all start period 2 with a prior with a lower entropy than their period 1 prior. That is, as long as households believe that η has some persistence, they will incorporate some period 1 information into their period 2 prior.

In period 2 this model diverges from the static case. Remember that agents choose how much information to process (κ) such that the marginal benefits of more information equal the marginal cost ψ. Information in prior beliefs is a substitute for information from new signals. This means that the marginal benefits of information are increasing in the entropy of the prior: if prior beliefs are already very informative (low entropy) there is little extra benefit from more information. In period 2, the entropy of each household's prior is lower than it was in period 1. The households therefore process less information in period 2 than they did in period 1. This implies that the households rely more on their priors to guide their decisions in period 2 than they did in period 1, because the period 2 priors contain more information. If they chose a low morning consumption in period 1, their posteriors in period 1 must have suggested that η_1 was low. With little new information processing in period 2, the agent is therefore more likely to choose a low morning consumption in period 2 than a household who chose a higher morning consumption in period 1. This inertia in choices is studied in detail by Matějka, Steiner and Stewart (2017). This is why our model features persistence in individual expectations, as seen in the data.

As long as $0 < \rho < 1$, period 2 priors are more dispersed than period 1 posteriors[16], and are less dispersed than the uniform prior from period 1. The households therefore start period 2 with more precise priors than period 1, but process less information to

[16]If $\rho = 1$, households simply take their posterior and use it as the next period's prior without adding any extra noise. In period 1 (with a uniform prior), they process information until the marginal benefit of being informed equals the marginal cost ψ. In the next period, their prior belief would contain the same amount of information as the period 1 posterior. Any further information processing would therefore have a marginal benefit below ψ. If a hiring rate is an equilibrium in the period with a uniform prior, it is therefore trivially a steady state, as no household changes their actions in any future periods. Similarly, if $\rho = 0$ households do not update their prior beliefs between periods and the model is a sequence of unconnected static problems, so any equilibrium in the static model is also a steady state.

update from that prior to a posterior. It is therefore not the case that priors degenerate to the true distribution of η over time. In fact, that cannot be the result: each period the households process enough information to return them to the point where the marginal benefit of information equals ψ. We know from period 1 that this occurs before posteriors completely pin down η_t.

The key component of our model that prevents prior beliefs collapsing to the true distribution of the hiring rate is that households don't learn from past realisations of the hiring rate, they learn from past signals they received about η (by assumption they never observe past values of the hiring rate). These signals are costly, and the marginal cost of more information is constant. The marginal benefit of more information falls as priors become more informative, so learning in our model does not lead to beliefs converging to something close to the true distribution of η. If instead households observed the history of realisations of the hiring rate without cost, prior beliefs would converge to the true distribution of the hiring rate[17].

For this simple model, we simulate to find candidate steady states. We choose a level of $\eta_t = \bar{\eta}$ and fix it over time. In period 1, a proportion L choose the low consumption c_l, and $H = (1 - L)$ choose the high c_h, in the optimal menu. The aggregate c_{m1} is simply $Lc_l + Hc_h$. In period 2, that means that a proportion L have priors which bias the agents towards low consumption, and H are biased towards high c. We solve the RI problem for both of these groups, and obtain new posteriors and proportions of households holding each posterior. We iterate this process until aggregate consumption and the composition of beliefs in the population are stable[18]. There are two types of prior belief that emerge in steady state, which are shown in figure 3.4 below.

[17]To be precise, they would converge to a distribution which is statistically indistinguishable from the true distribution. This is the 'near-rationality' in Adam and Marcet (2011).

[18]By this, we mean that there is a steady state set of prior beliefs which all households move between, and that the proportions of households on each prior belief each period in the set is constant, although individual households churn between priors. The steady state set of beliefs is such that if a household holds a prior belief in the set, their optimal information acquisition will always lead them to a prior belief in the next period which is an element of the set (up to a small degree of approximation).

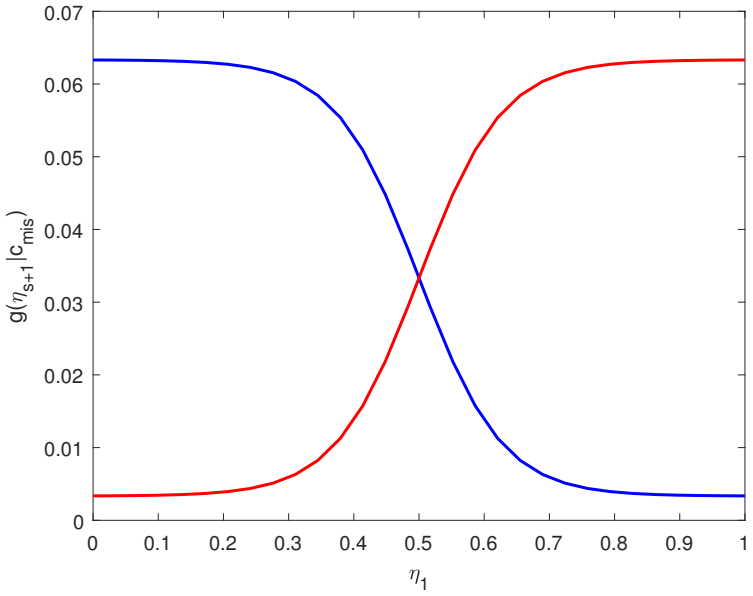


Figure 3.4: Prior beliefs seen in steady state with $\psi = 0.002$ and $\rho = 0.9$.

A household which receives a signal in period s that the hiring rate is low forms the prior belief plotted in blue in the next period $s+1$. This is biased towards low hiring rates because the household has incorporated that information from the previous period into their priors: we refer to a household with these prior beliefs as *pessimistic*. In period $s+1$ that household processes a small amount of new information, and because the prior belief is still bounded the optimal signal structure contains a discrete number of possible realisations. In fact, with the information cost parameter we use for our graphs the optimal signal contains two possible realisations, as it did in the initial period with a uniform prior[19]. This household, with a prior biased towards low hiring rates, does not process very much new information in period $s+1$, and so the posterior belief they form if they again receive a signal that the hiring rate is low is very close[20] to their posterior from the previous period s. The posterior they form if they receive a signal that the hiring rate is high is biased instead towards high η. This implies a prior belief in period $s+2$ as plotted in red in figure 3.4. We refer to households with this prior belief as *optimistic*: like pessimists they too process very little new information each period, and their optimal signal structure has two possible realisations, which lead them to either stay as optimists or transition to a pessimistic prior belief.

Figure 3.5 below plots the candidate steady states of the model when $\rho = 0.9$: the aggregate morning consumption c_m that comes out of holding η fixed at each possible level and iterating until aggregate consumption and the proportions of households on

[19]Running this problem with an information cost such that the optimal signal structure with a uniform prior has three possible realisations gives three types of steady state prior, with the optimal information strategy from each of those priors involving a signal with three possible realisations.

[20]That is, the same up to a small degree of error from the numerical solution method.

each prior belief are constant. Again, the equilibrium condition from the firm side of the model 3.15 is in black, and the model has steady states where the blue set of candidate steady states meets this equilibrium condition.

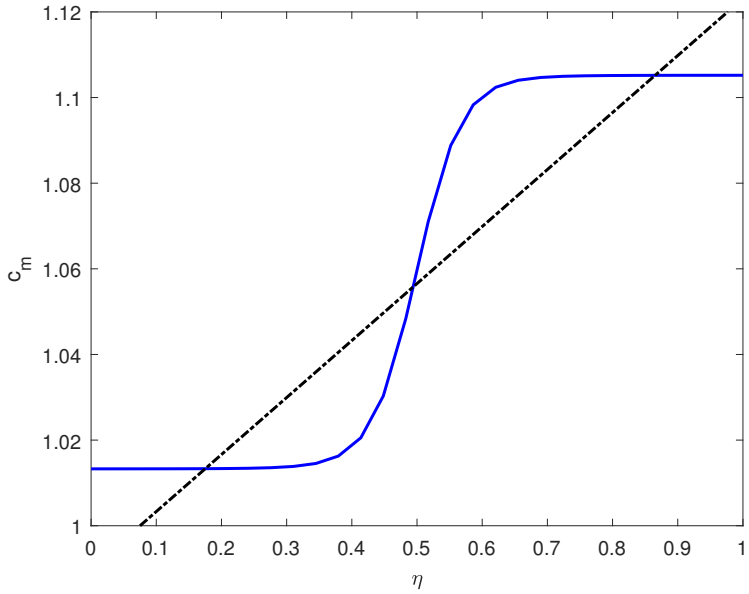


Figure 3.5: Candidate steady state c_m for each η.

If beliefs did degenerate to the truth, this curve would coincide with the linear full information consumption function[21] in figure 3.2. In fact we retain the logistic shape of the aggregate response curve from the static problem, so this model generates multiple steady states. There is a high employment trap steady state with $\eta = 0.89$, a middle steady state with $\eta = 0.49$, and an unemployment trap steady state with $\eta = 0.18$. Note that the blue curve here is closely related to the consumption function in figure 3.2. The figure 3.2 consumption function plots the combinations of the hiring rate and aggregate consumption which are consistent with household optimal information and consumption decisions. That is also true for the aggregate consumption function in figure 3.5, which adds to household optimality that the combination of η_t and c_{mt} must be consistent with a stable distribution of household prior beliefs.

The condition that the distribution of prior beliefs must be stable does not imply that each household makes the same decision each period. There is churn at the level of individual households, but the distribution of beliefs in the population is constant at every point along the blue steady state aggregate consumption function. The tables below show the transition probabilities between the two types of prior beliefs seen in steady state (pessimistic **P** or optimistic **O**) at the three steady states.

[21] We hold the hiring rate constant, so the true distribution of the hiring rate is simply a point. If households know this, with no further information processing they choose the optimum consumption for that hiring rate.

Table 3.1: Prior belief transition matrices

(a) Unemployment trap

	$\mathbf{P}_{s+1}$	$\mathbf{O}_{s+1}$
$\mathbf{P}_s$	0.9995	0.0005
$\mathbf{O}_s$	0.8919	0.1081

(b) Middle steady state

	$\mathbf{P}_{s+1}$	$\mathbf{O}_{s+1}$
$\mathbf{P}_s$	0.9514	0.0486
$\mathbf{O}_s$	0.0565	0.9435

(c) High employment steady state

	$\mathbf{P}_{s+1}$	$\mathbf{O}_{s+1}$
$\mathbf{P}_s$	0.0619	0.9381
$\mathbf{O}_s$	0.0002	0.9998

In the middle steady state, both optimistic and pessimistic beliefs are very persistent, so there is a great deal of belief heterogeneity in steady state. Each period 54% of households hold pessimistic prior beliefs and 46% hold optimistic beliefs, and individual households occasionally switch from one type to the other. In the unemployment trap, the hiring rate is so low that even households with optimistic prior beliefs are very likely to receive a signal that η is low, and so to switch to pessimistic beliefs. There is therefore much less heterogeneity in prior beliefs in the unemployment trap steady state: 99.94% of households hold pessimistic beliefs each period, though even here there is a small amount of churn in beliefs. Similarly, in the high employment steady state 99.98% of households hold optimistic beliefs.

The amount of belief dispersion in steady state is regulated by the information cost parameter ψ: with a lower information cost households would choose signals with more than two possible realisations, so households would try to distinguish between several different low values of the hiring rate. This would create more belief dispersion in steady states with a low η. This would not remove the multiple steady states: the rational inattention aggregate consumption function would still have the step-like shape displayed in figure 3.5, there would simply be more steps. In this case rational inattention may generate more than three steady states.

3.2.9 Stability under belief mutations

After finding multiple steady states in this model, a natural follow-up question is how the system behaves when shocked away from those steady states. We analyse the stability of each steady state to shocks to the distribution of prior beliefs in the population, and find that the unemployment trap and high employment trap steady states are locally stable, while the middle steady state is unstable.

In our simple model, the only variable linking separate periods is the prior beliefs of households, which are formed using the posterior beliefs formed after processing signals in the previous period. The steady states are therefore defined by the distribution of prior beliefs in the population. In the spirit of Kandori, Mailath and Rob (1993)(KMR) we define a *mutation*[22] as a perturbation of the proportions of households holding pessimistic

[22]KMR study a 2x2 game with multiple Nash equilibria. They consider a large population of players

and optimistic prior beliefs, and we study the response of the endogenous model variables over time to such a shock. Specifically, suppose that at the beginning of each period a shock arrives that causes some pessimistic households to switch to holding optimistic priors, and some optimists to switch to pessimistic priors, before the households choose their information strategy for the period. The mutation γ_t is therefore an exogenous shock to the proportion of households who hold pessimistic prior beliefs before processing signals in period t.

The interpretation of these mutations in KMR is that each period a fraction of households leave the economy, to be replaced by new households who have a different probability of being optimistic or pessimistic than the households they replaced. Alternatively, we could think that in the afternoon of each period households start to doubt the signals that led them to their morning consumption choice. When they come to form their priors the next morning, some decide they must have been wrong in the previous period and act as if they had received the opposite signal realisation[23].

We study the stability of each steady state to these shocks using simulations[24]. Starting from each steady state, we subject the distribution of prior beliefs to a small mutation shock in an initial period t_0, then return γ_t to 0 for all future periods $t > t_0$.

Each period, we take the distribution of prior beliefs and find the aggregate consumption function, and therefore the equilibria of the model in that period. If there are multiple equilibria, we select the equilibrium in which the hiring rate η_t is closest to its previous period value η_{t-1}, following the inertia assumption in KMR. Combining the prior belief distribution at the start of the period with the realised hiring rate we find the distribution of posterior beliefs in the population, from which we find the prior beliefs for the start of the next period. We repeat these steps for many periods after the period t_0 mutation, tracking the evolution of the hiring rate and the distribution of prior beliefs.

We find that after small mutations (positive or negative γ_{t_0}) from the unemployment trap and high employment steady states, the model quickly returns to that steady state. In contrast, the middle steady state is unstable with respect to any mutation away from the steady state belief distribution. A small positive γ_{t_0} (increase in pessimistic households) causes the model to transition to the unemployment trap, and a negative γ_{t_0} causes a shift to the high employment trap.

meeting at random, and perturb the proportion of players choosing each strategy. While our setting differs from this, we similarly consider a large population of players (households) with two possible 'types' - prior beliefs instead of the strategies in KMR.

[23]When a household chooses their information and action strategies at the start of each period they choose the possible realisations of their signals, and how they will respond to them, so households know precisely what the 'opposite signal realisation' would have meant.

[24]All simulations use the parameter values used to draw all of the simple model figures, which are detailed in appendix C.2.

The intuition for this can be seen in the steady state diagram in figure 3.5. All households who receive a signal that η is high choose $c_{mh} = 1.105$, and those households have optimistic priors in the next period. Those receiving a signal that η is low always choose $c_{ml} = 1.013$. Suppose we start from one of the steady states, then consider a 1 percentage point increase in the proportion of households with pessimistic priors. If that new composition of priors was to be sustained in steady state, aggregate consumption must therefore fall from the original level by $0.01 \cdot (c_{mh} - c_{ml}) = 0.00092$. This required consumption fall is the same whichever steady state we start from.

Suppose consumption really did fall by that required amount after the shock. The steady state aggregate consumption function tells us how much the hiring rate would need to adjust for that new proportion of pessimistic priors to be maintained into the next period, and the firm equilibrium condition tells us how much the hiring rate *actually* adjusts, since unlike the steady state consumption function it must be satisfied every period. If we start from the unemployment trap, the steady state consumption function is shallower than the firm condition, so the hiring rate falls by less than is necessary for the new, mutated proportions to be sustained. The proportion of pessimists therefore falls in the period after the shock, back towards the unemployment trap steady state. The same is true at the high employment steady state. At the middle steady state, however, the aggregate consumption function is the steeper of the curves, and so for that given fall in consumption the hiring rate actually falls by a greater amount than would be necessary to sustain that new belief composition. The proportion of households with pessimistic priors therefore grows, pushing the economy towards the unemployment trap[25].

This analysis focuses on small mutations. It is, however, possible to cause the model to transition away from the unemployment and high employment traps if the mutations are large enough, or if there is a series of mutations. For one-off mutations, the proportion of pessimists needs to decrease by at least 44 percentage points to move from the unemployment trap to the high employment steady state. That is, the initial mutation must be such that the proportion of pessimistic households in the population falls below that seen in the middle (unstable) steady state.

Shifting just beyond the proportions in the middle steady state from the high employment steady state (with a 56 percentage point increase in the proportion of pessimists), however, is not sufficient to cause a transition to the unemployment trap. This is be-

[25]Note that this does not necessarily mean that the middle steady state is unstable with respect to all shocks. We study these kinds of mutation shocks because they leave the household information and choice problems, the steady state prior 'types', and the firm side of the model identical, which makes them tractable. This will not generally be true of other shocks. Movements in the firm equilibrium condition, or other aspects of the household problem than simply the priors, will cause shifts in either or both of the two curves defining equilibrium, and it is not clear whether the middle steady state would remain unstable with respect to those kinds of movements.

cause there are three equilibria possible when beliefs are distributed in that way: one at $\eta = 0.49$ (the middle steady state), one with $\eta = 0.69$, and another with $\eta = 0.87$. The inertia assumption means we select this highest equilibrium hiring rate, which means most households continue to receive a signal that η is high, and so in the next period the prior belief distribution is again close to that in the high employment trap, so the model returns to that steady state. To transition to the unemployment trap the mutation must be sufficiently large that such a high employment equilibrium ceases to exist, which for these parameters occurs when the proportion of pessimists is increased in period t_0 by 88 percentage points.

3.3 HANK model

Here we study the HANK model in Ravn and Sterk (2020) (RS) with the addition of rational inattention to the hiring rate and prior beliefs which do not match the true equilibrium hiring rate distribution, to show that our results continue to hold in richer models than that studied above. The RS model is particularly useful because it remains tractable, despite featuring the uninsurable labour market risk which is necessary to generate a precautionary savings motive.

The tractability is achieved by assuming that there are 'asset-rich' households who are risk neutral, own firms and do not participate in the labour market. The remaining 'asset-poor' households supply labour to firms in a frictional labour market, and cannot borrow. Bonds are in zero net supply, so all asset-poor households hold zero wealth in equilibrium. Employed households are on their Euler equation and are all identical, and unemployed households are hand-to-mouth, consuming their home production. There are therefore only three types of households in the model, not the full distribution seen in other models with uninsurable idiosyncratic risk (e.g. Kaplan, Moll, Violante 2018). Employed asset-poor households are the critical households as they remain on their Euler equation, so they price the bond in equilibrium.

We make one significant change to the Ravn and Sterk model: we remove the constraint that vacancy posting cannot be negative. In RS, this constraint places a kink in the Phillips Curve at a hiring rate of zero, which gives rise to an unemployment trap steady state with zero hiring. Without this, the model has a unique steady state under full information[26]. We show that rational inattention generates multiple steady states with positive hiring in this environment.

[26]Note that vacancy posting will never be negative in steady states we find. Removing the explicit constraint simply removes the kink in the Phillips Curve which allows for a zero hiring steady state.

3.3.1 Households

Employed households choose how much to consume and save each period based on their current wealth and their expectations about the future hiring rate. Unemployed households face a binding no-borrowing constraint, so consume all of their income from home production each period. Each period an exogenous proportion ω of employed households lose their job, and an endogenous proportion η_s of unemployed households find a job. This is the same as the household problem in Ravn and Sterk, except that here employed households are allowed to borrow[27]. Note that there is no morning/afternoon distinction as in section 3.2: if a household loses their job at the end of period s, they start to search for new employment in period $s+1$.

The consumption Euler equation under perfect information is as in RS:

$$c_{es}^{-\mu} = \beta \frac{R_s}{\Pi_{s+1}} \Big(\omega(1-\eta_{s+1}) c_{u,s+1}^{-\mu} + (1-\omega(1-\eta_{s+1})) c_{e,s+1}^{-\mu} \Big) \tag{3.17}$$

Employed households choose consumption in period s, c_{es}, such that the marginal utility of consumption equals the expected marginal utility of future consumption over the two possible future labour market states, adjusted for discounting and the real interest rate[28]. Unemployed households are always at their borrowing constraint, and so their problem never matters for equilibrium determination.

To get the steady state Euler equation (**EE**) note that since equilibrium asset holdings are always zero $c_{es} = w$ and $c_{us} = \vartheta$, where ϑ is the payoff to home production. Then substitute for a simple interest rate rule[29]:

$$R_s = \bar{R} \bar{\Pi}^{-\delta_\pi} \Pi^{\delta_\pi} \tag{3.18}$$

This gives:

$$1 = \beta R^* \Pi^{\delta_\pi - 1} \Big[\omega(1-\eta) \Big(\frac{\vartheta}{w} \Big)^{-\mu} + 1 - \omega(1-\eta) \Big] \tag{3.19}$$

Where $R^* = \bar{R} \bar{\Pi}^{-\delta_\pi}$.

[27] If an employed household chooses to borrow, and then loses their job at the end of that period, they must return to weakly positive saving in their first period of unemployment. With full information allowing employed households to borrow makes no difference as if one household wants to borrow they all do, so there is no-one to lend to them. Equilibrium asset holdings will therefore still have to be zero for each household. The assumption will, however, matter under rational inattention.

[28] For the derivation see appendix C.3.1.

[29] This is as in RS, abstracting from the zero lower bound on interest rates and setting the coefficient on the hiring rate in the interest rate rule to 0, as RS do in drawing their figure 3.

3.3.2 Firms

Firms in our model are identical to those in RS, except they do not face a non-negativity constraint on vacancies. Firms set prices and choose how many vacancies to post each period to maximise profits. There are quadratic price adjustment costs as in Rotemberg (1982), firms are monopolistic and households have CES preferences. The price setting and vacancy posting decisions lead to a Phillips Curve, which in steady state becomes[30]:

$$\phi(1-\beta)(\Pi-1)\Pi = 1-\gamma+\gamma(w+k\eta^{\frac{\alpha}{1-\alpha}}(1-\beta(1-\omega))) \tag{3.20}$$

Here ϕ measures the extent of price adjustment costs, γ is the elasticity of substitution between goods in the consumer's problem, k is the cost of posting a vacancy, ω is the (fixed) job separation rate, η is the hiring rate and q is the vacancy filling rate, equal to $\eta^{\frac{-\alpha}{1-\alpha}}$, where α is the elasticity of the Cobb-Douglas labour matching function wrt job searchers.

Equation 3.20 is identical to equation (**PC**) in RS, except that assuming constant wages means that real wage w is not a function of η, and λ_f is dropped. This is the Lagrange multiplier on the constraint that vacancies must be weakly positive. We drop this constraint, so λ_f is always zero in our version of the model. In steady state, vacancies will always be positive in our model.

3.3.3 Rational Inattention

As in section 3.2.2, we now amend the household problem so that processing information about the next period hiring rate has marginal cost ψ. Again, we start with uniform prior beliefs $g(\eta) \sim U(0,1)$ for all households, then allow households to update their prior beliefs using information processed in the previous period, as in section 3.2.8. The equations of the household problem therefore mirror the setup explained in section 3.2.2. They are set out in appendix C.3.4.

Relaxing the no-borrowing constraint on the employed means that heterogeneity in the signals households receive implies a non-degenerate wealth distribution among the employed in steady state. Newly unemployed households may therefore have some savings in some cases, but their Euler equation is still unimportant for the determination of steady state because they all hit their borrowing constraint when they become unemployed for all inflation rates and hiring rates consistent with steady state and equilibrium[31].

[30] For the derivation see appendix C.3.2.

[31] That is, no unemployed household will choose to save if they have any of the saving levels chosen by employed households at any point along the black dashed line in figure 3.6. If a household is unemployed in period s they will therefore have $b_{s+1} = 0$, regardless of their history prior to period s. Any household moving from unemployment to employment therefore enters the bond market with zero starting wealth.

The steady states of the model are plotted below, for a calibration based on that in Ravn and Sterk (2020) appendix A.1.2[32]. The steady state Phillips Curve and Euler Equation (with either full information or rational inattention) pin down the steady state as they both contain just two endogenous variables: inflation and the hiring rate.

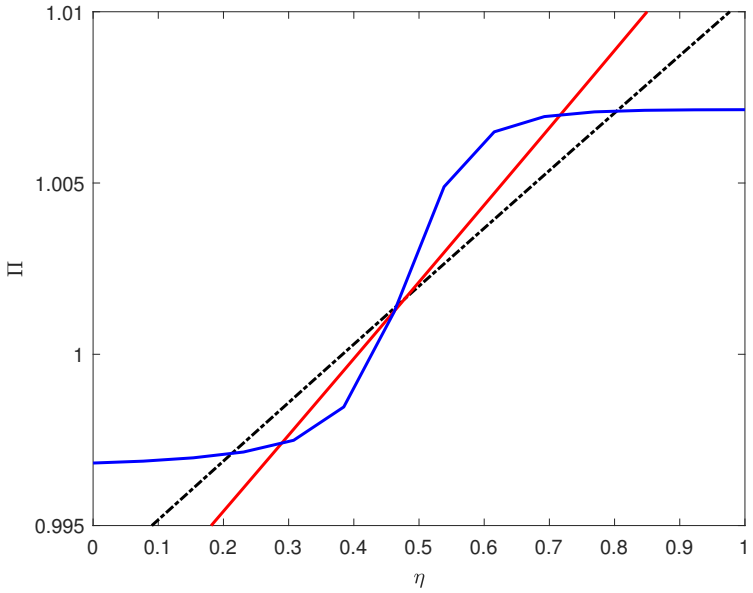


Figure 3.6: Steady state relations under perfect information and rational inattention ($\psi = 0.0025$)

The black (dashed) and red (solid) lines are the Phillips Curve and full information Euler Equation. They are identical to figure 2 panel 1 in RS, except we have not plotted separate relationships for regions where the zero lower bound on nominal interest rates and the non-negativity constraint on vacancies bind, as we abstract from those constraints. The blue solid line is the steady state Euler Equation with rational inattention[33]. Along the steady state Euler Equation with rational inattention we require that the distributions of wealth and prior beliefs are both stable. This does not mean that wealth or beliefs are static: each household receives signals with idiosyncratic noise, which implies churn in both wealth and beliefs underneath their stable distributions.

The Euler Equation is upward sloping because a higher steady state hiring rate η decreases the desire for precautionary saving[34]. To keep the bond market in equilibrium employed households must therefore be encouraged to save more. A higher rate of inflation leads to a higher interest rate, so a higher η is associated with more inflation in steady state. The upward slope of the Phillips Curve is a consequence of quadratic price

[32]We set this out in appendix C.3.3

[33]Details of the solution method for this curve are given in appendix C.3.5.

[34]To be precise, a higher hiring rate only decreases the precautionary savings motive when labour income risk is countercyclical, which is the case here because we assume that wages are independent of the hiring rate. If labour income risk is procyclical, the Euler equation would slope downwards.

adjustment costs (Ascari and Rossi, 2012).

Under perfect information the interaction of these curves leads to a unique steady state, with $\eta = 0.48$ and $\Pi = 1.0016$. With rational inattention there are three steady states, a high employment steady state with $\eta = 0.80$, $\Pi = 1.0071$, a middle steady state with $\eta = 0.47$, $\Pi = 1.0014$, and a low employment steady state with $\eta = 0.22$, $\Pi = 0.9973$.

These steady states share many properties with the simple model in section 3.2. The outer two steady states are locally stable with respect to belief mutations while the middle steady state is unstable. Large mutations to beliefs can however shift the economy between the steady states, and substantially larger mutations are needed to shift the economy out of the high employment steady state than are needed to escape the unemployment trap.

Just as in the simple model, underneath each steady state there is heterogeneity and churn in prior beliefs. Households are divided into optimists and pessimists based on the previous realisation of their signal about the hiring rate, and each period some households transition between the two priors. This heterogeneity and idiosyncratic variation in beliefs is consistent with the survey evidence presented in the next section.

3.4 Survey Expectations

In the Survey of Consumer Expectations, employed households are asked the following question:

> Suppose you were to lose your main job this month. What do you think is the percent chance that within the following 3 months, you will find a job that you will accept, considering the pay and type of work?

This is precisely the hiring rate studied in our models, and in the model of Ravn and Sterk (2020). In this section we show that the survey responses display several features which are present in our model. The figure below shows the histogram of responses as deviations from the mean response that month.

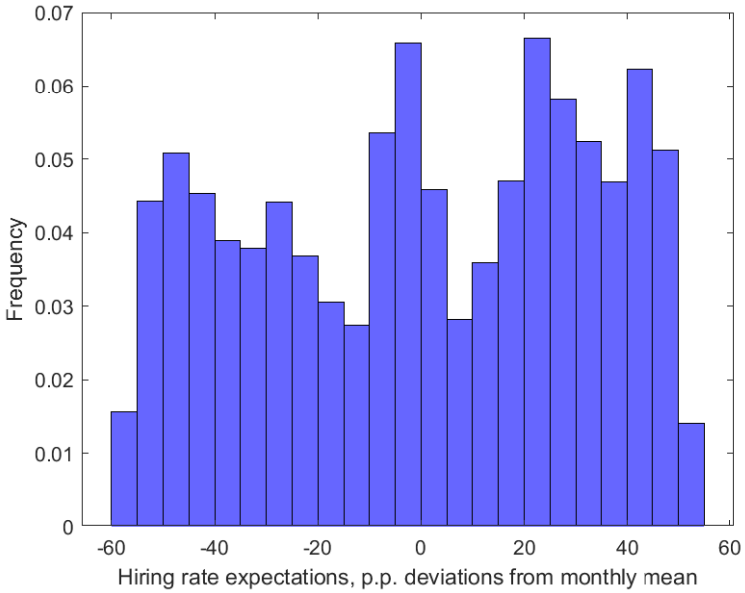


Figure 3.7: Histogram of hiring rate expectations from the Survey of Consumer Expectations, deviations from the average for the month of the interview.

There is a great deal of dispersion. In the average month the mean response is 53.5%, and the standard deviation is 32.1%. However, this disagreement could all be due to household heterogeneity, not differences in information. High and low education households, for example, could agree exactly about aggregate labour market conditions, but expect different hiring rates because they are making predictions for different segments of that labour market. To explore this, we run the following regression:

$$\mathbb{E}_{it}\eta_{it,t+3} = \alpha_0 + \alpha_1 X_{it} + \varepsilon_{it} \tag{3.21}$$

The dependent variable is household i's expectation of their own hiring rate for the months from t to $t+3$. X_{it} collects every personal characteristic available in the SCE[35]. The R^2 for this regression is 0.084. Adding time fixed effects to pick up any variation due to macroeconomic factors only increases this a small amount, to 0.094. The vast majority of heterogeneity in labour market expectations does not come from the dimensions of household heterogeneity recorded in the SCE. There are of course other dimensions of household heterogeneity which are not collected in the survey, which could explain more of the heterogeneity, but it is unlikely that this would account for all of the currently unexplained heterogeneity.

Beliefs could, however, be heterogeneous and still co-ordinate on a high (or low)

[35]The controls are age, age^2, income, income^2, education, gender, race, job tenure, financial distress, state, home ownership and number of times the household has been in the survey. Financial distress is measured as the percentage chance that the household will struggle to pay their bills in the next three months.

unemployment steady state. If all households agree in some period that the future hiring rate has fallen by 5% relative to their expectations in the previous period, that would give the same self-fulfilling expectations mechanism described in Ravn and Sterk (2020) and Heathcote and Perri (2018), even if the cross-sectional distribution of beliefs features lots of heterogeneity each period. To explore this, we use the panel nature of the SCE to study changes in the unexplained part of expectations, $\Delta\varepsilon_{it}$. If there is strong co-ordination in beliefs, then changes in household-level expectations should be largely explained by changes in average expectations. We therefore decompose individual belief updates into the average belief update that period and an idiosyncratic term, which are orthogonal to each other by construction:

$$\Delta\varepsilon_{it} = \Delta\bar{\varepsilon}_t + u_{it} \tag{3.22}$$

If households do agree on movements in the hiring rate but have different (stable) degrees of optimism or pessimism, the first differencing in this model would remove the individual-specific constants, and the variance of average belief updates ($\Delta\bar{\varepsilon}_t$) would account for most of the variance of individual belief updates ($\Delta\varepsilon_{it}$). That is, the fraction $\frac{\sigma^2_{\Delta\bar{\varepsilon}_t}}{\sigma^2_{\Delta\bar{\varepsilon}_t}+\sigma^2_{u_{it}}}$ would be very close to one. In fact, when we carry out this variance decomposition we find that less than 0.2% of the variation in revisions to household expectations is explained by movements in average expectations. There is therefore a great deal of idiosyncratic variation in *updates* to hiring rate expectations, and only weak co-ordination of beliefs. This is a natural feature of our model, because households form their expectations by observing very noisy signals.

Our model also produces a large amount of persistence in household beliefs through the updating of prior beliefs. This can be seen most clearly in table 3.1 above: even in the middle steady state, which features very dispersed expectations, individual households are extremely unlikely to revise their beliefs from one period to the next[36]. To investigate this persistence in the data, we run the following regression:

$$\varepsilon_{it} = \gamma_1\varepsilon_{it-1} + W_t + e_{it} \tag{3.23}$$

Here W_t are month fixed effects, which pick up any variation in expectations due to macroeconomic factors, including average beliefs. The results are displayed in table 3.2 below.

[36]This remains true in the HANK model. The equivalent belief transition matrices are in table C.1 in appendix C.3.6.

Table 3.2: The unexplained part of hiring rate expectations is persistent at the household level.

	(1)
	Hiring Rate Residual
L.Hiring Rate Residual	0.719***
	(0.00390)
Observations	31830
R^2	0.521

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The coefficient γ_1 is large and significant. This implies that hiring rate expectations are indeed persistent at the household level, as predicted by our model.

A final observation from this data is that the variance of expectations is an order of magnitude larger than the variance of the actual (aggregate) hiring rate, calculated using labour flows in the CPS. Estimating an AR(1) model on the hiring rate, we estimate the long run standard deviation[37] of this series at 4.7%. In contrast, the standard deviation of the unexplained part of hiring rate expectations ϵ_{it} is 30.7%. If households knew the true distribution of the hiring rate and received noisy signals (as in a standard rational inattention model a la Sims (2003)), we would expect the standard deviation of beliefs to be *lower* than the standard deviation of the data[38]. The data therefore suggests that households do not know the true equilibrium distribution of the hiring rate, and instead hold prior beliefs which are significantly more dispersed than that true distribution[39]. This supports our assumption that households hold 'ignorance priors' about the future hiring rate.

3.5 Conclusion

We have proposed a model of a self-fulfilling expectation driven unemployment trap which is driven by rational inattention. The model features households who face costs of processing information about the future hiring rate, and who do not know the true

[37]The hiring rate is found to be stationary at the 1% level. Labour flows data begins in 1990.

[38]This follows from the Law of Total Variance. If household i collects a signal s_i about the hiring rate η, they form a posterior expectation $\mathbb{E}(\eta|s_i)$. The variance of these conditional expectations across i is given by $V(\mathbb{E}(\eta|s_i)) = V(\eta) - \mathbb{E}(V(\eta|s_i))$. If prior beliefs are correct, the unconditional variance $V(\eta)$ is the true variance of the hiring rate, and so the variance of expectations must be weakly less than this true variance.

[39]This statement relies on our conditioning variables picking up the majority of expectation variation due to heterogeneity in the actual hiring rates experienced by households. Since the variance of expectations is an order of magnitude larger than that of the observed average hiring rate, we are confident that household beliefs are more dispersed than the true distribution of the labour market variables they are trying to estimate.

equilibrium distribution of that rate. The households optimally choose to process noisy signals which imply a highly nonlinear response of aggregate consumption to changes in labour market conditions, which leads to the possibility of multiple steady states. We have shown that in the HANK model of Ravn and Sterk (2020) there is a unique steady state with positive hiring if households have full information. Rational inattention generates three steady states: an unemployment trap with mild deflation and a low (but positive) job hiring rate, a middle steady state with higher employment and moderate inflation, and an 'employment trap' with very high employment and high inflation. Expectations in the model are consistent with key properties of survey expectations: labour market beliefs are heterogeneous, persistent, and display greater variance than that implied by accurate prior beliefs.

Bibliography

[1] Acharya, Sushant ; Wee, Shu Lin (2019) Rational Inattention in Hiring Decisions. *American Economic Journal: Macroeconomics*, Vol.12(1), pp.1-40

[2] Adam, Klaus (2007) Optimal Monetary Policy with Imperfect Common Knowledge, *Journal of Monetary Economics* 54.2 (2007): 267-301.

[3] Adam, Klaus (2009) Monetary Policy and Aggregate Volatility, *Journal of Monetary Economics* 56.Supp. (2009): S1-S18

[4] Adam, Klaus ; Marcet, Albert (2011) Internal rationality, imperfect market knowledge and asset prices *Journal of Economic Theory*, 2011, Vol.146(3), pp.1224-1252

[5] Adam, Klaus ; Kuang, Pei ; Marcet, Albert (2012) House Price Booms and the Current Account. *NBER Macroeconomics Annual* 2012, Vol.26(1), pp.77-122

[6] Adams, Paul ; Hunt, Stefan ; Palmer, Christopher ; Zaliauskas, Redis (2019) Testing the Effectiveness of Consumer Financial Disclosure: Experimental Evidence from Savings Accounts. *National Bureau of Economic Research Working Paper* No. 25718

[7] Adolfson, M ; Laseen, S ; Lindé, J ; Villani, M (2007) Bayesian estimation of an open economy DSGE model with incomplete pass-through, *Journal of International Economics*, Vol. 72, pages 481–511

[8] Aegon (2017) Only 8% of people speak to an adviser about financial decisions, press release. https://www.aegon.co.uk/content/ukpaw/news/only_8_of_peoplespeaktoanadviseraboutfinancialde

[9] Andrade, Philippe ; Gaballo, Gaetano ; Mengus, Eric ; Mojon, Benoît (2019) Forward guidance and heterogeneous beliefs. *American Economic Journal: Macroeconomics*, vol. 11(3), 1-29, July 2019

[10] Angeletos, George-Marios ; La'O, Jennifer (2009) Noisy Business Cycles, *National Bureau of Economic Research Working Paper* No. 14982

[11] Angeletos, George-Marios ; La'O, Jennifer (2020) Optimal Monetary Policy with Informational Frictions, *The Journal of Political Economy* 128.3 (2020): 1027-064.

[12] Armantier, Olivier ; Bruine de Bruin, Wändi ; Topa, Giorgio ; Klaauw, Wilbert ; Zafar, Basit (2015) Inflation Expectations and Behavior: Do Survey Respondents Act on Their Beliefs? *International Economic Review*, Vol. 56, No. 2, May 2015

[13] Arrow, Kenneth. J (1987) The Demand for Information and the Distribution of Income, *Probability in the Engineering and Informational Sciences*, 1987, Vol.1(1), pp.3-13

[14] Ascari, Guido ; Rossi, Lorenza (2012) Trend Inflation and Firms Price-Setting: Rotemberg Versus Calvo. *Economic Journal*, September 2012, Vol.122(563), pp.1115-1141

[15] Auclert, Adrien ; Rognlie, Matthew ; Straub, Ludwig (2020) Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model. *National Bureau of Economic Research Working Paper* No.26647

[16] Barrett, Claer (2018) What would it take for you to switch your bank account?, *Financial Times*, London. July 27, 2018

[17] Baye, M. R, Morgan, J, Scholten, P. A. (2006). Persistent Price Dispersion in Online Markets. In Jansen, Dennis (Eds.), (pp. 122-143). Northampton, MA: . Edward Elgar

[18] BBC (2005) Banks and Building Societies. http://news.bbc.co.uk/2/hi/business/2965412.stm

[19] Beaudry, Paul ; Galizia, Dana ; Portier, Franck (2017) Reconciling Hayek's and Keynes' Views of Recessions. *The Review of Economic Studies*, 2018, Vol. 85(1), pp.119-156

[20] Beaudry, Paul ; Galizia, Dana ; Portier, Franck (2020) Putting the Cycle Back into Business Cycle Analysis. *American Economic Review* 110.1 (2020): 1-47.

[21] Benhabib, Jess ; Bisin, Alberto ; Luo, Mi (2018) Wealth Distribution and Social Mobility: A Quantitative Analysis of U.S. Data *American Economic Review*, 109 (5): 1623-47.

[22] Best, James ; Dogra, Keshav (2018) Capital Management and Inequality, Manuscript

[23] Bilbiie, Florin (2008) Limited Asset Market Participation, Monetary Policy and (Inverted) Aggregate Demand Logic, *Journal of Economic Theory*, 140 (1): 162-196

[24] Bilbiie, Florin (2019a) The New Keynesian Cross, *Journal of Monetary Economics*, March 2019

[25] Bilbiie, Florin (2019b) Monetary Policy and Heterogeneity: An Analytical Framework. Manuscript

[26] Bilbiie, Florin ; Monacelli, Tommaso ; Perotti, Roberto (2013) Public Debt and Redistribution with Borrowing Constraints *Economic Journal*, February 2013, Vol.123(566), pp.F64-F98

[27] Blanchard, Olivier ; Summers, Lawrence (2019) Introduction, in *Evolution or Revolution? Rethinking Macroeconomic Policy after the Great Recession*, edited by Olivier Blanchard and Lawrence Summers. MIT Press.

[28] Branzoli, Nicola (2016) Price dispersion and consumer inattention: evidence from the market of bank accounts. *Working Papers of the Bank of Italy*, N. 1082, 2016

[29] Briggs, Joseph ; Cesarini, David ; Lindqvist, Erik ; Östling, Robert (2020) Windfall Gains and Stock Market Participation, *Journal of Financial Economics*, forthcoming

[30] Burdett, K.; Judd, K. (1983) Equilibrium Price Dispersion, *Econometrica*, 51(4), 955-969

[31] Campanale, Claudio (2007) Increasing Returns to Savings and Wealth Inequality, *Review of Economic Dynamics*, October 2007, 10 (4), 646-675.

[32] Challe, Edouard ; Ragot, Xavier (2016) Precautionary Saving over the Business Cycle. *Economic Journal*, 126 (February), 135–164

[33] Challe, Edouard ; Matheron, Julien ; Ragot, Xavier ; Rubio-Ramirez, Juan Francisco (2017) Precautionary Saving and Aggregate Demand. *Quantitative Economics* 8 (2017), 435–478

[34] Chari, V. V. ; Kehoe, Patrick J. ; McGrattan, Ellen R. (2009) New Keynesian Models: Not Yet Useful for Policy Analysis. *American Economic Journal: Macroeconomics* 1.1 (2009): 242-66.

[35] Chavaz, M. ; Slutzky, P. (2018) Paying Attention to Banks: Evidence from Offshore Deposits. Manuscript.

[36] Christiano, Lawrence J. ; Eichenbaum, Martin ; Trabandt, Mathias (2015) Understanding the Great Recession. *American Economic Journal: Macroeconomics*, 7, 110–67.

[37] Cogley, Timothy ; Colacito, Riccardo ; Sargent, Thomas (2007) Benefits from U.S. Monetary Policy Experimentation in the Days of Samuelson and Solow and Lucas *Journal of Money, Credit, and Banking*, Feb 2007, Vol.39(s1), pp.67-991

[38] Coibion, Olivier ; Gorodnichenko, Yuriy ; Hong, Gee Hee (2015) The Cyclicality of Sales, Regular and Effective Prices: Business Cycle and Policy Implications. *American Economic Review*, Vol.105(3), pp.993-1029

[39] Connon, Heather (2007) The man who made the tills ring at Moneysupermarket. *The Observer*, London, Sun 22 Jul 2007

[40] Cook, Malcolm ; Earley, Fionnuala ; Smith, Sarah ; Ketteringham, Jody (2002) Losing interest-how much can consumers save by shopping around for financial products?, FSA Occasional Paper, No. 19, October. Gilbert, R A

[41] Cooper, Russell ; John, Andrew (1988) Coordinating Coordination Failures in Keynesian Models. *The Quarterly Journal of Economics*, Vol. 103, No. 3 (Aug., 1988), pp. 441-463

[42] Cruickshank, Don (2000) Competition in UK banking : a report to the Chancellor of the Exchequer. Norwich : Stationery Office

[43] Dasgupta, Kunal ; Mondria, Jordi (2018) Inattentive importers. *Journal of International Economics*, May 2018, Vol.112, pp.150-165

[44] Debortoli, Davide ; Galí, Jordi (2018) Monetary Policy with Heterogeneous Agents: Insights from TANK models, Manuscript

[45] Diamond, Peter A. (1982) Aggregate Demand Management in Search Equilibrium. *Journal of Political Economy*, 1982, vol. 90, no. 5

[46] Diebold, F.X. ; Sharpe, S.A. (1990) Post-Deregulation Bank-Deposit-Rate Pricing: The Multivariate Dynamics, *Journal of Business and Economic Statistics*, 8(3), pp. 281-291.

[47] Driscoll, J. C. ; Judson, R. A. (2013) Sticky Deposit Rates: Data and Implications for Models of Price Adjustment, *Federal Reserve Board, Finance and Economics Discussion Series working paper*

[48] Drechsler, Itamar ; Savov, Alexi ; Schnabl, Philipp (2017) The Deposits Channel of Monetary Policy. *The Quarterly Journal of Economics*, Vol. 132(4), pp.1819-1876

[49] Dupor, William (2017) How the U.S. Debt-to-GDP Ratio Has Changed, St. Louis Fed On the Economy Blog, January 2017. https://www.stlouisfed.org/on-the-economy/2017/january/how-us-debt-gdp-ratio-changed

[50] Eggertsson, Gauti ; Juelsrud, Ragnar ; Summers, Lawrence H. ; Wold, Ella G. (2019) Negative Nominal Interest Rates and the Bank Lending Channel. *National Bureau of Economic Research Working Paper* No.25416

[51] Eusepi, Stefano ; Preston, Bruce (2011) Expectations, Learning, and Business Cycle Fluctuations. *The American Economic Review*, Oct 2011, Vol.101(6), pp.2844-2872

[52] Eusepi, Stefano ; Preston, Bruce (2018) The Science of Monetary Policy: An Imperfect Knowledge Perspective, *Journal of Economic Literature* 56.1 (2018): 3-59

[53] Fagereng, Andreas ; Guiso, Luigi ; Malacrino, Davide ; Pistaferri, Luigi (2019) Heterogeneity and Persistence in Returns to Wealth, *Econometrica*, forthcoming

[54] Financial Conduct Authority (2015) Cash savings market study report

[55] Fisher, Jonas D.M. (2015) On the Structural Interpretation of the Smets–Wouters "Risk Premium" Shock. *Journal Of Money, Credit, And Banking* 47.2/3 (2015): Pp511-516.

[56] Financial Services Authority (2000) Better informed consumers: Assessing the implications for consumer education of research by BMRB

[57] Fox, Craig R. ; Clemen, Robert T. (2005) Subjective Probability Assessment in Decision Analysis: Partition Dependence and Bias toward the Ignorance Prior *Management Science*, 1 September 2005, Vol.51(9), pp.1417-1432

[58] Galí, Jordi (2008) Monetary policy, inflation, and the business cycle : an introduction to the new Keynesian framework. Princeton, N.J. : Princeton University Press

[59] Hall, Robert E. (2005) Employment Efficiency and Sticky Wages: Evidence from Flows in the Labor Market. *Review of Economics and Statistics* 87.3 (2005): 397-407

[60] Harrison, Richard ; Oomen, Özlem (2010) Evaluating and estimating a DSGE model for the United Kingdom, *Bank of England Quarterly Bulletin*, First Quarter 2010, Vol.50(1), p.52

[61] Heathcote, J., and Perri, F. (2018) Wealth and Volatility. *The Review of Economic Studies*, 2018, Vol. 85(4), pp.2173-2213

[62] Hohler, Emily (2007) My first million: Simon Nixon of Moneysupermarket. *MoneyWeek*, London. https://moneyweek.com/31394/my-first-million-simon-nixon-of-moneysupermarket/

[63] Holston, Kathryn ; Laubach, Thomas ; Williams, John C. (2017) Measuring the natural rate of interest: International trends and determinants, *Journal of International Economics*, May 2017, Vol.108, pp.S59-S75

[64] Jung, Junehyuk ; Kim, Jeong-Ho ; Matějka, Filip ; Sims, Christopher A (2015) Discrete behavior in information-constrained tracking problems and portfolio choice. *Review of Economic Studies* (2019) 01, 1–28

[65] Kandori, Michihiro ; Mailath, George ; Rob, Rafael (1993) Learning, Mutation, and Long Run Equilibria in Games. *Econometrica* (1986-1998) 61.1 (1993): 29-56.

[66] Kaplan, G. ; Menzio, G. (2016). Shopping Externalities and Self-Fulfilling Unemployment Fluctuations. *The Journal of Political Economy*, 124(3), 771-825.

[67] Kaplan, Greg ; Moll, Benjamin ; Violante, Giovanni L. (2018) Monetary policy according to HANK, *American Economic Review*, 2018, Vol.108(3), pp.697-743

[68] Kaplan, Greg ; Violante, Giovanni L. (2018) Microeconomic Heterogeneity and Macroeconomic Shocks, *Journal of Economic Perspectives*, Vol.32(3), pp.167-194

[69] Kelly, Jemima (2019) How Monzo is banking on customer apathy, *Financial Times Alphaville*, London. April 23, 2019

[70] Kreps, David (1998) Anticipated Utility and Dynamic Choice. 1997 Schwartz Lecture, in *Frontiers of Research in Economic Theory*, edited by Donald P. Jacobs, Ehud Kalai, and Morton Kamien, Cambridge University Press, Cambridge, England.

[71] Lei, Xiaowen (2019) Information and Inequality, *Journal of Economic Theory* 184 (2019)

[72] Luo, Yulei (2008) Consumption Dynamics under Information Processing Constraints. *Review of Economic Dynamics*, 2008, Vol.11(2), pp.366-385

[73] Luo, Yulei; Nie, Jun; Wang, Gaowang; Young, Eric R. (2017) Rational Inattention and Dynamics of Consumption and Wealth in General Equilibrium *Journal of Economic Theory*, November 2017, Vol.172, pp.55-87

[74] Lusardi, Annamaria; Michaud, Pierre-Carl ; Mitchell, Olivia S (2017) Optimal financial knowledge and wealth inequality, *Journal of Political Economy*, 2017, 125 (2), 431-477.

[75] Maćkowiak, Bartosz ; Wiederholt, Mirko (2009) Optimal Sticky Prices under Rational Inattention, *American Economic Review*, 99, 769-803

[76] Maćkowiak, Bartosz ; Wiederholt, Mirko (2015) Business Cycle Dynamics under Rational Inattention *Review of Economic Studies*, 82, 1502-1532.

[77] Martín-Oliver, Alfredo ; Salas-Fumás, Vicente ; Saurina, Jesús (2009) Informational differentiation, interest rate dispersion and market power. *Applied Economics Letters*, 30 October 2009, Vol.16(16), pp.1645-1649

[78] Martinez-Miera, David ; Repullo, Rafael (2017) Search for Yield. *Econometrica*, March 2017, Vol.85(2), pp.351-378

[79] Matějka, Filip ; McKay, Alisdair (2015) Rational inattention to discrete choices: A new foundation for the multinomial logit model, *American Economic Review*, 1 January 2015, Vol.105(1), pp.272-298

[80] Matějka, Filip (2017) Rationally Inattentive Seller: Sales and Discrete Pricing. *Review Of Economic Studies* Volume 83 Issue 3 Page 1125

[81] Matějka, Filip ; Steiner, Jakub ; Stewart, Colin (2017) Rational Inattention Dynamics: Inertia and Delay in Decision Making, *Econometrica*, March 2017, Vol.85(2), pp.521-553

[82] Maundrell, Hannah (2017), quoted in The Independent newspaper: Bank or Building Society? Published 15/02/2017. https://www.independent.co.uk/money/bank-or-building-society-a7581421.html

[83] McKay, Alisdair (2013) Search for financial returns and social security privatization, *Review of Economic Dynamics*, 4/2013, Vol.16(2), pp.253-270

[84] Meeks, Roland ; Monti, Francesca (2019) Heterogeneous beliefs in the Phillips curve, *Bank of England Staff Working Paper* No 807

[85] Menzio, Guido ; Trachter, Nicholas (2015) Equilibrium Price Dispersion with Sequential Search, *Journal of Economic Theory* 160.C (2015): 188-215

[86] Monacelli, Tommaso ; Perotti, Roberto (2011) Redistribution and the Multiplier, *IMF Economic Review*, 2011, Vol.59(4), pp.630-651

[87] Moneyfacts (2019) Company Overview, Moneyfacts Facebook Page, accessed 28/05/2019

[88] Morris, Stephen ; Shin, Hyun Song (2000) Rethinking multiple equilibria in macroeconomic modeling. *NBER Macroeconomics Annual* 2000, Vol.15(1), pp.118-161

[89] Office of Fair Trading (2008) Current account research, Annexe D of *Personal current accounts in the UK – an OFT market study*

[90] Paciello, Luigi ; Wiederholt, Mirko (2014) Exogenous Information, Endogenous Information, and Optimal Monetary Policy, *Review of Economic Studies* 81.1 (2014): 356-388

[91] Ravn, Morten O. ; Schmitt-Grohe, Stephanie ; Uribe, Martin (2006) Deep Habits, *Review of Economic Studies*, Vol.73(1)

[92] Ravn, Morten O. ; Sterk, Vincent (2020) Macroeconomic Fluctuations with HANK & SAM: An Analytical Approach. *Journal of the European Economic Association*, forthcoming

[93] Reis, Ricardo (2006a) Inattentive Consumers, *Journal of Monetary Economics* 53.8 (2006): 1761-800

[94] Reis, Ricardo (2006b) Inattentive Producers, *Review of Economic Studies* 73.3 (2006): 793-821

[95] Rotemberg, Julio J. (1982) Sticky Prices in the United States. *Journal of Political Economy*, 1 December 1982, Vol.90(6), pp.1187-1211

[96] Schmitt-Grohe, Stephanie ; Uribe, Martin (2003) Closing Small Open Economy Models, *Journal of International Economics*, 2003, Vol.61(1), pp.163-185

[97] Sims, Christopher A. (2003) Implications of Rational Inattention. *Journal of Monetary Economics*, 2003, Vol.50(3), pp.665-690

[98] Sims, Christopher A. (2006) Rational inattention: Beyond the linear-quadratic case. *American Economic Review*, 2006, Vol.96(2), pp.158-163

[99] Smets, Frank ; Wouters, Rafael (2007) Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review*, Vol.97(3), pp.586-606

[100] Stevens, Luminita (2019) Coarse pricing policies, *Review of Economic Studies*, 2019, Vol. 87(1), pp.420-453

[101] Stigler, G. (1961) The Economics of Information, *Journal of Political Economy*, 69(3), 213-225.

[102] Vickers, J. (2011). Independent Commission on Banking : Final report : Recommendations. London: TSO.

[103] Wiederholt, Mirko (2017) Empirical Properties of Inflation Expectations and the Zero Lower Bound, Manuscript

[104] Yankov, Vladimir (2018) In Search of a Risk-free Asset. Manuscript.

APPENDIX

Appendix to Chapter 1

A.1 Alternative Mechanisms

A.1.1 The core mechanism in alternative models

Here I show that the main mechanism of the inattention model of section 1.2 is also present in a broad class of models in which households can pay a cost to increase the interest rate they face. This includes a model with frictional search for savings products, as in McKay (2013). For simplicity I assume an exogenously fixed distribution of interest rates.

Consider an infinitely lived household who chooses consumption and saving each period to maximise expected lifetime utility subject to a standard budget constraint, where income comes from an endowment y_t and asset income. Households can choose in period t to pay a cost to increase the interest rate they face i_t^e. That is, to achieve i_t^e they must pay a cost $C(i_t^e)$, where C is an increasing convex function. I will consider two specifications for this cost, one in which the cost is an additively separable cost in the utility function, and another in which it is a monetary cost entering the budget constraint. The utility cost specification could be thought of as time or effort spent searching for products, while the monetary cost would be paying an advisor or intermediary to search on their behalf. The specification in use is determined by the binary variable ϕ: when $\phi = 0$ the cost is a utility cost, when $\phi = 1$ we are studying the monetary cost specification.

$$\max_{c_t, b_t, i_t^e} \mathbb{E} \sum_{t=0}^{\infty} \beta^t \left[u(c_t) - (1-\phi) C(i_t^e) \right] \tag{A.1}$$

subject to

$$c_t + b_t + \phi C(i_t^e) = y_t + b_{t-1}(1 + i_{t-1}^e) \tag{A.2}$$

We obtain a familiar consumption Euler equation, and a first order condition on i_t^e:

$$u'(c_t) = \beta(1 + i_t^e)\mathbb{E}_t u'(c_{t+1}) \tag{A.3}$$

$$\beta b_t \mathbb{E}_t u'(c_{t+1}) = (1-\phi)C'(i_t^e) + \phi u'(c_t)C'(i_t^e) \tag{A.4}$$

The household problem in section 1.2 is a special case of this problem. The household equates the marginal utility of higher asset income with the marginal cost of achieving such a rise in interest rates. With a diminishing marginal utility of consumption, when expected future consumption falls the marginal utility of higher interest rates rises. If $\phi = 0$ households will respond by paying to increase their interest rate, since C is convex. If $\phi = 1$, households will only pay to increase i_t^e (and so $C'(i_t^e)$) if expected future consumption has fallen *relative to* current consumption, as increasing future asset income is achieved by sacrificing current consumption.

After a persistent contractionary shock, expected future consumption will fall, so households will pay to increase their interest rate[1], which will cause current consumption to fall further through the consumption Euler equation, amplifying the shock. This is the mechanism explored in section 1.2: the rational inattention problem is a tractable way to motivate and model the cost $C(i_t^e)$ as a utility cost, and allows for the distribution of interest rates available to be endogenised as a bank pricing equilibrium. It is not, however, the only way to do this. I now show that a model with frictional search for banks also fits into this class of models.

Suppose that the household is made up of many individuals. Many banks offer savings products, with interest rates that are distributed according to some CDF $F(i)$. Individuals can only choose a bank for their saving if they have observed its interest rate. All individuals observe one bank drawn at random from F, then with probability ψ they observe a second bank (again drawn at random) before choosing where to place their savings. The meeting rate ψ is an increasing function of the search effort of the individual, denoted e, which is decided by the household.

If an individual observes the interest rates of two banks, they choose the bank offering the higher interest rate, so the interest rate chosen has distribution $(F(i))^2$. The expected interest rate for an individual before we know how many banks they will observe, that is the effective interest rate faced by the household overall, is therefore:

$$i_t^e = (1-\psi(e_t)) \int i f(i) di + 2\psi(e_t) \int i f(i) F(i) di \tag{A.5}$$

This is increasing in the probability of seeing a second bank $\psi(e_t)$, as the expected maximum of two draws from a distribution must be (weakly) greater than the expectation of a single draw. We can rearrange this to express search effort in terms of the interest

[1] In the monetary cost specification households will only increase their interest rate if future consumption is expected to fall by more than current consumption. In many business cycle models, including that in section 1.5, internal persistence gives rise to 'hump-shaped' dynamics after shocks, which imply that households would pay to increase rates after a contractionary shock in both cost specifications.

rate the household ends up facing:

$$e_t = \psi^{-1}\left(\frac{i_t^e - \int i f(i) di}{2\int i f(i) F(i) di - \int i f(i) di}\right) \tag{A.6}$$

The fraction inside the inverse ψ function increases linearly in i_t^e. If there are diminishing returns to effort (ψ is concave) then effort will be a convex function of the desired interest rate. If we think of effort as being (psychologically) costly in its own right, or because it uses up valuable time, then the costs of increasing i_t^e will be a direct cost in the household utility function. As long as there are weakly diminishing returns to effort, and the cost of effort is weakly convex in effort, and at least one of those two curvatures is strict, then we obtain the first specification discussed above: there is a direct cost in utility which is convex in the desired (chosen) level of the interest rate. Formally, if the cost of effort in the utility function is $C_e(e)$, then we have:

$$C(i_t^e) = C_e\left(\psi^{-1}\left(\frac{i_t^e - \int i f(i) di}{2\int i f(i) F(i) di - \int i f(i) di}\right)\right) \tag{A.7}$$

$$C''(i_t^e) > 0 \text{ if } C_e''(i_t^e) \geq 0 \text{ and } \psi''(e_t) \leq 0, \text{ one inequality strict} \tag{A.8}$$

A.1.2 Loans and misallocation

Here I discuss two alternative channels through which attention to financial product choice could affect the business cycle: attention to loan choice and misallocation of credit. I argue that they are potentially less powerful than the consumption channel of attention to savings that I study in the main body of the paper.

If attention to both saving and loan choices rises in contractions, then savers will face higher interest rates and so reduce their consumption (the main channel studied in this paper), but borrowers will on average find out about lower interest rate loans, and so will have an incentive to increase their consumption. Attention to loans may therefore counteract the savings channel, but there are two reasons to expect that attention to loans does not operate in the same way, and is less powerful than attention to savings.

Firstly, the most significant debt for the majority of indebted households is a mortgage, and the evidence from the FCA (2019) suggests that there is strong price competition leading to only limited interest rate dispersion in mortgages. The scope for attention to drive interest rate changes is therefore small, and indeed one reason why this might be the case is that the large sums of money involved lead almost all mortgagors to pay a large amount of attention to their choice of product whatever the state of the economy.

Secondly, it is not clear that attention to loan choice will in fact rise in contractions.

For savings, I find that the marginal utility of income is very important in determining the extent of attention, and for savers the marginal utility of income is high in (demand-driven) contractions for two reasons: labour income and asset income are both low, as wages and interest rates are low. In contrast, in such a contraction a debtor sees their labour income fall, but the decline in interest rates leads to lower debt repayments, and so to a greater disposable income. It is not therefore clear that attention to loan choice will rise in contractions: for the most indebted households a fall in interest rates will increase disposable income so much that the marginal utility of income could even fall. Unfortunately, the Moneyfacts data on loans is not suitable for an empirical examination of attention to credit products, as the products tend to be very complicated. The equivalent Bank of England data on quoted household interest rates therefore averages over a set of products with substantially different characteristics, and so the comparison of this with the Moneyfacts panel does not accurately reflect search or attention behaviour.

The second alternative mechanism relates to what banks do with deposits. If higher interest rates reflect more productive investment opportunities for the bank, then as households pay more attention to their savings choices in recessions there will be a reduction in loan misallocation, dampening the output effects of the contraction. This is a very interesting potential mechanism, and I leave detailed study of it to future research.

In the specific case of retail savings, however, it is unlikely that this channel has much effect. Particularly in the UK, retail banks take in deposits in order to fund residential lending. Since there is very little interest rate dispersion in mortgages, it is unlikely that banks offering higher deposit rates are doing so because they have access to more profitable lending opportunities.

This argument does not apply to other forms of saving, such as saving in equities. Greater attention to equity choices in recessions should indeed lead to lower misallocation, which would mitigate the amplification of the business cycle that I find through the consumption channel. Note, however, that less than 20% of equities in the UK are owned by UK individuals[2] (ONS 2020), and even that figure masks the fact that many of those individuals hold their equities through managed funds and other institutions. This means that an increase in attention by households can only have a small effect on misallocation, as the majority of equity investment decisions are controlled by professional investors, who should spend all of their time paying as much attention to their choices as possible.

[2]The ONS does not distinguish between foreign individual and foreign institutional investors, but even with these included it is clear that the majority of equities are not under the direct control of households.

A.2 Proofs

A.2.1 The household FOCs are sufficient for utility maximisation

Here I prove that the household first order conditions are sufficient for utility maximisation in the simple model (section 1.2), and in the quantitative model (section 1.5). First, write the household problem as an unconstrained maximisation by substituting out for consumption using the budget constraint:

$$\max_{b_t, i_t^e, X_t} U = \sum_t \beta^t \left(u\left(\frac{b_{t-1}}{\Pi_t}(1 + i_{t-1}^e) + y_t(X_t) - b_t \right) - \mu \mathcal{I}(i_t^e) + v(X_t) \right) \tag{A.9}$$

Here I have summarised all choice variables other than saving b_t and the effective interest rate i_t^e in the vector X_t. In the simple model there are no other choice variables, so X_t is empty and non-asset income y_t is exogenous. In the quantitative model X_t includes wage setting, investment in capital and foreign bonds, capital utilisation, and money holdings. Inflation erodes real bond holdings as in the quantitative model. Since it does not feature in the simple model, this proof corresponds to that model if Π_t is set to 1 for all t.

I begin by defining H_s as the Hessian matrix of second-order partial derivatives of this utility function with respect to each choice variable that would result if there was no information friction, and so i_t^e was not a choice variable. The Hessian matrix for the full problem is then:

$$H = \begin{bmatrix} & & & 0 \\ & H_s & & \vdots \\ & & & 0 \\ & & & \frac{\partial^2 U}{\partial b_t \partial i_t^e} \\ 0 & \dots & 0 & \frac{\partial^2 U}{\partial b_t \partial i_t^e} \quad \frac{\partial^2 U}{\partial i_t^{e\,2}} \end{bmatrix} \tag{A.10}$$

Here I have used the fact that the only choice variable that i_t^e interacts with in the utility function is b_t. For all other choice variables X_t, $\frac{\partial^2 U}{\partial X_t \partial i_t^e} = 0$. The first order conditions are sufficient for utility maximisation if U is weakly concave, which is true if for any vector x:

$$xHx' = x_s H_s x_s' + 2yz \frac{\partial^2 U}{\partial b_t \partial i_t^e} + z^2 \frac{\partial^2 U}{\partial i_t^{e\,2}} \leq 0 \tag{A.11}$$

Where $x_s = [x_1, ..., y]$ and $x = [x_s, z]$. If households cannot influence effective interest rates the utility function is concave, as then this is a standard household maximisation problem (identical to that in Harrison and Oomen (2010) in the quantitative model). This implies that $x_s H_s x_s' < 0$.

Assuming a diminishing marginal utility of consumption we have that:

$$\frac{\partial^2 U}{\partial b_t^2} = u''(c_t) + \beta \mathbb{E}_t \frac{u''(c_{t+1})(1+i_t^e)^2}{\Pi_{t+1}^2} < 0 \tag{A.12}$$

It is therefore sufficient for the concavity of U to show that for any y, z:

$$y^2 \frac{\partial^2 U}{\partial b_t^2} + 2yz \frac{\partial^2 U}{\partial b_t \partial i_t^e} + z^2 \frac{\partial^2 U}{\partial i_t^{e^2}} \leq 0 \tag{A.13}$$

Using the definition of U this condition becomes:

$$\begin{aligned} y^2 u''(c_t) + y^2 \beta \mathbb{E}_t u''(c_{t+1}) \frac{(1+i_t^e)^2}{\Pi_{t+1}^2} + 2yz\beta \mathbb{E}_t u''(c_{t+1}) \frac{(1+i_t^e) b_t}{\Pi_{t+1}^2} \\ + 2yz\beta \mathbb{E}_t u'(c_{t+1}) \frac{1}{\Pi_{t+1}} - z^2 \mu \mathcal{I}''(i_t^e) + z^2 \beta \mathbb{E}_t u''(c_{t+1}) \frac{b_t^2}{\Pi_{t+1}^2} \leq 0 \end{aligned} \tag{A.14}$$

The two terms that don't depend on c_{t+1} are both negative by definition. Assuming CRRA utility, so $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$, the remaining terms can be written as:

$$\begin{aligned} \beta \mathbb{E}_t \frac{u''(c_{t+1}) b_t^2}{\Pi_{t+1}^2} \left(\frac{y^2 (1+i_t^e)^2}{b_t^2} + z^2 - 2yz \left(\frac{c_{t+1} \Pi_{t+1} b_t - \gamma (1+i_t^e)}{b_t \gamma} \right) \right) \\ = \beta \mathbb{E}_t \frac{u''(c_{t+1}) b_t^2}{\Pi_{t+1}^2} \left(\frac{y^2 c_{t+1} \Pi_{t+1}}{b_t \gamma^2} \left(2\gamma (1+i_t^e) - c_{t+1} \Pi_{t+1} b_t \right) \right. \\ \left. + \left(z - y \left(\frac{c_{t+1} \Pi_{t+1} b_t - \gamma (1+i_t^e)}{b_t \gamma} \right) \right)^2 \right) \end{aligned} \tag{A.15}$$

Since $u''(c_{t+1}) < 0$, U is concave if the terms inside the brackets are positive. The final term is positive by definition. Using the functional form for utility, a sufficient condition for U to be concave is therefore that:

$$-\frac{\beta y^2 b_t}{\gamma} \mathbb{E}_t \frac{c_{t+1}^{-\gamma}}{\Pi_{t+1}} \left(2\gamma (1+i_t^e) - c_{t+1} \Pi_{t+1} b_t \right) \leq 0 \tag{A.16}$$

Therefore the first order conditions are sufficient for utility maximisation as long as consumption, inflation and savings are not too large relative to the coefficient of risk aversion and the effective interest rate.

The qualitative results in section 1.2 hold as long as this condition is satisfied. In the quantitative model this is easily the case for plausible parameterisations. There $b_t = 1$, and γ is estimated to be significantly greater[3] than 1. Since steady state consumption

[3]It is the inverse of σ^c in table A.8, so the posterior mode for γ is 1.44

and inflation are 0.662 and 1, and steady state $i_t^e = 1/\beta - 1 = 0.003$, in the region of the steady state this condition is comfortably satisfied[4]. Consumption and inflation would have to be implausibly high, and interest rates implausibly low, to violate this condition, and indeed the estimation never suggests we approach such a region. The condition for the first order conditions to be sufficient for utility maximisation is therefore weak.

A.2.2 Equilibrium interest rates rise when attention increases

Here I show that when attention rises, the interest rate distribution shifts up, and rate dispersion falls, just as it does in models based on Burdett and Judd (1983).

First, differentiate the first order condition for bank n (equation 1.15) with respect to λ_t, denoting $\mathcal{S}_t^n = \frac{\exp(i_t^n/\lambda_t)}{\sum_{k=1}^{N}\exp(i_t^k/\lambda_t)}$ as the market share of bank n in period t, and $d_t^n = i_t^{CB} - i_t^n - \chi_t^n$ as the profit bank n makes per bond sold:

$$-d_t^n \frac{d\mathcal{S}_t^n}{d\lambda_t} - (1-\mathcal{S}_t^n)\frac{di_t^n}{d\lambda_t} = 1 \tag{A.17}$$

Using the definition of $\mathcal{S}_t^n$:

$$\frac{d\mathcal{S}_t^n}{d\lambda_t} = \frac{\mathcal{S}_t^n(1-\mathcal{S}_t^n)}{\lambda_t}\frac{di_t^n}{d\lambda_t} - \frac{\mathcal{S}_t^n(1-\mathcal{S}_t^n)i_t^n}{\lambda_t^2} + \mathcal{S}_t^n\left(\sum_{j\neq n}\frac{\mathcal{S}_t^j}{\lambda_t^2}(i_t^j - \lambda_t\frac{di_t^j}{d\lambda_t})\right) \tag{A.18}$$

Substituting this in to equation A.17 and rearranging we obtain:

$$\frac{di_t^n}{d\lambda_t} = \frac{1}{\lambda_t(1-\mathcal{S}_t^n)(\lambda_t + d_t^n\mathcal{S}_t^n)}\left[i_t^n d_t^n \mathcal{S}_t^n(1-\mathcal{S}_t^n) - \lambda_t^2 - d_t^n\mathcal{S}_t^n\big(\sum_{j\neq n}\mathcal{S}_t^j(i_t^j - \lambda_t\frac{di_t^j}{d\lambda_t})\big)\right] \tag{A.19}$$

Finally, from equation 1.15 we can write $\lambda_t = d_t^n(1-\mathcal{S}_t^n)$. Using this to substitute out for d_t^n we obtain:

$$\frac{di_t^n}{d\lambda_t} = \frac{1}{\lambda_t}\left[i_t^n\mathcal{S}_t^n - \lambda_t - \frac{\mathcal{S}_t^n}{1-\mathcal{S}_t^n}\big(\sum_{j\neq n}\mathcal{S}_t^j(i_t^j - \lambda_t\frac{di_t^j}{d\lambda_t})\big)\right] \tag{A.20}$$

The first term inside the square brackets is positive, leading higher attention (lower λ_t) to imply lower interest rates. This comes about because banks with high interest rates see their market share rise when attention rises if all interest rates are constant. This provides an incentive for those banks to decrease interest rates, accept a lower market share and make more profit per bond. For banks with lower interest rates (higher costs) this effect is smaller, and may be outweighed by the second term, which is negative.

[4]With c_{t+1} and Π_{t+1} at steady state and γ at its posterior mode in the estimation then the term in brackets is positive for all $i_t^e > -0.770$.

There is also an indirect effect on the profit maximising interest rate coming from the behaviour of other banks in the market, which is summarised in the final term of equation A.20. Firstly, if there are competitor banks with higher interest rates than bank n, an increase in attention will cause bank n's market share to drop, encouraging an increase in interest rates to offset this. Secondly, if other banks increase their interest rates when attention rises then $\frac{di_t^j}{d\lambda_t} < 0$, which pushes $\frac{di_t^n}{d\lambda_t}$ down: interest rates are strategic complements. As competitors raise their interest rates bank n becomes less competitive and starts to lose market share, so raises their own interest rate to compete.

For this reason the interest rate rises even at the lowest cost bank in the market for plausible values of attention[5]. This bank gains market share when attention rises, pushing them to cut their interest rate and increase profit per bond, but this is more than outweighed by the incentive to raise rates along with their higher cost competitors. Banks with greater market share and higher interest rates relative to their competitors do however increase their rates by less than those with small shares, which implies that a rise in attention leads to a fall in the dispersion of interest rates, just as a rise in search effort reduces price dispersion in Burdett and Judd (1983).

A.3 Persistent bank costs

A.3.1 Modelling persistent bank costs

Here I show how persistent bank costs affect equilibrium attention, interest rates, and individual choice probabilities. For simplicity, I keep to the case of $N = 2$ banks, though the intuition holds for greater numbers.

Suppose that, as in section 1.5, each period one bank is 'good' (cost χ^g) and the other is 'bad'(cost $\chi^b > \chi^g$). There are two possible states of the world: in state 1 bank 1 is good and bank 2 is bad, and in state 2 the ordering is reversed. Unlike in section 1.5, assume that there is persistence in the state. Specifically, the state of the world , denoted s_t, follows a two-state Markov process, in which $\Pr(s_{t+1} = s|s_t = s) = g$, where $g \geq 0.5$.

Assume that savers know the previous state of the world: they observe whether they chose correctly or not when the interest rate payouts occur[6]. Their choice problem in period t therefore remains a static problem. The persistence in s_t shows up as a prior

[5]With the estimated parameters in the quantitative model, all interest rates rise with attention as long as the probability of choosing the highest rate bank remains below 0.92, well above the steady state of 0.56. As the probability an individual identifies the lowest cost bank approaches 1, interest rates approach the Bertrand equilibrium: all banks set price equal to marginal cost except the lowest cost bank which sets their interest rate marginally above the next highest rate and captures all of the market.

[6]An exploration of this kind of problem without the assumption that individuals know the history of states (but with exogenous payoffs) can be found in Matějka, Steiner and Stewart (2017).

belief biased towards the previous period's realised state, which I assume without loss of generality to be state 1. I drop time subscripts to simplify notation. Savers know the bank policy functions, and so they know what interest rate each bank will set in each state of the world. They therefore face the payoff matrix:

Table A.1: Payoff matrix with observed previous state

	s_1	s_2
a_1	$i^{1,1}$	$i^{1,2}$
a_2	$i^{2,1}$	$i^{2,2}$
Prior prob.	g	$1-g$

Here a_n indicates choosing bank n, and $i^{n,s}$ is the interest rate offered by bank n in state s. This matrix is not, in general, symmetric, because bank policy functions depend on both their costs (i.e. the state of the world) and saver predispositions, so bank 1 will set different interest rates in state 1 than bank 2 would in state 2 if $g \neq 0.5$.

With a marginal cost of information of λ, the probability a saver chooses bank n in state s is as in equation 1.10:

$$P(n|i^{n,s}, i^{-n,s}, s) = \frac{\mathcal{P}_n^0 \exp(\frac{i^{n,s}}{\lambda})}{\mathcal{P}_n^0 \exp(\frac{i^{n,s}}{\lambda}) + (1-\mathcal{P}_n^0)\exp(\frac{i^{-n,s}}{\lambda})} \tag{A.21}$$

The unconditional choice probabilities (predispositions) are found as the solution to two normalization conditions:

$$\frac{\exp(\frac{i^{1,1}}{\lambda})g}{\mathcal{P}_1^0 \exp(\frac{i^{1,1}}{\lambda}) + (1-\mathcal{P}_1^0)\exp(\frac{i^{2,1}}{\lambda})} + \frac{\exp(\frac{i^{1,2}}{\lambda})(1-g)}{\mathcal{P}_1^0 \exp(\frac{i^{1,2}}{\lambda}) + (1-\mathcal{P}_1^0)\exp(\frac{i^{2,2}}{\lambda})} = 1 \tag{A.22}$$

$$\frac{\exp(\frac{i^{2,1}}{\lambda})g}{\mathcal{P}_1^0 \exp(\frac{i^{1,1}}{\lambda}) + (1-\mathcal{P}_1^0)\exp(\frac{i^{2,1}}{\lambda})} + \frac{\exp(\frac{i^{2,2}}{\lambda})(1-g)}{\mathcal{P}_1^0 \exp(\frac{i^{1,2}}{\lambda}) + (1-\mathcal{P}_1^0)\exp(\frac{i^{2,2}}{\lambda})} = 1 \tag{A.23}$$

The $\mathcal{P}_1^0$ that satisfies these conditions is:

$$\mathcal{P}_1^0 = \frac{e^{\frac{i^{21}}{\lambda}} e^{\frac{i^{22}}{\lambda}} - (1-g)e^{\frac{i^{21}}{\lambda}} e^{\frac{i^{12}}{\lambda}} - g e^{\frac{i^{11}}{\lambda}} e^{\frac{i^{22}}{\lambda}}}{e^{\frac{i^{11}}{\lambda}} e^{\frac{i^{12}}{\lambda}} - e^{\frac{i^{21}}{\lambda}} e^{\frac{i^{12}}{\lambda}} - e^{\frac{i^{11}}{\lambda}} e^{\frac{i^{22}}{\lambda}} + e^{\frac{i^{21}}{\lambda}} e^{\frac{i^{22}}{\lambda}}} \tag{A.24}$$

Since savers observe past states of the world, their priors are entirely determined by the true previous state and the transition probabilities, neither of which the banks can influence. The bank problem therefore remains static: banks choose interest rates to maximise their instantaneous expected profit, giving the same first order condition as in section 1.2.1:

$$\frac{d}{di_t^n} P(n|s_t) \cdot (i_t^{CB} - i_t^n - \chi_t^n) = P(n|s_t) \tag{A.25}$$

I assume that banks take saver predispositions as given when deciding their interest rates.

Intuitively, this is similar to the assumptions in the deep habits model of Ravn, Schmitt-Grohe and Uribe (2006), in which consumption habits evolve very slowly over time, so firms have limited ability to influence them in the short run. I take this to the extreme and assume that banks cannot influence predispositions at all in the short run. While predispositions must be consistent with interest rate policies in the long run, banks do not take this into account in their decisions. The bank first order condition is then as in section 1.2:

$$\big(1 - P(n|s_t)\big) \cdot (i_t^{CB} - i_t^n - \chi_t^n) = \lambda_t \tag{A.26}$$

The only difference is that $\Pr(n|s)$ here includes the predisposition, which comes from the prior beliefs, which are in turn driven by the persistence of bank costs.

To find equilibrium, take equation A.26 and equation A.21 for each of the four combinations of bank and state, and equation A.24 to give 9 equations in 9 variables: the four interest rates, four conditional choice probabilities, and the predisposition towards bank 1. Since this allows $\mathcal{P}_1^0$ to vary in response to interest rates, this equilibrium can be taken as the steady state of the system after predispositions have had time to adjust.

All of the results from the static cost model still hold: as attention rises dispersion falls and average rates rise. This is true both in steady state when $\mathcal{P}_1^0$ can adjust and after short-run changes when $\mathcal{P}_1^0$ is constant. Graphs showing this with some example parameters are in figures A.1 and A.2 below.

On top of those results, we have two new results[7]. First, increasing the persistence of bank costs reduces the amount of attention savers pay each period, as priors become more informative. This causes bank 1 (which is increasingly likely to be low cost) to offer lower interest rates, as savers will come to them with a high probability anyway. Conversely, bank 2 offers higher rates to try and maintain their market share.

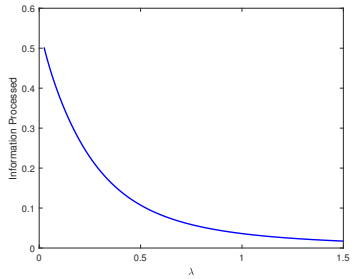
The second result is that the effective interest rate faced by the large household depends on whether the state of the world is the same as in the previous period or not. Bank 1 is more likely to be the low cost bank, so savers are predisposed to choose them. Bank 1 responds to this predisposition by offering lower interest rates. This only partially offsets the prior belief effect, so savers have $\mathcal{P}_1^0 > 0.5$ in equilibrium. This means that if the state stays the same (bank 1 is low cost), savers are more likely to correctly identify the low cost bank than they are if the state changes. This increases the effective interest rate in state 1. At the same time, interest rates at the low cost bank are lower if that low cost bank is bank 1, as they are reacting to savers predispositions. Average interest rates are therefore higher in state 2, which increases the effective interest rate in state 2 relative to state 1. Which effect dominates depends on the parameter values of the model, but in either case there will be two possible effective interest rates, and whenever

[7]Both results are obtained from a large number of simulations using different parameters.

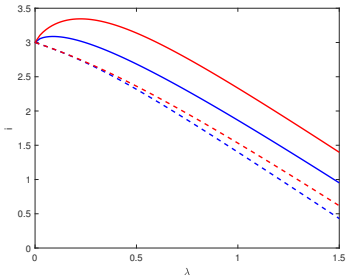
there is a transition from one state to the other the effective interest rate will change. In most calibrations the second effect dominates, so effective interest rates are higher in the period immediately after a state transition.

State transitions therefore produce i.i.d. shocks to the household effective interest rate. The shocks are i.i.d. because individuals observe the state at the end of each period. If there was a transition from $s = 1$ to $s = 2$ in period t, in period $t + 1$ prior beliefs would have $\Pr(s = 2) = g$, so the effective interest rate would only be at the transition level for one period before returning to the level associated with an unchanged state - unless there was another transition in period $t+1$. The probability of being at the transition level of i^e is therefore the probability of a state transition happening, which is the same every period. These shocks are the key qualitative difference between this model and the static cost model in sections 1.2 and 1.5.

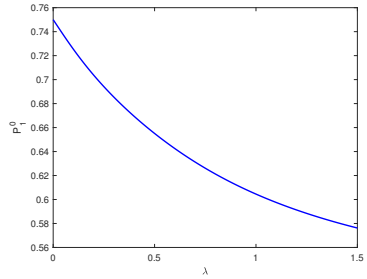
Figure A.1: Long run equilibrium varies with λ



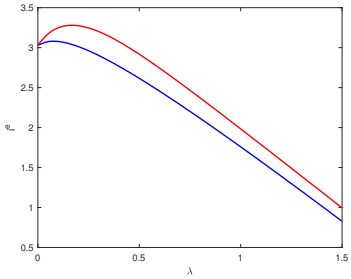
(a) Amount of information processed



(b) Equilibrium interest rates for bank 1 (blue) and bank 2 (red) when they are low (solid) and high (dashed) cost.

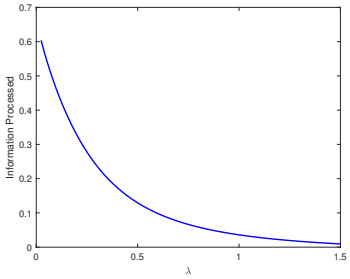


(c) Predisposition to choosing bank 1

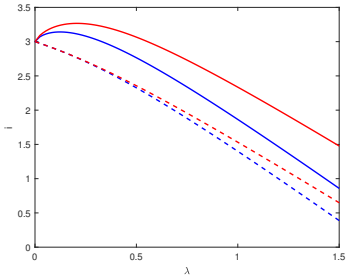


(d) Effective interest rate if the low cost bank is bank 1 (blue) and bank 2 (red).

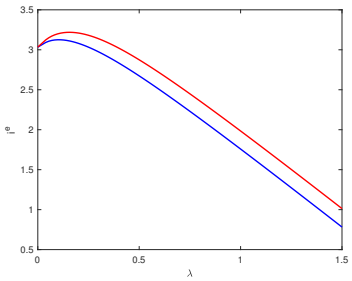
Figure A.2: Short run equilibrium varies with λ



(a) Amount of information processed



(b) Equilibrium interest rates for bank 1 (blue) and bank 2 (red) when they are low (solid) and high (dashed) cost.



(c) Effective interest rate if the low cost bank is bank 1 (blue) and bank 2 (red).

A.3.2 Persistence of interest rate rankings in the data

In sections 1.2 and 1.5 I assume that the ranking of a bank in the interest rate distribution has no persistence. Table A.2 shows the bank transition probabilities between quintiles of the interest rate distribution of the products studied in section 1.4 over a month and a year. The length of a period in section 1.5 is one month, but the annual transition probabilities are also relevant since these products have a term of one year, so individual savers buying these products only return to the decision a year later.

Without persistence, every transition probability would equal 0.2. The values on the diagonal of the transition matrices are all greater than this, so there is some persistence in the data. However, the persistence is limited, even in the top and bottom quintiles where it is strongest. If a saver chose a bank in the top quintile of the interest rate distribution in a given period, then a year later when their product matures there is only a 33% probability of that bank still being in the top quintile.

	1	2	3	4	5
1	0.77	0.15	0.04	0.02	0.02
2	0.21	0.47	0.20	0.08	0.04
3	0.04	0.27	0.44	0.19	0.07
4	0.01	0.08	0.30	0.40	0.21
5	0.01	0.03	0.09	0.22	0.65

(a) Monthly

	1	2	3	4	5
1	0.59	0.21	0.11	0.06	0.04
2	0.28	0.25	0.22	0.14	0.11
3	0.17	0.24	0.26	0.19	0.15
4	0.08	0.21	0.26	0.23	0.22
5	0.07	0.15	0.21	0.24	0.33

(b) Annual

Table A.2: Bank quintile transition matrices. In each table the cell (n, m) indicates the probability of transitioning from the nth quintile to the mth quintile in the following period.

I test if these transition matrices are significantly different from a matrix where every element is 0.2 (the no-persistence case) with a likelihood ratio test:

$$2 \ln \left(\frac{\prod_{n=1}^{5} \prod_{m=1}^{5} p_{n,m}}{\prod_{n=1}^{5} \prod_{m=1}^{5} 0.2} \right) \sim \chi^2_{19} \tag{A.27}$$

The critical value of the test statistic for 5% significance is 30.1. The monthly and annual transition matrices give test statistics of 31.2 and 7.2 respectively. We therefore cannot reject the hypothesis of no persistence at an annual frequency, and we only marginally reject that hypothesis at the monthly frequency.

A.4 Relationship between attention and φ

A.4.1 N=2 banks

Here I show that there is an exact correspondence between attention and φ in the model with $N = 2$ banks, as used in section 1.5. More attention means φ rises because individuals choose the higher interest rate in the market with a greater probability, and so the effective rate faced by the household rises relative to the distribution of rates on offer.

As in section 1.5, assume priors are uninformative and define p_t^g as the probability an individual chooses the high interest rate bank in period t:

$$p_t^g = \frac{\exp(\frac{i_t^g}{\lambda_t})}{\exp(\frac{i_t^g}{\lambda_t}) + \exp(\frac{i_t^b}{\lambda_t})} \tag{A.28}$$

Individuals paying no attention to bank choice choose bank n with probability $\mathcal{P}_n = 0.5$, so the benchmark no-attention rate in the model is the unweighted mean of the available interest rates:

$$i_t^{NA} = \mathcal{P}_1 i_t^1 + (1 - \mathcal{P}_1) i_t^2 = 0.5(i_t^1 + i_t^2) \tag{A.29}$$

In this two-bank case of the model, attention is given by:

$$\mathcal{I}_t = \log(2) + p_t^g \log p_t^g + (1 - p_t^g) \log(1 - p_t^g) \tag{A.30}$$

Attention is a monotonically increasing function of p_t^g (as $p_t^g \geq 0.5$). The empirical statistic φ is:

$$\varphi_t = \frac{p_t^g i_t^g + (1 - p_t^g) i_t^b - \frac{1}{2}(i_t^g + i_t^b)}{\frac{1}{2}(i_t^g - i_t^b)} \tag{A.31}$$

This simplifies to:

$$\varphi_t = \frac{p_t^g (i_t^g - i_t^b) - \frac{1}{2}(i_t^g - i_t^b)}{\frac{1}{2}(i_t^g - i_t^b)} = 2p_t^g - 1 \tag{A.32}$$

That is, in this simple case φ_t is a linear function of the probability an individual successfully chooses the higher interest rate bank. With no information processing, $p_t^g = 0.5$ and so $\varphi_t = 0$. If they process enough information to ensure that they always identify the high interest rate bank, $p_t^g = 1$ and $\varphi_t = 1$. The same intuition holds for the N bank case, though the relationship is no longer so precise. This is shown in appendix A.4.2.

A.4.2 N>2 banks

Here I show that φ and attention $\mathcal{I}$ are closely related in the model with N banks and uninformative priors. There are no dynamics to the relationship, so for ease I drop all time subscripts. Denoting the unweighted mean interest rate (which is again the model's no-attention rate) as $\bar{i}$, and the standard deviation of interest rates as $\sigma(i)$, the model-implied φ is:

$$\varphi = \frac{\sum_n i^n \Pr(\text{choose } n) - \bar{i}}{\sigma(i)} = \frac{\frac{\sum_n i^n \exp(\frac{i^n}{\lambda})}{\sum_m \exp(\frac{i^m}{\lambda})} - \bar{i}}{\sigma(i)} \tag{A.33}$$

First, note that as $\mathcal{I}$ approaches 0, λ tends to infinity, and so when attention is 0, $\varphi = 0$:

$$\lim_{\lambda \to \infty} \varphi = \frac{\frac{1}{N}\sum_n i^n - \bar{i}}{\sigma(i)} = 0 \tag{A.34}$$

If attention $\mathcal{I}$ reaches $\log(N)$, then each individual can perfectly identify the highest interest rate bank with probability 1, so denoting this as bank 1 (without loss of generality) we have $\varphi > 0$:

$$\varphi(\mathcal{I} = \log(N)) = \frac{i^1 - \bar{i}}{\sigma(i)} = \frac{\frac{1}{N}\sum_n (i^1 - i^n)}{\sigma(i)} > 0 \tag{A.35}$$

The information constraint is continuous for $\mathcal{I} \in (0, \log(N))$, so the statements above guarantee that $\mathcal{I}$ and φ are positively related at least in some portions of this range.

To make further progress, consider how φ changes in the model assuming that interest

rates are held fixed. The equilibrium interest rate response will be incorporated below. We use the chain rule to write:

$$\frac{d\varphi}{d\mathcal{I}} = \frac{d\varphi}{d\lambda}\frac{d\lambda}{d\mathcal{I}} \tag{A.36}$$

I start with $\frac{d\lambda}{d\mathcal{I}}$. Substituting the optimal choice probabilities into the information constraint 1.5 gives:

$$\mathcal{I} = \log(N) + \frac{i^e}{\lambda} - \log(\sum_n \exp(\frac{i^n}{\lambda})) \tag{A.37}$$

Differentiate this with respect to λ:

$$1 = \frac{d\lambda}{d\mathcal{I}}\Big[- \frac{i^e}{\lambda^2} + \frac{1}{\lambda}\frac{di^e}{d\lambda} - \frac{d}{d\lambda}\log(\sum_n \exp(\frac{i^n}{\lambda}))\Big] \tag{A.38}$$

Take the final term in the square brackets:

$$\frac{d}{d\lambda}\log(\sum_n \exp(\frac{i^n}{\lambda})) = -\frac{\sum_n i^n \exp(\frac{i^n}{\lambda})}{\lambda^2 \sum_m \exp(\frac{i^m}{\lambda})} = -\frac{i^e}{\lambda^2} \tag{A.39}$$

Substituting this back into equation A.38 the first and third terms in the square brackets cancel, giving:

$$1 = \frac{d\lambda}{d\mathcal{I}}\frac{1}{\lambda}\frac{di^e}{d\lambda} \tag{A.40}$$

The sign of $\frac{d\lambda}{d\mathcal{I}}$ is therefore the same as the sign of $\frac{di^e}{d\lambda}$:

$$\frac{di^e}{d\lambda} = \frac{\big(\sum_n i^n \exp(\frac{i^n}{\lambda})\big)^2 - \big(\sum_n \exp(i^{n^2}\frac{i^n}{\lambda})\big)\big(\sum_m \exp(\frac{i^m}{\lambda})\big)}{\lambda^2\big(\sum_n \exp(\frac{i^n}{\lambda})\big)^2} \tag{A.41}$$

The denominator of this fraction is always positive, so the sign is determined by the sign of the numerator, which after expanding the terms in brackets is:

$$\sum_n i^{n^2}\exp(\frac{2i^n}{\lambda}) + \sum_{m\neq n} i^n i^m \exp(\frac{i^n + i^m}{\lambda}) - \sum_n i^{n^2}\exp(\frac{2i^n}{\lambda}) - \sum_{m\neq n} i^{n^2}\exp(\frac{i^n + i^m}{\lambda})$$
$$= -\sum_{m\neq n}(i^{n^2} - i^n i^m)\exp(\frac{i^n + i^m}{\lambda}) \tag{A.42}$$

Inside the sum, each pair of banks $\{j,k\}$ appear twice: when $m = k, n = j$ and when $m = j, n = k$. For each distinct pair of banks $\{j,k\}$, the terms inside the sum are equal to:

$$\exp(\frac{i^j + i^k}{\lambda})(i^{j^2} - i^j i^k + i^{k^2} - i^k i^j) = \exp(\frac{i^j + i^k}{\lambda})(i^j - i^k)^2 > 0 \tag{A.43}$$

Each pair of terms inside the sum in equation A.42 is therefore positive, and so $\frac{di^e}{d\lambda}$ is

negative. That is, when the shadow cost of information in the individual problem falls, the effective interest rises, if we hold the distribution of interest rates constant.

This implies that $\frac{d\lambda}{d\mathcal{I}}$ is also negative. If attention rises, then holding the distribution of interest rates constant the shadow price of attention falls.

Now consider $\frac{d\varphi}{d\lambda}$. Since we have already shown that $\frac{di^n}{d\lambda} < 0$ we have:

$$\frac{d\varphi}{d\lambda} = \frac{1}{\sigma(i)} \frac{di^n}{d\lambda} < 0 \tag{A.44}$$

Therefore holding the distribution of interest rates constant, φ monotonically increases with attention.

To include equilibrium interest rate changes, write the full derivative of φ with respect to $\mathcal{I}$ as:

$$\frac{d\varphi}{d\mathcal{I}} = \frac{\partial \varphi}{\partial \mathcal{I}}|_{i^n} + \sum_n \frac{d\varphi}{di^n} \frac{di^n}{d\mathcal{I}} \tag{A.45}$$

We have shown that the first term is positive. As discussed in section 1.2.3 and appendix A.2, $\frac{di^n}{d\mathcal{I}} > 0$ for the range of attention encountered in simulations of the model. However, the sign of $\frac{d\varphi}{di^n}$ is ambiguous. To see why, it is helpful to examine the numerator and denominator of φ separately.

The numerator is $\sum_n \Pr(\text{choose } n) i^n - \frac{1}{N} \sum_n i^n$. This will rise when a given bank's rate i^j rises if $\Pr(\text{choose } j) > \frac{1}{N}$. For all $\mathcal{I} > 0$ this is true for the lowest cost bank, and is not true for the highest cost bank. For intermediate banks the sign of this term will depend on the rate distribution and the level of attention. The second-lowest cost bank, for example, will be chosen with a probability greater than $\frac{1}{N}$ at low and modest levels of attention, and rises in attention will initially increase this probability further. However, as attention gets very high individuals will accurately distinguish even between the top two interest rates in the market, and so the probability of this bank being chosen will fall below $\frac{1}{N}$, eventually reaching 0 when $\mathcal{I} = \log(N)$.

The derivative of the denominator of φ (i.e. $\sigma(i)$) with respect to any individual rate i^j is:

$$\frac{d\sigma(i)}{di^j} = \frac{1}{N\sigma(i)} \Big(i^j - \frac{1}{N} \sum_n i^n \Big) \tag{A.46}$$

The denominator of φ therefore rises with i^j if i^j is above the mean interest rate in the market. At moderate levels of attention, both the numerator and denominator of φ rise with the interest rates of the banks offering the highest rates, and fall when the lower rates in the market rise. The overall effect on φ of a rise in any individual rate is therefore ambiguous without further specification of the levels of attention and costs.

Simulations of the model with $N > 2$ banks suggest that these equilibrium interest

rate response terms are not sufficient to outweigh the direct effect discussed above that $\frac{\partial \varphi}{\partial \mathcal{I}}|_{i^n} > 0$, so there is a positive monotonic relationship between attention and φ in the model even outside of the simple case with $N = 2$ banks.

Note, however, that this derivation still holds the policy rate constant. The policy rate affects the standard deviation of interest rates in the quantitative model if $\tau_2 \neq 0$. This means that an increase in attention when policy rates are low will have a different effect on φ than it would if the increase happened when policy rates are high. As there is not a one-to-one mapping from policy rates to attention (attention decisions co-move differently with the policy rate with different shocks), the link from attention to φ would not be one-to-one if there were more than 2 banks in the quantitative model.

A.5 Robustness for section 1.4

A.5.1 Alternative measures of φ

Here I present two alternatives to the household choice statistic φ, which corroborate the evidence in section 1.4.3 that households move up through the distribution of interest rates when unemployment is high and the level of average rates is low.

First, I define a new variable φ_{best} in a similar way to φ, but rather than comparing the average rate achieved by households each month with the rate at the big four banks, I compare it with the highest interest rate available in the market. Intuitively, rather than comparing choices to a 'no attention' benchmark, this compares choices to a full information benchmark.

$$\varphi_{best} = \frac{\mathbb{E}_h i_t - i_t^{best}}{\sigma(i_t)} \tag{A.47}$$

Second, I define φ_{pct} to be the percentile of the interest rate distribution at which the average interest rate achieved by the household sits. This is even more model-free than φ and φ_{best}, taking no stance on the appropriate benchmark for choices. As with the previous two statistics, it is homogeneous of degree 0. The downside is that it does not consider the shape of the rate distribution either side of the average rate achieved by households.

$$\varphi_{pct} = \Pr(i_t^n < \mathbb{E}_h i_t) \tag{A.48}$$

When households are more successful at choosing the higher interest rate products in the market, φ_{best} is low and φ_{pct} is high. The pairwise correlations between each of the three statistics on household choice (φ, φ_{best}, φ_{pct}), unemployment and mean interest rates are shown in table A.3 below[8].

[8] As in section 1.4, all correlations are between the HP-filtered cyclical components of the variables.

When unemployment is high and interest rates are low, φ_{pct} and φ are high, while φ_{best} is low. All correlations are strongly significant. The two alternative measures of household choice success therefore deliver the same qualitative implications as those found in section 1.4: in contractions households move up within the distribution of interest rates, away from the low rate offered by the big four banks and towards the highest rate in the market.

Table A.3: Pairwise contemporaneous correlations

	φ	φ_{best}	φ_{pct}	U	$\bar{i}$
φ	1				
φ_{best}	-0.707^{***}	1			
φ_{pct}	0.657^{***}	-0.736^{***}	1		
U	0.413^{***}	-0.310^{***}	0.371^{***}	1	
$\bar{i}$	-0.382^{***}	0.275^{***}	-0.394^{***}	-0.785^{***}	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.5.2 Market size cannot explain fluctuations in φ_t

Drechsler, Savov and Schnabl (2017) (DSS) show that when the Federal Funds Rate rises in the US, retail banks increase their deposit spreads and deposits flow out of the retail market. Here I show that such switching out of the deposit market and into other asset types cannot explain the cyclicality of φ_t, because the proportion of households who hold fixed interest rate bonds does not vary significantly through the Great Recession.

In principle, the switching identified by DSS could drive my empirical findings. If households differ in their propensity to pay attention to savings, then it could be that when the level of interest rates rises the high-attention households switch out of the retail deposit market. The savers that remain buying fixed-rate savings bonds from banks are the low-attention households, and so the average attention of households in the market falls without any individual household changing their attention.

To explore if this switching is occurring, I study waves 1-3 (2006, 2008, 2010) of the Wealth and Assets Survey (WAS). This survey asks a large number of households about their assets, including whether they hold fixed interest rate savings bonds, and if so how large their deposits are in such products. As the three waves span the Great Recession, if the DSS switching effect is driving the cyclicality of φ_t we should find that the proportion of households who holding products studied in section 1.4 increases over the recession.

The products considered in section 1.4 were those available with an investment of £5000. In the majority of cases, this means that the minimum investment was £5000, and the maximum was £9999, as above that level the banks usually offer a different product with a potentially different interest rate. I therefore study the proportion of

households who hold fixed-rate savings bonds with balances in this range. Table A.4 shows the results from regressing a dummy variable indicating whether a household owns a fixed-rate bond of the appropriate size on the wave of the WAS they are in, with wave 1 (2006) as the baseline. The probability of holding the bond is not significantly different in each wave. This remains true if I widen the range of deposit sizes, which I do as a check since a minority of products in the sample have a minimum investment below £5000, or don't have a corresponding product for balances above £10000.

This is not inconsistent with the mechanism in DSS. Table A.5 shows that the proportion of households holding very large balances in fixed-rate bonds increased through the Great Recession (following the same method as table A.4 with larger balances). While a minority of products studied in section 1.4 would allow these higher balances, in general households could get higher interest rates by depositing larger balances in banks offering specific large-balance bonds, so if these households are high-attention types they are unlikely to be buying products in the set studied above.

Finally, if the proportion of savers in the fixed-rate market who are high-attention types increases in recessions, we should expect to see an increase through the Great Recession in the average education of those who hold fixed-rate bonds. Table A.6 shows that the year in which they participated in the survey does not significantly correlate with the education of a holder of a fixed-rate bond of the size considered in section 1.4.

Table A.4: Proportion of households holding the relevant fixed-rate bonds does not change significantly over the Great Recession

	(1)	(2)
	Hold bond £5,000-£9,999	Hold bond £2,500-£12,499
Wave=2	-0.000575	0.000431
	(-0.56)	(0.28)
Wave=3	-0.00112	-0.000229
	(-1.05)	(-0.15)
Constant	0.0129***	0.0275***
	(19.85)	(29.08)
Observations	72197	72197

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.5: Proportion of households holding larger fixed-rate bonds does change significantly over the Great Recession

	(1)	(2)
	Hold bond £25,000-£49,999	Hold bond £50,000+
Wave=2	0.00290**	0.00169
	(2.91)	(1.84)
Wave=3	0.00262**	0.00428***
	(2.67)	(4.40)
Constant	0.0114***	0.0114***
	(19.14)	(20.23)
Observations	72197	72197

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.6: Educational attainment conditional on holding a fixed-rate bond with balance between £5,000 and £9,999 does not change significantly over the Great Recession

	(1)	(2)
	Has degree level education or above	Has some educational qualification
Wave=2	0.0466	0.00341
	(1.23)	(0.11)
Wave=3	0.0216	-0.0294
	(0.54)	(-0.80)
Constant	0.269***	0.809***
	(12.00)	(40.34)
Observations	1052	1052

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.6 Estimation Details

A.6.1 Data Sources and Treatment

There are 11 standard observable variables: domestic (UK) GDP, consumption, inflation, the 3-month treasury bill rate, investment, real wages, hours worked, and foreign inflation, industrial production, interest rates, and relative export prices. The foreign variables are trade-weighted averages of the other G7 countries. On top of these I add 3 observables from the Moneyfacts data: the mean and standard deviation of deposit rates, and φ. I use data from 1993-2009.

I follow Harrison and Oomen to source and detrend the standard observables. See their paper for details of the data series. The only difference to their approach is that I use the treasury bill rate at a monthly frequency, from the same source as them. I extract a log-linear trend from all real variables, and transform inflation and interest rates into monthly rates before de-meaning[9]. For the average and standard deviation of interest rates in Moneyfacts I follow the same procedure used for the treasury bill rate. For φ I take logs and extract a linear trend, as with the other real variables. I choose τ_1 to match the average gap between the highest and the (unweighted) mean interest rate available.

A.6.2 Foreign VAR

Foreign variables are assumed to follow a VAR process estimated outside of the model, as in Adolfson et al (2007). Denoting the vector of foreign variables as Y_t, the VAR process is:

$$F_0 Y_t = F_1 Y_{t-1} + F_2 Y_{t-2} + ... + u_t \tag{A.49}$$

In Adolfson et al (2007) Y_t consists of foreign inflation, output, and interest rates. They impose that:

$$F_0 = \begin{bmatrix} 1, & 0, & 0 \\ 0, & 1, & 0 \\ -\gamma_\pi, & -\gamma_y, & 1 \end{bmatrix} \tag{A.50}$$

Output and inflation are assumed to be unaffected by contemporaneous shocks to anything other than themselves, but interest rates respond to both. The VAR is over-identified, and they cannot reject the over-identifying restrictions.

I have one extra variable, relative export prices. I begin with the Adolfson et al (2007) restrictions, and add that inflation and output also do not respond contemporaneously to shocks to relative export prices. Furthermore, I assume that the foreign interest

[9]Harrison and Oomen take separate means for each 'regime' of the data identified by Benati (2006), but the final regime begins in 1993, which is when my data starts.

rate doesn't respond contemporaneously to shocks to relative export prices, but that relative export prices can respond contemporaneously to all variables. The idea is that the exchange rate can vary rapidly in response to shocks, and this will affect the relative export price. This gives:

$$F_0 = \begin{bmatrix} 1, & 0, & 0, & 0 \\ 0, & 1, & 0, & 0 \\ -\gamma_\pi, & -\gamma_y, & 1, & 0 \\ -\gamma_\pi^p, & -\gamma_y^p, & \gamma_r^p, & 1 \end{bmatrix} \tag{A.51}$$

The model is over-identified. We cannot reject the over-identifying restrictions (p-value 0.33). The AIC suggests that 12 lags is optimal here, which fits with the quarterly models with 4 lags in both Adolfson et al and Harrison and Oomen.

A.6.3 Estimation Results

Table A.7 gives descriptions of each estimated parameter. Tables A.8 and A.9 show the estimation results for the baseline model and the full information model in section 1.5.4 respectively. The variance decomposition in the full information and inattention models discussed in section 1.5.4, is shown in figure A.3.

Table A.7: Description of estimated parameters

Parameter	Description	Prior Distribution
σ^c	Intertemporal elasticity of substitution	$N(0.66, 0.20)$
ψ^{hab}	Consumption habit parameter	Beta$(0.88, 0.05)$
σ^h	Labour supply elasticity	$N(0.43, 0.11)$
χ^k	Capital adjustment cost constant	$N(2000, 600)$
ϵ^k	Indexation to past capital adjustment in capital adjustment cost	Beta$(0.79, 0.1)$
σ^z	Capital utilization cost elasticity	$N(0.56, 0.168)$
χ^{hv}	Domestic goods price adjustment cost	$N(3268, 980.4)$
ϵ^{hv}	Domestic goods inflation indexation	Beta$(0.64, 0.1)$
χ^{xv}	Export goods price adjustment cost	$N(526.5, 12.5)$
ϵ^{xv}	Export goods inflation indexation	Beta$(0.52, 0.05)$
ψ^{pm}	Imported goods Calvo parameter	Beta$(0.80, 0.15)$
ϵ^m	Imported goods inflation indexation	Beta$(0.55, 0.05)$
ψ^w	Wage Calvo parameter	Beta$(0.74, 0.05)$
ϵ^w	Wage inflation indexation	Beta$(0.83, 0.1)$
θ^p	Taylor Rule inflation weight	$N(0.87, 0.1305)$
θ^y	Taylor Rule output weight	$N(0.11, 0.0275)$
θ^{rg}	Taylor Rule persistence	Beta$(0.95, 0.01)$
μ	Marginal cost of information	Gamma$(0.01, 0.5)$
τ_2	Elasticity of inefficient bank costs to the policy rate	$U(0, 0.5)$
$\rho_t fp$	Persistence of AR(1) TFP shock	Beta$(0.96, 0.01)$
σ_{tfp}	s.d. TFP shock	InvGamma$(0.0056, 2)$
ρ_g	Persistence of AR(1) government spending shock	Beta$(0.99, 0.01)$
σ_g	s.d. government spending shock	InvGamma$(0.0085, 2)$
ρ_x	Persistence of AR(1) shock x	$U(0.79, 0.29)$
σ_{κ^h}	s.d. labour disutility shock	InvGamma$(0.01, 2)$
σ_{rg}	s.d. monetary policy shock	InvGamma$(0.000707, 2)$
$\sigma_{\zeta^h b}$	s.d. price markup shock	InvGamma$(0.006, 2)$
σ_{ζ^k}	s.d. capital adjustment cost shock	InvGamma$(0.06, 2)$
σ_y	s.d. shock y	InvGamma$(0.025, 2)$

$x = \kappa^h, \zeta^c, \zeta^h b, \zeta^k, \mu, \tau_2$ refers to the shock to labour disutility, the risk premium, price markups, capital adjustment costs, information costs, and interest rate dispersion.
$y = \zeta^c, \mu, \tau_1, \tau_2$ is a subset of x, plus τ_1 the measurement error in mean rates.

Table A.8: Estimated parameters in baseline model

Parameter	Mode	s.d.	Parameter	Mode	s.d.
σ^c	0.694	0.002	ρ_{κ^h}	0.793	0.001
ψ^{hab}	0.944	0.002	ρ_{ζ^c}	0.778	0.001
σ^h	0.467	0.011	$\rho_{\zeta^{hb}}$	0.815	0.001
χ^k	1807.600	12.000	ρ_{ζ^k}	0.740	0.003
ϵ^k	0.826	0.001	ρ_μ	0.790	0.001
σ^z	0.580	0.002	ρ_{τ_1}	NA	NA
χ^{hv}	3477.152	9.150	ρ_{τ_2}	0.809	0.001
ϵ^{hv}	0.633	0.001	μ	0.012	0.001
χ^{xv}	526.443	0.002	τ_2	-0.139	0.001
ϵ^{xv}	0.559	0.002	σ_g	0.040	0.002
ψ^{pm}	0.816	0.001	σ_{κ^h}	0.132	0.005
ϵ^m	0.541	0.001	σ_{rg}	0.000	0.000
ψ^w	0.751	0.001	σ_{tfp}	0.005	0.001
ϵ^w	0.854	0.001	σ_{ζ^c}	0.007	0.003
θ^p	0.779	0.003	$\sigma_{\zeta^{hb}}$	0.001	0.000
θ^y	0.112	0.000	σ_{ζ^k}	0.286	0.009
θ^{rg}	0.976	0.002	σ_μ	0.066	0.002
ρ_{tfp}	0.983	0.001	σ_{τ_1}	0.475	0.015
ρ_g	0.987	0.001	σ_{τ_2}	0.003	0.005

Table A.9: Estimated parameters in full information model

Parameter	Mode	s.d.	Parameter	Mode	s.d.
σ^c	0.829	0.001	ρ_{κ^h}	0.796	0.002
ψ^{hab}	0.899	0.000	ρ_{ζ^c}	0.893	0.001
σ^h	0.434	0.001	$\rho_{\zeta^{hb}}$	0.533	0.001
χ^k	1956.000	0.800	ρ_{ζ^k}	0.592	0.002
ϵ^k	0.713	0.000	ρ_μ	NA	NA
σ^z	0.560	0.000	ρ_{τ_1}	NA	NA
χ^{hv}	3161.463	4.575	ρ_{τ_2}	NA	NA
ϵ^{hv}	0.637	0.001	μ	NA	NA
χ^{xv}	525.653	0.027	τ_2	NA	NA
ϵ^{xv}	0.529	0.000	σ_g	0.025	0.002
ψ^{pm}	0.646	0.001	σ_{κ^h}	0.053	0.004
ϵ^m	0.558	0.000	σ_{rg}	0.000	0.000
ψ^w	0.681	0.000	σ_{tfp}	0.005	0.001
ϵ^w	0.796	0.000	σ_{ζ^c}	0.004	0.001
θ^p	0.792	0.001	$\sigma_{\zeta^{hb}}$	0.001	0.000
θ^y	0.100	0.000	σ_{ζ^k}	0.359	0.02
θ^{rg}	0.957	0.000	σ_μ	NA	NA
ρ_{tfp}	0.961	0.000	σ_{τ_1}	NA	NA
ρ_g	0.984	0.000	σ_{τ_2}	NA	NA

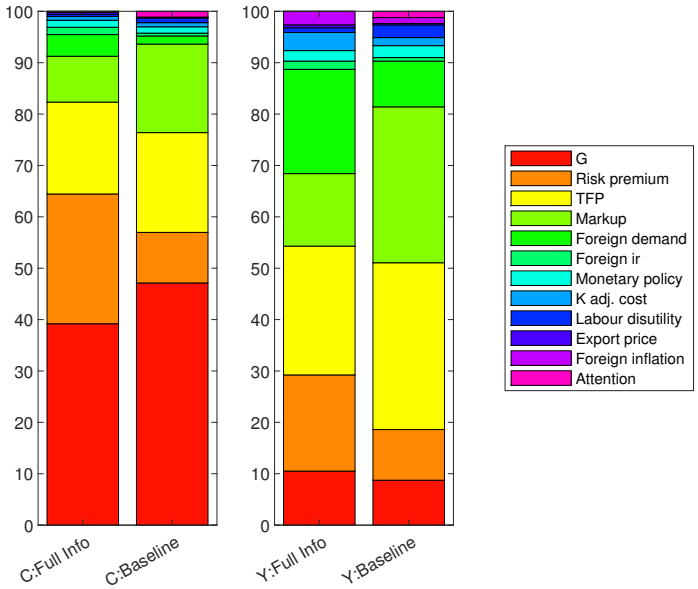


Figure A.3: Percentage of the variance of consumption and output due to each shock in the full information and variable attention models.

The risk premium shock is displayed here in orange. As described in section 1.5.4, without information frictions the risk premium shock explains 25% of the variance of consumption, more than any other shock except for government spending (39%). It explains the third largest share of output variance (19%), after TFP shocks (25%) and foreign demand shocks (20%).

In moving to the baseline model with inattention the risk premium shock becomes substantially less important, explaining 10% of both consumption and output. This is not picked up by shocks to attention, which explain 1% of the variance of both consumption and output in the baseline inattention model. Rather, TFP and price markup shocks explain greater shares of output variance: with full information they explain 25% and 14%, but with inattention they explain 32% and 30%. For the variance of consumption, government spending, TFP and price markup shocks explain 39%, 18% and 9% with full information, and 47%, 19% and 17% in the baseline model.

Appendix to Chapter 2

B.1 Proofs

B.1.1 Derivation of equation 2.3

Suppose that there are N bond dealers, of which G are good. Each dealer is good with the same probability $g = \frac{G}{N}$. There are therefore S possible states of the world, where each state is a unique combination of matched dealers and interest rates, and each state is equally likely. S is given by:

$$S = \frac{N!}{G!(N-G)!} \tag{B.1}$$

Information processing is defined as the expected entropy reduction from prior to posterior beliefs, as is common in the rational inattention literature (Sims, 2003):

$$\mathcal{I} = H(h) - \mathbb{E}_s H(f|s) \tag{B.2}$$

Here H is the entropy function, a measure of the dispersion in a distribution:

$$H(\cdot) = -\sum_{k=1}^{S} \Pr(\text{state } k) \log \Pr(\text{state } k) \tag{B.3}$$

The distribution h is the prior belief, which is that each state is equally likely - i.e. no bond dealer is more likely to be good than any other. The distribution $f|s$ is the posterior, the distribution of beliefs implied by the signals households process, embodied in the decisions they took.

In the prior each state is equally likely, so:

$$H(h) = -\sum_{k=1}^{S} \frac{1}{S} \log \frac{1}{S} = \log S \tag{B.4}$$

Households choose signals that maximise their probability of successfully choosing a good bond dealer q for the minimum information cost. This means that if a signal instructs

a household member to choose dealer d^*, their posterior belief will be that each state in which d^* is good is equally likely. There is no use in distinguishing between those states. The posterior therefore attaches probability $\frac{qN}{SG}$ to each of those states, where $\frac{SG}{N}$ is the number of states in which d^* is good. Similarly, they assign probability $\frac{(1-q)N}{N-SG}$ to each state of the world in which d^* is bad. The entropy of this posterior is therefore:

$$H(f|s \Rightarrow \text{choose } d^*) = -q \log\big(\frac{qN}{SG}\big) - (1-q)\log\big(\frac{(1-q)N}{N-SG}\big) \tag{B.5}$$

This is independent of the arbitrary d^*, so this expression also gives the expected entropy across all such signals. The total information processed to attain a success probability of q is therefore:

$$\mathcal{I} = \log S + q \log\big(\frac{qN}{SG}\big) + (1-q)\log\big(\frac{(1-q)N}{N-SG}\big) \tag{B.6}$$

Splitting apart the logs and cancelling terms we find:

$$\mathcal{I} = \log N - q \log G - (1-q)\log(N-G) + q \log q + (1-q)\log(1-q) \tag{B.7}$$

Now replace $G = gN$ and simplify to obtain equation 2.3:

$$\mathcal{I} = q \log\left(\frac{q}{g}\right) + (1-q)\log\left(\frac{1-q}{1-g}\right) \tag{B.8}$$

All terms involving N, the specific number of bond dealers, have cancelled out, leaving only terms in g, the proportion of bond dealers who are offering good interest rates. This implies that as $\mathcal{I}$ approaches $-\log(g)$, the success probability q approaches 1. While the information processing formula is not defined for $q = 1$, q gets arbitrarily close to 1 for this finite amount of information processing, which does not depend on the number of bond dealers in the market. $\mathcal{I}'(q) \to \infty$ as $q \to 1$.

B.1.2 The household first order conditions are sufficient for utility maximisation

Here I prove that the household first order conditions (equations 2.7 - 2.9) are sufficient for utility maximisation. I only consider a household away from the no-borrowing constraint, as at the constraint they fix attention at zero and their problem becomes identical to the poor household problem in standard TANK papers (e.g. BMP, Debortoli and Galí (2018)).

For an unconstrained household, their utility maximisation problem can be written as an unconstrained maximisation by substituting out for consumption using the budget

constraint:

$$\max_{n_t,b_t,i_t^e} U = \sum_t \beta^t \left(\frac{\left(\frac{b_{t-1}}{\Pi_t} + w_t n_t + \mathcal{D}_t - T_t - \frac{b_t}{1+i_t^e}\right)^{1-\gamma}}{1-\gamma} - \frac{n_t^{1+\nu}}{1+\nu} - \mu\mathcal{I}(i_t^e) \right) \tag{B.9}$$

The first order conditions are:

$$\frac{\partial U}{\partial b_t} = -\frac{1}{1+i_t^e}\left(\frac{b_{t-1}}{\Pi_t} + w_t n_t + \mathcal{D}_t - T_t - \frac{b_t}{1+i_t^e}\right)^{-\gamma} + \mathbb{E}_t \frac{\beta}{\Pi_{t+1}}\left(\frac{b_t}{\Pi_{t+1}} + w_{t+1} n_{t+1} + \mathcal{D}_{t+1} - T_{t+1} - \frac{b_{t+1}}{1+i_{t+1}^e}\right)^{-\gamma} = 0 \tag{B.10}$$

$$\frac{\partial U}{\partial n_t} = w_t\left(\frac{b_{t-1}}{\Pi_t} + w_t n_t + \mathcal{D}_t - T_t - \frac{b_t}{1+i_t^e}\right)^{-\gamma} - n_t^{\nu} = 0 \tag{B.11}$$

$$\frac{\partial U}{\partial i_t^e} = \frac{b_t}{(1+i_t^e)^2}\left(\frac{b_{t-1}}{\Pi_t} + w_t n_t + \mathcal{D}_t - T_t - \frac{b_t}{1+i_t^e}\right)^{-\gamma} - \mu\mathcal{I}'(i_t^e) = 0 \tag{B.12}$$

The elements of the Hessian matrix are therefore (substituting consumption back in using the budget constraint where it simplifies the expressions):

$$\frac{\partial^2 U}{\partial b_t^2} = -\frac{\gamma}{(1+i_t^e)^2} c_t^{-\gamma-1} - \mathbb{E}_t \frac{\gamma\beta}{\Pi_{t+1}^2} c_{t+1}^{-\gamma-1} \tag{B.13}$$

$$\frac{\partial^2 U}{\partial n_t^2} = -w_t^2 \gamma c_t^{-\gamma-1} - \nu n_t^{\nu-1} \tag{B.14}$$

$$\frac{\partial^2 U}{\partial i_t^{e2}} = -\frac{\gamma b_t^2}{(1+i_t^e)^4} c_t^{-\gamma-1} - \frac{2b_t}{(1+i_t^e)^3} c_t^{-\gamma} - \mu\mathcal{I}''(i_t^e) \tag{B.15}$$

$$\frac{\partial^2 U}{\partial b_t \partial n_t} = \frac{\gamma w_t}{1+i_t^e} c_t^{-\gamma-1} \tag{B.16}$$

$$\frac{\partial^2 U}{\partial b_t \partial i_t^e} = \frac{1}{(1+i_t^e)^2} c_t^{-\gamma} + \frac{\gamma b_t}{(1+i_t^e)^3} c_t^{-\gamma-1} \tag{B.17}$$

$$\frac{\partial^2 U}{\partial n_t \partial i_t^e} = -\frac{\gamma b_t w_t}{(1+i_t^e)^2} c_t^{-\gamma-1} \tag{B.18}$$

The first order conditions are sufficient for utility maximisation of U is concave. This is the case if the Hessian matrix is negative semi-definite, i.e. if for any real-valued vector $[x \quad y \quad z]$:

$$[x \quad y \quad z] \begin{bmatrix} \frac{\partial^2 U}{\partial b_t^2} & \frac{\partial^2 U}{\partial b_t \partial n_t} & \frac{\partial^2 U}{\partial b_t \partial i_t^e} \\ \frac{\partial^2 U}{\partial b_t \partial n_t} & \frac{\partial^2 U}{\partial n_t^2} & \frac{\partial^2 U}{\partial n_t \partial i_t^e} \\ \frac{\partial^2 U}{\partial b_t \partial i_t^e} & \frac{\partial^2 U}{\partial n_t \partial i_t^e} & \frac{\partial^2 U}{\partial i_t^{e2}} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \leq 0 \tag{B.19}$$

Multiplying this out and simplifying we obtain:

$$-\gamma c_t^{-\gamma-1}\left(\frac{x}{1+i_t^e}-yw_t-\frac{zb_t}{(1+i_t^e)^2}\right)^2 - \mathbb{E}_t\frac{\gamma\beta x^2 c_{t+1}^{-\gamma-1}}{\Pi_{t+1}^2}+\frac{2xzc_t^{-\gamma}}{(1+i_t^e)^2}-\nu y^2 n_t^{\nu-1}-\frac{2z^2 b_t c_t^{-\gamma}}{(1+i_t^e)^3}-\mu z^2\mathcal{I}''(i_t^e)\leq 0 \quad \text{(B.20)}$$

This expression would clearly be true without the term in xz, which comes from $\frac{\partial^2 U}{\partial b_t \partial i_t^e}$. This is to be expected: it is the feedback between saving and attention (so interest rates) that is the cause for our concern that U may not be concave. I proceed by showing a condition under which this feedback is sufficiently weak that U remains concave, and so the first order conditions remain sufficient for utility maximisation.

Factorising the two terms containing $c_t^{-\gamma}$ gives:

$$\frac{2xzc_t^{-\gamma}}{(1+i_t^e)^2}-\frac{2z^2 b_t c_t^{-\gamma}}{(1+i_t^e)^3}=\frac{c_t^{-\gamma}}{(1+i_t^e)^3}\big(2xz(1+i_t^e)-z^2\big)=-\frac{c_t^{-\gamma}}{(1+i_t^e)^3}\big(z-x(1+i_t^e)\big)^2+\frac{x^2 c_t^{-\gamma}}{1+i_t^e} \quad \text{(B.21)}$$

Putting this back into the full expression, the condition for concavity becomes:

$$-\gamma c_t^{-\gamma-1}\left(\frac{x}{1+i_t^e}-yw_t-\frac{zb_t}{(1+i_t^e)^2}\right)^2 + x^2\left(\frac{c_t^{-\gamma}}{1+i_t^e}-\mathbb{E}_t\frac{\gamma\beta c_{t+1}^{-\gamma-1}}{\Pi_{t+1}^2}\right)-\frac{c_t^{-\gamma}}{(1+i_t^e)^3}\big(z-x(1+i_t^e)\big)^2-\nu y^2 n_t^{\nu-1}-\mu z^2\mathcal{I}''(i_t^e)\leq 0 \quad \text{(B.22)}$$

Assuming a non-negative information cost μ and an effective interest rate i_t^e which is greater than -100% (so $1+i_t^e>0$), a sufficient condition for the concavity of the utility function is therefore:

$$\frac{c_t^{-\gamma}}{1+i_t^e}-\mathbb{E}_t\frac{\gamma\beta c_{t+1}^{-\gamma-1}}{\Pi_{t+1}^2}\leq 0 \quad \text{(B.23)}$$

At the first order conditions, the Euler equation gives $c_t^{-\gamma}$ in terms of expected future consumption. I complete the proof by substituting this in to condition B.23 and showing that for plausible calibrations, the condition holds. This shows that the first order conditions are a local maximum of the utility function. Since there is only one stationary point of the utility function and the function is continuous, this local maximum must be the global maximum.

Substituting the Euler equation in to expression B.23 and dividing by β we obtain:

$$\mathbb{E}_t\frac{c_{t+1}^{-\gamma}}{\Pi_{t+1}}\leq\mathbb{E}_t\frac{c_{t+1}^{-\gamma}}{\Pi_{t+1}}\cdot\frac{\gamma}{\Pi_{t+1}c_{t+1}} \quad \text{(B.24)}$$

Therefore if the product of consumption and inflation is always less than γ, the first order conditions will be sufficient for utility maximisation. In all calibrations in this paper, $\gamma = 1$ and inflation remains close to its steady state at $\Pi = 1$. Steady state consumption of wealthy households is considerably less than 1 (it is 0.862 in the calibration I use to draw the figures in this paper), and so for all the exercises I consider we remain well within this region, where the first order conditions are sufficient for utility maximisation.

B.1.3 Proposition 1

Suppose that in the long run the poor household expects to return to their original level of consumption, the level of the representative agent steady state, which I denote $\bar{c}$. Iterating the consumption Euler equation forward we find that in any period t:

$$c_t^{p^{-\gamma}} = \bar{c}^{-\gamma} \mathbb{E}_t \prod_{s=t}^{\infty} \frac{\beta(1 + i_s^{e,p})}{\Pi_{s+1}} \tag{B.25}$$

The response of current consumption and saving depends on the expected future path of the effective interest rates facing the poor household.

Substituting the poor household period t budget constraint, and the labour supply condition into the first order condition on attention we obtain another condition relating poor household consumption to the interest rate they will earn, with no other choice variables present:

$$c_t^{p^{-\gamma}} (\frac{b_{t-1}^p}{\Pi_t} + w_t^{\frac{1+\nu}{\nu}} c_t^{p^{-\frac{\gamma}{\nu}}} - c_t^p + \mathcal{D}_t - T_t^p) = \mu(1 + i_t^{e,p}) \mathcal{I}(i_t^{e,p}) \tag{B.26}$$

Consider the possibility that if the poor household enters period t with less income than the wealthy household (either due to higher taxes or lower savings from the previous period), they still maintain the same attention, and so the same effective interest rates, as the wealthy households. From the Euler equation, poor households must therefore have the same consumption in period t as wealthy households. In that case, the left hand side of equation B.26 will be smaller for poor households than wealthy households, as income is lower and consumption the same. For equation B.26 to hold, poor households cannot therefore face the same interest rates as wealthy households.

If poor households face lower interest rates than wealthy households in period t, because they pay less attention, then from the Euler equation they will consume more in period t than wealthy households. This is consistent with equation B.26: the left hand side of that equation will be lower for poor households due to both lower income and higher consumption. Poor households expecting to return to the representative agent

steady state in the long run therefore pay less attention to choosing between different bond dealers and pay a higher price per bond.

I now proceed to show the conditions under which this implies that the gap between the incomes of the households will grow without further shocks. Let the savings of poor households in period t be $b^p_{t-1} = b^w_{t-1} - \epsilon_b$, with $\epsilon_b > 0$. The only reason incomes will differ between households in period $t+1$ is because poor households save less in period t and so have less asset income in the next period. From the Euler equation, the minimum that poor households could consume in period t is given by:

$$c^p_t = c^w_t \left(\frac{1 + i^{e,w}_t}{1 + i^{e,p}_t} \right)^{\frac{1}{\gamma}} \tag{B.27}$$

This is the consumption we would see if the poor household effective interest rate equals that of the wealthy households from period $t+1$ onwards. This provides a maximum for poor household saving, so if the gap in saving between households grows for this level of c^p_t it must grow for all possible c^p_t.

Using the budget constraints of the two households and the labour supply conditions we can find that the difference between the savings of the two households in period t with this minimum c^p_t is:

$$\begin{aligned} b^w_t - b^p_t = &\left(\frac{b_{t-1}}{\Pi_t} + \mathcal{D}_t \right)(i^{e,w}_t - i^{e,p}_t) + {w_t}^{\frac{1+\nu}{\nu}} {c^w_t}^{-\frac{\gamma}{\nu}} (1 + i^{e,w}_t) \left(1 - \left(\frac{1 + i^{e,p}_t}{1 + i^{e,w}_t} \right)^{\frac{1+\nu}{\nu}} \right) \\ &- c^w_t (1 + i^{e,w}_t) \left(1 - \left(\frac{1 + i^{e,p}_t}{1 + i^{e,w}_t} \right)^{\frac{\gamma-1}{\gamma}} \right) + \frac{\epsilon_b}{\Pi_t} \left(1 + i^{e,p}_t + (1-\omega)(i^{e,w}_t - i^{e,p}_t) \left(1 - \frac{\Pi_t}{1 + i_t} \right) \right) \end{aligned} \tag{B.28}$$

Where I have also substituted out for taxes using the government budget constraint, noting that in period $t-1$ government debt is equal to $\bar{b}_{t-1} = \omega b^w_{t-1} + (1-\omega) b^p_{t-1} = b^w_{t-1} - (1-\omega)\epsilon_b$.

The first term is always positive, as we have already shown that $i^{e,w}_t > i^{e,p}_t$.

The second term must also be positive, as $i^{e,p}_t < i^{e,w}_t$ and $\nu > 0$. The sign of the third term depends on the magnitude of γ, the inverse elasticity of intertemporal substitution. If $\gamma = 1$ (log utility), this term becomes zero. If $\gamma \in (0,1)$ then $\left(\frac{1+i^{e,p}_t}{1+i^{e,w}_t}\right)^{\frac{\gamma-1}{\gamma}} > 1$, so the whole term is positive. If $\gamma > 1$ then $\left(\frac{1+i^{e,p}_t}{1+i^{e,w}_t}\right)^{\frac{\gamma-1}{\gamma}} < 1$ and the term is negative.

A sufficient condition for the sum of terms 2 and 3 to be positive is if:

$$\frac{1 + i^{e,p}_t}{1 + i^{e,w}_t} > \left(\frac{(\gamma - 1)\nu}{\gamma(1+\nu)} \right)^{\frac{\gamma\nu}{\gamma+\nu}} {w_t}^{-\frac{\gamma(1+\nu)}{\gamma+\nu}} {c^w_t}^{\gamma} \tag{B.29}$$

To obtain this condition, first note that the sum of the terms is positive if and only if:

$$F := {w_t}^{\frac{1+\nu}{\nu}} {c_t^w}^{-\frac{\gamma+\nu}{\nu}} \left(1 - \left(\frac{1+i_t^{e,p}}{1+i_t^{e,w}}\right)^{\frac{1+\nu}{\nu}}\right) - \left(1 - \left(\frac{1+i_t^{e,p}}{1+i_t^{e,w}}\right)^{\frac{\gamma-1}{\gamma}}\right) > 0$$

If $i_t^{e,p} = i_t^{e,w}$, $F = 0$ and the two middle terms in equation B.28 sum to zero. F is maximised with respect to $\frac{1+i_t^{e,p}}{1+i_t^{e,w}}$ at:

$$\left(\frac{1+i_t^{e,p}}{1+i_t^{e,w}}\right)^* = \left(\frac{(\gamma-1)\nu}{\gamma(1+\nu)}\right)^{\frac{\gamma\nu}{\gamma+\nu}} {w_t}^{-\frac{\gamma(1+\nu)}{\gamma+\nu}} {c_t^w}^{\gamma} \tag{B.30}$$

If condition B.29 holds, we therefore have that $\frac{1+i_t^{e,p}}{1+i_t^{e,w}}$ is above the value that maximises F. Since we know $\frac{1+i_t^{e,p}}{1+i_t^{e,w}} < 1$, this also ensures that the unique maximum of F occurs at $\left(\frac{1+i_t^{e,p}}{1+i_t^{e,w}}\right)^* < 1$, and since $F = 0$ when $i_t^{e,p} = i_t^{e,w}$ and F is continuous, we know that this maximum has $F > 0$. With $\frac{1+i_t^{e,p}}{1+i_t^{e,w}}$ between the level that maximises F and 1, we therefore have that $F \in (0, F^*)$. The second and third terms in equation B.28 sum to a positive number.

Income differentials grow if:

$$\frac{b_t^w - b_t^p}{\Pi_{t+1}} > \frac{b_{t-1}^w - b_{t-1}^p}{\Pi_t} = \frac{\epsilon_b}{\Pi_t} \tag{B.31}$$

If the sum of terms 1-3 in equation B.28 is positive then a sufficient condition for income differentials to grow is that the real interest rate earned by poor households is positive:

$$\frac{b_t^w - b_t^p}{\Pi_{t+1}} > \frac{\epsilon_b}{\Pi_t} \quad \text{if} \quad \frac{1+i_t^{e,p}}{\Pi_{t+1}} > 1 \tag{B.32}$$

This is because the right hand side of equation B.28 is strictly greater than the final term in isolation, which in turn is strictly greater than $\frac{\epsilon_b}{\Pi_t}(1+i_t^{e,p})$ (as long as $1+i_t > \Pi_t$, which will be assured by the Taylor Rule as long as Π_t doesn't deviate too far from steady state[1]).

We know that in the representative agent steady state the real interest rate equals $\beta^{-1} > 1$, so for small shocks this condition will hold. This is also an extremely small subset of the conditions under which income differentials will grow, as it misses the effects of the other terms in equation B.28.

[1]To see this, notice that since $q^w < 1$, $\bar{i} > \bar{i}^{e,w} = \beta^{-1} - 1 > 0$. Therefore $1 + i_t > \Pi_t$ if $\beta^{-1} > \Pi_t + (1 - \frac{\Pi_t}{\bar{\Pi}}^{\delta_\Pi} \frac{Y_t}{\bar{Y}}^{\delta_Y})$. Close to steady state the right hand side of this expression will be close to Π_t, which will be close to $\bar{\Pi} = 1$, and so the condition will be satisfied. Another approach is to note that $1+i_t > \Pi_t$ if and only if lump sum taxes are positive in period t (from the government budget constraint assuming asset supply is the same in periods $t-1$ and t). In steady state with $\bar{i} > \bar{i}^{e,w} = \beta^{-1} - 1$, $\bar{T} > 1 - \beta > 0$. Close to steady state the condition will therefore be satisfied.

Condition B.29 will hold for a wide range of sensible calibrations. If $\gamma \leq 1$ it will hold, as the right hand side will be (weakly) negative while the left hand side is always strictly positive. All of the numerical results in the paper are for log utility ($\gamma = 1$), so the condition holds.

In general condition B.29 is more likely to be satisfied when consumption is low, wages are high, and γ and ν are low, but these are not very restrictive bounds. The non-shaded area in figure B.1 shows the range of w_t and c_t^w for which condition B.29 holds for any of the possible values of $i_t^{e,w}$ and $i_t^{e,p}$, given $q^w, q^p \in (g, 1)$, if the policy rate is at its steady state level, for the extremely conservative calibration $\gamma = 3$, $\nu = 5$. The red diamond shows the steady state w, c^w in the representative agent steady state, which is easily within the region in which condition B.29 is satisfied. In fact for that combination of w, c^w to be outside of the white region with these parameters, we would require $\frac{1+i_t^{e,p}}{1+i_t^{e,w}} < 0.75$, implying a more than 2500 basis point difference in the returns faced by the two households, which is implausibly large.

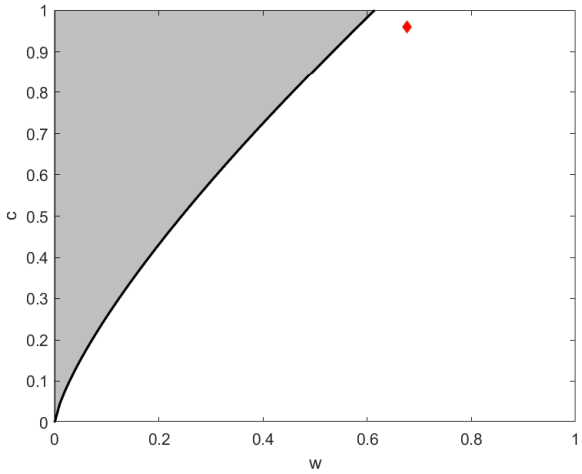


Figure B.1: Values of w_t, c_t^w such that condition B.29 holds for any possible interest rate differential, for $\gamma = 3$, $\nu = 5$ (non-shaded region). Steady state values as diamond.

B.1.4 Consumption responds less to tax cuts in the attention trap model

Consider a MIT shock to taxes (debt-funded or redistributive): the shock happens unexpectedly in period 1, and households know that there will be no further shocks from period 2 onwards. Furthermore, assume that the shock causes poor households to lift off their budget constraint for period 1 only in both the attention trap and heterogeneous β model.

I show that in this case poor household consumption is lower in period 1 in the

attention trap model than in the heterogeneous β model. This is sufficient to ensure that the impact of the fiscal policy shock on aggregate consumption is smaller in the attention trap model, because I restrict the attention trap model to calibrations in which wealthy households pay sufficient attention to choosing bond dealers that they successfully identify a good (high interest rate) bond dealer with probability $q^w \approx 1$ each period, and any variations with shocks are negligible. This means that the consumption problem of wealthy households is the same in both models, so if poor households behave the same in the two models all macro variables are also equal. If poor households consume less in one model, this directly reduces the aggregate consumption in that model, but also reduces it further through the effects of lower wages[2].

From period 3 onwards all state variables (the shocks and the wealth distribution) are at their steady states[3], and so therefore are all other variables. In period 2 the only state variable not at steady state is the distribution of asset income[4], as poor households were not against their borrowing constraint in period 1 and so have some asset income in period 2. The relative consumption levels across the two models in period 2 are determined by the levels of saving in period 1, which is in turn determined by the poor household Euler equation in each model.

Denoting variables in the attention trap model with a subscript A and those in the heterogeneous β model with subscript B, from the Euler equations in an arbitrary period t (and the definition of the poor household interest rate for the attention trap model) we have:

$$c_{t,A}^{p^\gamma} = \frac{\Pi_{t+1,A}(1+\chi(1-q_{t,A}^p))}{\beta(1+i_{t,A})} c_{t+1,A}^{p^\gamma}, \quad c_{t,B}^{p^\gamma} = \frac{\Pi_{t+1,B}}{\beta^p(1+i_{t,B})} c_{t+1,B}^{p^\gamma} \tag{B.33}$$

The two models are calibrated such that the steady state θ is the same in both models. Using equations 2.29 and 2.33 we therefore have that $\beta^p = \beta - \chi(1-g)$. Using this and the Euler equations above we have:

$$c_{1,A}^\gamma - c_{1,B}^\gamma = \frac{\Pi_{2,A} c_{2,A}^\gamma}{\beta(\beta-\chi(1-g))}\Bigg[\chi(1-g)\left(\beta\frac{1-q_{1,A}}{1-g} - 1 - \chi(1-q_{1,A})\right) + \beta\left(1 - \frac{\Pi_{2,B} c_{2,B}^\gamma}{\Pi_{2,A} c_{2,A}^\gamma}\right)\Bigg] \tag{B.34}$$

[2]The effects of changes in the waste due to households choosing bad savings products are negligible.

[3]This assumes there is no persistence in the shock. This therefore applies to the redistribution shocks I study, and to debt-funded tax cuts that are completely funded by higher taxes in period 2.

[4]If we are considering a debt-funded tax cut paid off with higher taxes in period 2 then the tax is also not at steady state in period 2, but this does not affect the results, as the only difference in the tax level between the models comes from differences in the degree to which the debt has been inflated away. With plausible levels of price stickiness inflation does not move far enough from steady state in either model for this difference to overturn the result.

Inside the square brackets, the first term is always negative since $q_{1,A} \geq g$ and $\beta < 1$. The second term is ambiguous. If in period 2 poor household consumption is lower in model A then so is aggregate consumption, output, and so inflation.

If, therefore $c_{2,A} \leq c_{2,B}$, we must have that $c^{\gamma}_{1,A} - c^{\gamma}_{1,B} < 0$, so the response of poor household consumption to the tax cut in the period of the shock is smaller in the attention trap model.

If instead $c_{2,A} > c_{2,B}$ we can obtain the same result through a different route. For this to be the case, poor households must have saved more in period 1 in model A, as asset income is the only state variable not at steady state in period 2 in both models. If poor households are saving more in period 1 in model A, they must be consuming less than in model B, so again we have that the impact of the tax cut is smaller in period 1 on poor, and so aggregate, consumption in the attention trap model.

Formally, denoting net household income from labour, dividends and taxes in period 1 for model x as $M_{1,x}$ we have from the budget constraint of poor households in each model:

$$b^{p}_{1,A} = (1 + i^{e,p}_{1,A})(M_{1,A} - c^{p}_{1,A}), \quad b^{p}_{1,B} = (1 + i^{p}_{1,B})(M_{1,B} - c^{p}_{1,B}) \tag{B.35}$$

If policy rates are the same in the two models then due to inattention we know that $1 + i^{e,p}_{1,A} < 1 + i^{p}_{1,B}$. This will remain true away from the case of equal policy rates as long as $i_{1,A}$ is not too far above $i_{1,B}$ relative to the losses from inattention. That is, as long as:

$$q_{1,A} < 1 + \frac{1}{\chi}\left(\frac{1}{1 + i_{1,A}} - \frac{1}{1 + i_{1,B}}\right) \tag{B.36}$$

Assuming this holds[5] then we have:

$$b^{p}_{1,A} - b^{p}_{1,B} < (1 + i_{1,B})(M_{1,A} - M_{1,B} + c^{p}_{1,B} - c^{p}_{1,A}) \tag{B.37}$$

For poor households to save more in period 1 in the attention trap model than the heterogeneous β model we therefore need:

$$M_{1,A} - M_{1,B} > c^{p}_{1,A} - c^{p}_{1,B} \tag{B.38}$$

It is only possible to have simultaneously higher consumption and saving in period 1 in

[5]This does hold with even small amounts of interest rate heterogeneity, since lower interest rate heterogeneity reduces the benefits of attention and so reduces $q_{1,A}$. With the calibration used in sections 2.4 and 2.5, even the large debt-funded tax cut that implies a 8.7 percentage point increase in debt-to-GDP implies that poor household effective interest rates increase to only marginally above the steady state policy rate (by 2 basis points), and so the same large policy would need to have almost no effect in the heterogeneous β model for it to be the case that $i^{e,p}_{1,A} > i_{1,B}$. That is not the case, as shown by BMP. The same shock in the heterogeneous β model used here causes the policy rate to rise by 6.5 basis points.

the attention trap model if incomes are substantially higher in the attention trap model, so much so that they outweigh the higher cost of bonds for poor households in that model.

If we were to have that $c^p_{1,A} > c^p_{1,B}$, then it is ambiguous whether $M_{1,A} > M_{1,B}$ or not. Higher labour demand in model A leads to higher wages, but this is offset by the higher consumption leading to a reduction in labour supply by the poor households through an income effect[6]. In addition, with higher wages dividend income will be lower in model A in this case. Therefore, as long as the wage increases are not extremely large relative to these other effects we can rule out that poor households both consume and save more in period 1 in the attention trap model than in the heterogeneous β model.

With larger shocks that cause poor households to be away from the borrowing constraint for many periods we have a similar logic. Equation B.34 holds in all periods, so we know that if consumption in period $t+1$ is lower in model A then so must be consumption in period t. There must be at least one initial period of lower consumption in model A to accumulate the extra saving that allows for higher consumption in future periods.

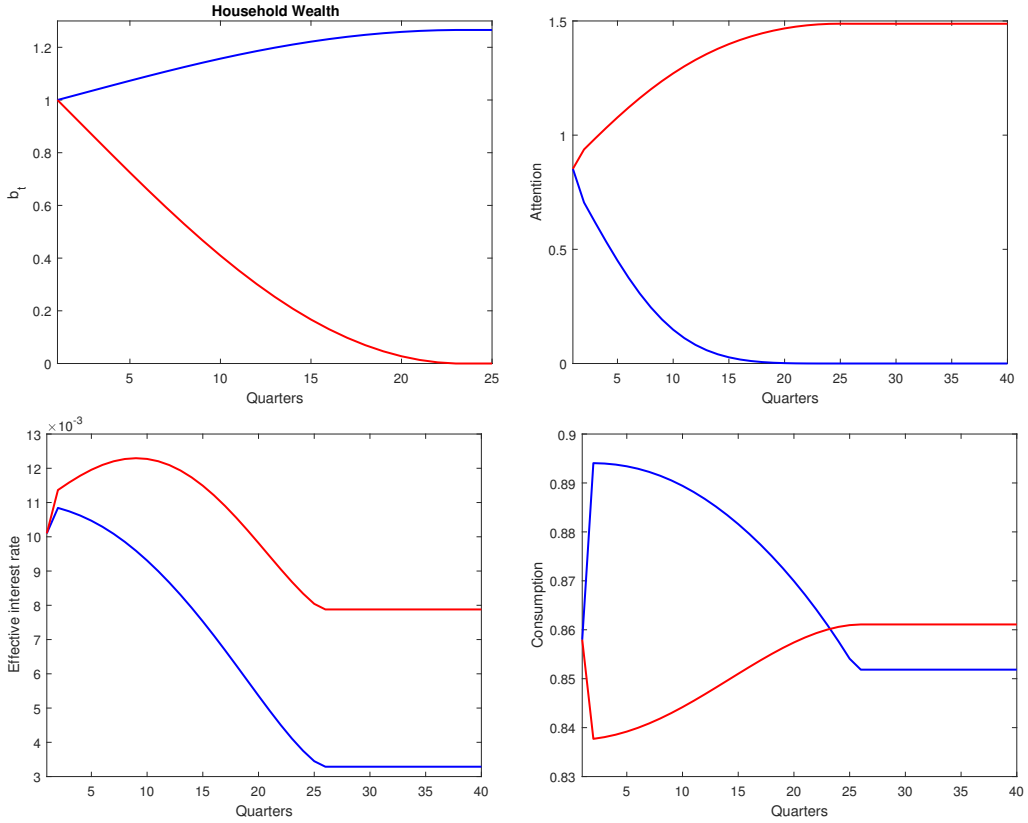
B.2 Steady State Transition Simulation

As long as the conditions for Proposition 1 hold, the representative agent steady state is unstable with respect to idiosyncratic shocks. The graphs below show the transition paths of key model variables after a one-off small idiosyncratic tax shock, starting at the representative agent steady state[7]. As described in section 2.3.2, household wealth and attention diverge and settle on the two-agent steady state. During the transition, interest rate heterogeneity grows and poor households initially consume more than wealthy households as they run down their savings. In the new steady state poor household consumption is below wealthy household consumption.

[6]The higher labour demand is met by wealthy households increasing their labour supply. Their consumption rises by less as they increase saving a lot to absorb the higher debt supply (or as their incomes fall due to higher taxes in the case of a redistributive shock), and so they respond to the higher wages with a greater labour supply. This willingness to expand labour supply also tempers the wage response to higher poor household consumption.

[7]The calibration used for this simulation is mostly the same as that used for the simulations in sections 2.4 and 2.5, except I increase the cost of information to $\mu = 0.0025$. This is to ensure that the amount of attention paid by households in the representative agent steady state is such that q is substantially below 1, to make it clear in the graphs that wealthy households increase their attention as the poor households decrease theirs. This is also true in the baseline calibration, but the increase in attention is from one level extremely close to $-\log(g)$ to one even closer than that. The difference in attention is therefore so small that it would not be visible on a graph - and for that reason the plots here show the case with a higher information cost. The other change I make is that I set $\bar{i}$ and $\bar{y}$ such that there is no inflation in the representative agent steady state. Aggregate consumption is a little different in the TANK steady state, so these values are different to those used to ensure $\Pi = 1$ in the TANK steady state in section 2.4. Inflation does not therefore revert exactly to 1 in the transition.

Figure B.2: Transition paths for wealthy (red) and poor (blue) households after a small idiosyncratic tax shock.



B.3 Comparison Model Equations

The household side of the model consists of an Euler equation, intratemporal labour supply condition, and a budget constraint for each type of household, $h \in \{w, p\}$:

$$c_t^{h^{-\gamma}} = \beta^h (1 + i_t) \mathbb{E}_t \left(\frac{c_{t+1}^{h^{-\gamma}}}{\Pi_{t+1}} \right) + (1 + i_t)\theta_t^h \tag{B.39}$$

$$c_t^{h^{\gamma}} n_t^{h^{\nu}} = w_t \tag{B.40}$$

$$c_t^h + \frac{b_t^h}{1 + i_t} = \frac{b_{t-1}^h}{\Pi_t} + w_t n_t^h + \mathcal{D}_t - T_t^h \tag{B.41}$$

The firm, policy and market clearing blocks of the model are as in section 2.2, except that there is no waste through bad bond dealer transaction costs in the goods market clearing condition. The equations consist of a Phillips curve, production function, government

budget constraint, tax rule, and market clearing conditions for bonds and goods:

$$Y_t(1-\epsilon)+\frac{\epsilon}{1-\alpha}w_t\left(\frac{Y_t}{A_t}\right)^{\frac{1}{1-\alpha}}-\psi(\Pi_t-1)\Pi_t Y_t+\psi\beta\mathbb{E}_t(\Pi_{t+1}-1)\Pi_{t+1}Y_{t+1}\left(\frac{\bar{c}_t}{\bar{c}_{t+1}}\right)^{\gamma}=0 \tag{B.42}$$

$$Y_t=A_t(\omega n_t^w+(1-\omega)n_t^p)^{1-\alpha} \tag{B.43}$$

$$\frac{\bar{b}_{t-1}}{\Pi_t}=\frac{\bar{b}_t}{1+i_t}+(\omega T_t^w+(1-\omega)T_t^p) \tag{B.44}$$

$$1+i_t=\bar{i}+\left(\frac{\Pi_t}{\bar{\Pi}}\right)^{\delta_\Pi}\left(\frac{Y_t}{\bar{Y}}\right)^{\delta_Y}+M_t \tag{B.45}$$

$$\bar{b}_t=\omega b_t^w+(1-\omega)b_t^p \tag{B.46}$$

$$T_t^w=T_t^p+z_t \tag{B.47}$$

$$Y_t=\omega c_t^w+(1-\omega)c_t^p+\frac{\psi}{2}(\Pi_t-1)^2Y_t \tag{B.48}$$

B.4 Calibration for sections 2.4 and 2.5

Most parameters are the same in the attention trap and heterogeneous β models. The common parameters are set at standard values, shown in the first block of the table below. After those parameters, the heterogeneous β model has just one novel parameter, the discount factor of poor (impatient) households. After displaying that the final portion of the table shows the parameter values for the information problem in the attention trap model. These have been chosen so that θ^p, the shadow value of relaxing the borrowing constraint for the poor household, is equal in steady state across the two models. The information cost has been set sufficiently low that q^w is arbitrarily close to 1 in all periods of the simulations.

Parameter	Description	Value
β	Discount factor	0.99
γ	Coefficient of relative risk aversion	1
ν	Disutility of labour	1
ω	Proportion of households which are wealthy	0.65
ϵ	Elasticity of substitution between intermediate good varieties	9
A	Steady state technology	1
α	Degree of diminishing returns in intermediate firms' production	0.25
ψ	Price adjustment costs	372.8
$\mathcal{B}$	Steady state government debt	1
$\bar{\Pi}$	Gross inflation target	1
d_{Π}	Elasticity of the interest rate to inflation	1.5
$\bar{Y}$	Output target	0.858954
$\bar{i}$	Interest rate target	0.01010101
d_Y	Elasticity of the interest rate to output	0.125
M	Steady state monetary policy shock	0
β^p	Poor household discount factor	0.985
μ	Information cost	0.0001
g	Fraction of good financial institutions	0.15
χ	Inefficient bond dealer costs	0.005882
χ^-	Lower inefficient bond dealer costs for section 2.5	0.002941

B.5 Intratemporal Redistribution

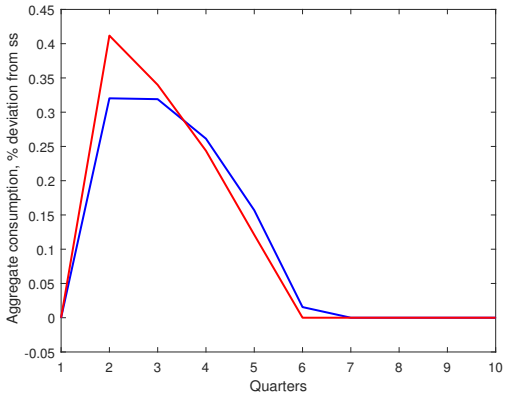
Here I show results for simulations of a one-period redistribution shock, which reduces taxes on poor households, funded by contemporaneous tax rises for wealthy households. This is expansionary in both models because poor households have a higher MPC than wealthy households.

As with the debt-funded tax cut, if the shock is sufficiently small that poor households remain at their borrowing constraint then the shock has identical implications in both the attention trap and heterogeneous β models. If the shock is large enough that poor households temporarily save positive amounts, aggregate consumption responds less on impact, but is more persistent, in the attention trap model, for the same reasons discussed for debt-funded tax cuts in section 2.4.

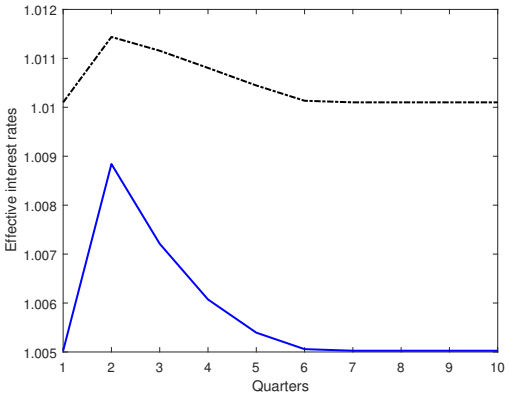
Using the same calibration as in section 2.4, figure B.3a below shows the response of aggregate consumption in the two models after a redistribution that causes the 35% of poor households to save 1% of the stock of government debt. Aggregate consumption rises by 22% less on impact in the attention trap model, but is more persistent, returning to

steady state 1 quarter later than in the heterogeneous β model. Figure B.3b shows why: effective interest rates facing poor households rise substantially relative to the policy rate in the attention trap model.

Figure B.3: Effects of a redistributive tax shock



(a) Aggregate consumption in the attention trap model (blue) and the heterogeneous β model (red).



(b) Effective interest rates facing poor (blue solid) and wealthy (black dashed) households in the attention trap model.

As with debt-funded tax cuts, larger shocks amplify the difference between the two models. A redistribution that causes poor households to save 2.5% of government debt supply has a 65% smaller impact on aggregate consumption in the attention trap model than in the heterogeneous β model.

Appendix to Chapter 3

C.1 Aggregate consumption function graphs for the static model

The graph below is the same as figure 3.2, with an extra curve added in black. This is the aggregate consumption function with a lower (but still positive) cost of information $\psi = 0.00064$. This implies $\kappa \approx 1$.

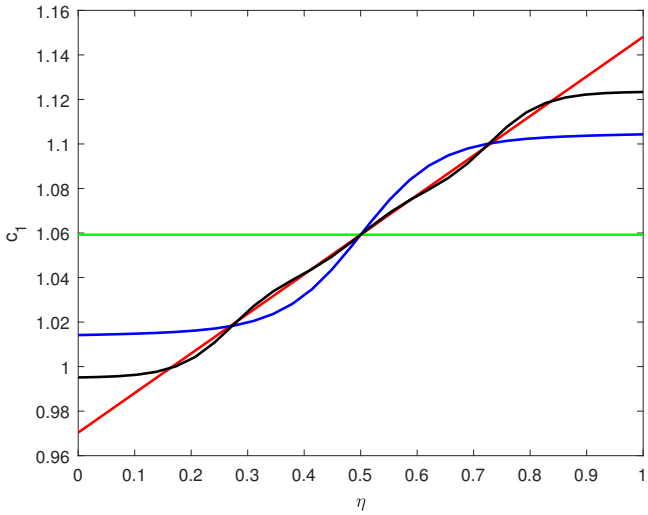


Figure C.1: Aggregate consumption function with $\kappa = 0$ (green), $\kappa = 0.5$ (blue), $\kappa = 1$ (black) and in the unconstrained case (red)

The shape of the aggregate response curve in the less constrained ($\psi = 0.00064$) case has the same form as the baseline case of $\psi = 0.002$, but with this greater information processing capacity agents choose from four levels of consumption, so there are four flat regions in the aggregate response curve.

C.2 Static Model: Firms

Here we derive the linear equilibrium relationship between aggregate consumption and the hiring rate in section 3.2.1.

In period 1, all households are employed at wage 1 and the firm sells $\bar{c}_1$ units of the consumption good. Assume that any unsold output in period 1 is wasted, and that the firm receives some subsidy for operating S. Firm profits are therefore equal to $\bar{c}_1 - 1 + S$.

New hires in period 2 (h) are determined by the number of vacancies posted (v) and the number of job seekers (u) through the matching function:

$$h = mv^{1-\alpha}u^{\alpha} \tag{C.1}$$

Now note that the number of job seekers at the start of period 2 is equal to the number of separations at the end of period 1, ω, since all households were employed in period 1. In order to hire, firms must post vacancies. The cost per vacancy is k. Using the matching function, we have that the cost of hiring one worker is[1]:

$$C(\eta) = km^{\frac{-1}{1-\alpha}}\omega\eta^{\frac{1}{1-\alpha}} \tag{C.2}$$

Assume that these costs must be paid out of period 1 profits, before firms do any period 2 production or sales. The profit per hire is:

$$D(\eta) = F - w - C(\eta) \tag{C.3}$$

Where F is the value of production from that worker, equal to period 2 aggregate consumption per worker plus the value of inventory at the end of period 2. Assume that this is sufficiently high that $D(\eta) > 0$ for all $\eta \in (0, 1)$. That is, in period 2, workers are very productive, and any output not sold is held as inventory, which has a high value to the firm. The firm would therefore always like to hire as many workers as possible, given the working capital constraint that they must use period 1 profits to pay for vacancies.

This creates an upward sloping relationship between $\bar{c}_1$ and η: when period 1 consumption is higher, the firm sells more in period 1, and makes more profit. That means the firm can post more vacancies in period 2, and so the hiring rate increases. Specifically, the hiring rate is pinned down by:

$$\bar{c}_1 - 1 + S = km^{\frac{-1}{1-\alpha}}\omega\eta^{\frac{1}{1-\alpha}} \tag{C.4}$$

In the graphs in section 3.2.1 I further assume that $\alpha = 0$, so the matching function only depends on the number of vacancies posted, which implies a linear relationship between

[1] I am assuming that firms are large, so the proportion of vacancies filled equals the probability of filling a vacancy.

$\bar{c}_1$ and η:

$$\bar{c}_1 = 1 - S + \frac{k\omega}{m}\eta \tag{C.5}$$

This is the linear relationship used in section 3.2.1. The parameters used there are $\beta = 0.975$, $R = \frac{1}{\beta}$, $\omega = 0.4$, $m = 0.25$, $k = 0.062w$, $w = 1.3$, $\theta = 0.4$, $S = 0.01$

C.3 HANK model details

C.3.1 Household problem

The employed household's problem under perfect information is:

$$\max_{c_{es}} V_s^e = \left(\frac{c_{es}^{1-\mu} - 1}{1-\mu} - \zeta\right) + \beta\mathbb{E}_s\left(\omega(1-\eta_{s+1})V_{s+1}^u\right) + \beta\mathbb{E}_s\left((1-\omega(1-\eta_{s+1}))V_{s+1}^e\right) \tag{C.6}$$

subject to

$$P_s c_{es} + \frac{b_{s+1}}{R_s} = P_s w + b_s \tag{C.7}$$

Here V_s^e and V_s^u are the values of being employed and unemployed in period s respectively. Employed households consume c_{es}, and have disutility of labour ζ. The coefficient of risk aversion is μ and the discount factor is β. Prices are P_s, the interest rate is R_s and the exogenous real wage[2] is w. Households save by buying one period bonds b_{s+1}.

The first order condition under perfect information, substituting inflation $\Pi_{s+1} = \frac{P_{s+1}}{P_s}$, is as in RS:

$$c_{es}^{-\mu} = \beta\frac{R_s}{\Pi_{s+1}}\left(\omega(1-\eta_{s+1})c_{u,s+1}^{-\mu} + (1-\omega(1-\eta_{s+1}))c_{e,s+1}^{-\mu}\right) \tag{C.8}$$

Unemployed households are always at their borrowing constraint, and so their problem never matters for equilibrium determination.

To get the steady state Euler equation (**EE**) note that since equilibrium asset holdings are always zero $c_{es} = w$ and $c_{us} = \vartheta$, where ϑ is the payoff to home production. Then substitute for a simple interest rate rule[3]:

$$R_s = \bar{R}\bar{\Pi}^{-\delta_\pi}\Pi^{\delta_\pi} \tag{C.9}$$

[2]Ravn and Sterk consider Nash bargaining over wages, so the wage changes with tightness in the labour market. Since we obtain a non-degenerate wealth distribution when we introduce rational inattention into this model, we keep to fixed wages to prevent the wealth distribution implying heterogeneous wages through heterogeneity in worker outside options, as that is not the focus of this paper. Assuming fixed wages places our model in the counter-cyclical labour income risk region of the RS model.

[3]This is as in RS, abstracting from the zero lower bound on interest rates and setting the coefficient on the hiring rate in the interest rate rule to 0, as RS do in drawing their figure 3.

This gives:

$$1 = \beta R^* \Pi^{\delta_\pi - 1} \Big[\omega(1-\eta) \Big(\frac{\vartheta}{w} \Big)^{-\mu} + 1 - \omega(1-\eta) \Big] \tag{C.10}$$

Where $R^* = \bar{R}\bar{\Pi}^{-\delta_\pi}$.

C.3.2 Firm problem

Firm j in period s chooses price P_{js} and vacancy posting v_{js} to maximise real profits. The firm's costs are wages, vacancy posting costs, and quadratic price adjustment costs:

$$\max_{P_{js}, v_{js}} \mathbb{E}_t \sum_{s=t}^{\infty} \Lambda_{j,t,t+s} \Big[\frac{P_{js}}{P_s} y_{js} - w_s n_{js} - k v_{js} - \frac{\phi}{2} \Big(\frac{P_{js} - P_{j,s-1}}{P_{j,s-1}} \Big)^2 \Big] \tag{C.11}$$

subject to

$$y_{js} = \left(\frac{P_{js}}{P_s} \right)^{-\gamma} y_s \tag{C.12}$$

$$y_{js} = \exp(A_s) n_{js} \tag{C.13}$$

$$n_{js} = (1-\omega) n_{j,s-1} + v_{js} q_s \tag{C.14}$$

Here y_{js} is firm j output and $y_s = \int y_{js} dj$ is total output. Firm j employs n_{js} units of labour. Each period it loses a fraction ω of its workforce through exogenous separations. The vacancy filling rate is q_s, so $v_{js} q_s$ is the number of new hires by firm j in period s.

Solving this profit maximisation we have:

$$1 - \gamma + \gamma m c_s = \phi(\Pi_s - 1)\Pi_s - \phi \mathbb{E}_s \Lambda_{s,s+1} \Big[\frac{y_{s+1}}{y_s} (\Pi_{s+1} - 1) \Pi_{s+1} \Big] \tag{C.15}$$

where

$$m c_s = \frac{w_s}{e^{A_s}} + \frac{k}{q_s} - (1-\omega) \mathbb{E}_s \Lambda_{s,s+1} \frac{k}{q_{s+1}} \tag{C.16}$$

Here we have dropped the j subscripts because all firms make the same choices.

The labour market matching function is:

$$M(u_s, v_s) = u_s^{\alpha} v_s^{1-\alpha} \tag{C.17}$$

The vacancy filling rate q_s and hiring rate η_s are given by $\frac{M_s}{v_s}$ and $\frac{M_s}{u_s}$ respectively, so:

$$q_s = \eta_s^{\frac{-\alpha}{1-\alpha}} \tag{C.18}$$

Also note that the firm owners are risk neutral, so the stochastic discount factor $\Lambda_{s,s+1} = \beta$ for all periods s. The technology shock A_s equals 0 in steady state. The steady state

Phillips Curve is therefore:

$$\phi(1-\beta)(\Pi-1)\Pi = 1-\gamma+\gamma(w+k\eta^{\frac{\alpha}{1-\alpha}}(1-\beta(1-\omega))) \tag{C.19}$$

C.3.3 Calibration

These parameter values are used to draw the figures in section 3.3. The calibration is monthly, and is based on Ravn and Sterk (2020) appendix A.1.2. We set the annualised inflation target to 2%, the nominal interest rate target such that the annualised real interest rate is 2% when inflation is at target, and we choose the price stickiness parameter ϕ such that the linearised Phillips Curve in a full-information, complete-insurance version of the model would have the same slope as the Phillips Curve in an equivalent model with Calvo price setting and an average price duration of 6 months. The parameter ψ, the marginal cost of information, is set to ensure that each household always chooses an optimal menu of savings choices with two discrete savings levels. We choose the wage so that the full information steady state has a monthly hiring rate of 0.48, as in Hall (2005), and inflation equal to target.

Parameter	**Name**	**Value**
ε	Elasticity of substitution	6
ϕ	Price adjustment cost	146.3
k	Vacancy posting cost	$0.1w$
α	Elasticity of matching function	0.5
ω	Separation rate	0.02
w	Wage	0.83254
θ	Home production value	$0.8w$
β	Discount factor	0.995
μ	Coefficient of risk aversion	2
ψ	Information processing cost	0.0025
$\bar{\Pi}$	Monthly inflation target	$1.02^{\frac{1}{12}}$
$\bar{R}$	Monthly nominal interest rate target	$1.02^{\frac{1}{24}}$
d_Π	Taylor Rule parameter	1.5

C.3.4 Household problem with rational inattention

We substitute the budget constraint C.7 into the value function, and add in information costs. The employed household problem with RI is therefore as follows:

$$f = \arg\max_{\hat{f}} \mathbb{E}[V_s^e(\eta_{s+1}, b_{s+1})] - \psi\kappa$$

$$= \arg\max_{\hat{f}} \int_{\eta_{s+1}} \int_{b_{s+1}} V_s^e(\eta_{s+1}, b_{s+1}) \hat{f}(\eta_{s+1}, b_{s+1}) d\eta_{s+1} db_{s+1} - \psi\kappa \quad \text{(C.20)}$$

subject to

$$\int_x \hat{f}(\eta_{s+1}, b_{s+1}) dx = g(\eta_{s+1}) \qquad \forall \eta_{s+1} \quad \text{(C.21)}$$

$$\hat{f}(\eta_{s+1}, b_{s+1}) \geq 0 \qquad \forall \eta_{s+1}, b_{s+1} \quad \text{(C.22)}$$

$$H[g(\eta_{s+1})] - \mathbb{E}_{b_{s+1}} H[\hat{f}(\eta_{s+1}|b_{s+1})] \leq \kappa \quad \text{(C.23)}$$

The function $H[.]$ is the entropy of the distribution over which it operates, defined in equation 3.13. These constraints are explained in section 3.2.2. As in the static model, we consider values for the information cost ψ that imply households optimally limit themselves to a menu with two choices of saving b_{s+1}. This makes the diagrams clearer, but it is not necessary for our results. A lower information cost would imply more choices in the optimal menu, and potentially more steady states than the two we find here.

We assume that agents know the current value of all parameters and variables except the hiring rate. They observe current interest rates and prices. We consider the problem in steady state, so we assume agents expect inflation to remain constant. This implies (through the interest rate rule in equation 3.18) a constant interest rate. Households either know this or simply expect interest rates to remain constant as they do with inflation. We normalise current prices P_s to 1, so current inflation enters the problem through R_s and assuming constant inflation pins down expectations of P_{s+1}. We run this problem many times, for a variety of possible inflation rates.

Specifically, households solving this problem are processing information about the hiring rate in the next period. For simplicity, we do not allow them to process information about the hiring rate in periods further in the future.

This implies that unemployed households derive no benefit from information about the hiring rate, since in the current period they are at their borrowing constraint. Information about the next period hiring rate has no effect on their decision. This means that unemployed households choose to process no information at all. In steady state the prior belief distribution of newly hired households is therefore identical to that of the population as a whole.

C.3.5 Solution

The solution to the rationally inattentive household problem for an employed household with zero wealth $b_s = 0$ facing a (known) monthly gross inflation of $\Pi = 1.005$ is plotted as an example in figure C.2. It shows how the savings choices of employed households with uniform prior beliefs vary with η_{s+1}.

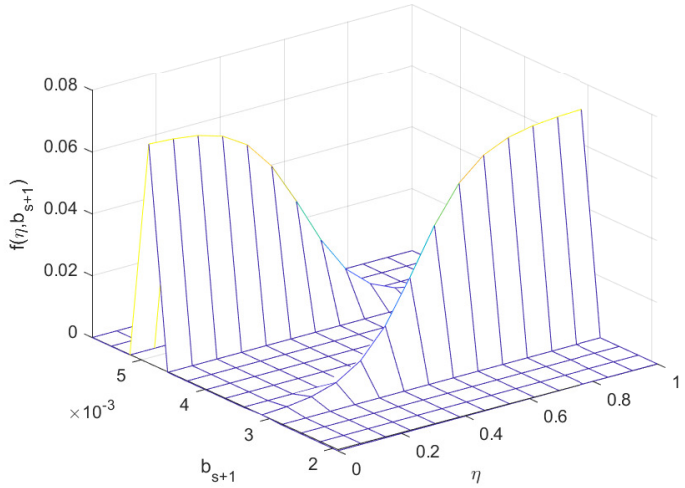


Figure C.2: Decision rule for $\Pi = 1.005$, $b_s = 0$, with a uniform prior belief.

Households limit themselves to a discrete number of savings choices. As the hiring rate rises, the probability that the household gets a signal that η is high rises. It therefore becomes more likely that they will choose the low level of savings in their optimal menu. This decision rule varies with the household wealth. The more savings a household has, the lower the marginal cost to saving further, as there are diminishing marginal returns to consumption in each period.

As in section 3.2.8 for the static model, we now allow households to update their beliefs over time using the simple rule:

$$g(\eta_{t+2}|b_{i,t+1}) = \rho f(\eta_{t+1}|b_{i,t+1}) + (1-\rho), \quad \rho \in (0,1) \tag{C.24}$$

That is, household i combines their posterior belief about the hiring rate (from information collected in period t which informed their savings choice $b_{i,t+1}$) with their original uniform prior, weighting the posterior by ρ and the uniform belief by $1-\rho$. A household with a uniform prior and zero wealth in period t has the decision rule plotted as an example in figure C.2, so has two possible realisations of $b_{i,t+1}$. This implies that there are two posteriors they could form in period t, one for each savings choice. This leads to two possible prior beliefs at the start of period $t+1$.

We solve the household problem using each of these two prior beliefs, and find the prior beliefs households could hold in period $t+2$. Iterating this process we find that there are two groups of priors[4] that emerge in steady state: households are divided into optimists, who hold prior beliefs biased towards a high η, and pessimists, whose beliefs are biased towards low η. As discussed in section 3.2.8, the discrete number of household 'types' is a consequence of the discreteness of the optimal decision rule. As in the static model, our choice of information cost implies that decision rules have two discrete levels of saving, which leads to there being just two groups of priors in steady state. The two types of prior belief present in steady state are plotted below. The prior belief plotted in blue is that of an optimist: a household who received a signal in the previous period that the hiring rate was high, and so their belief is biased upwards this periods. Similarly, pessimists hold the prior belief in red, which is biased towards a low hiring rate because in the previous period that household received a signal that the hiring rate was low. In steady state the proportions of optimists and pessimists in the population will be constant, but individual households will transition from one to the other each period.

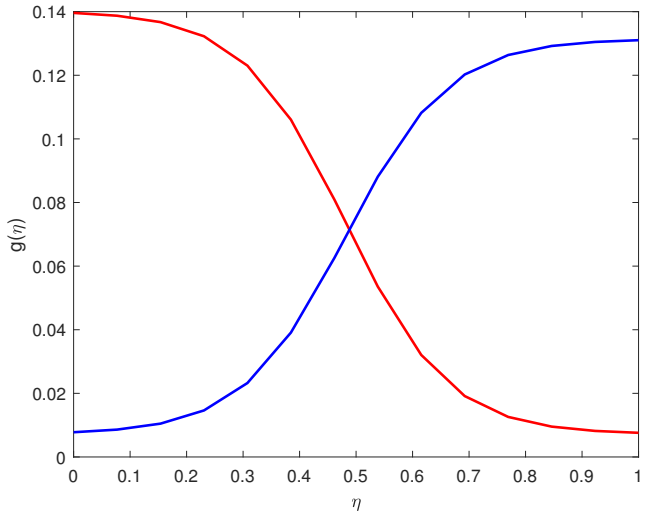
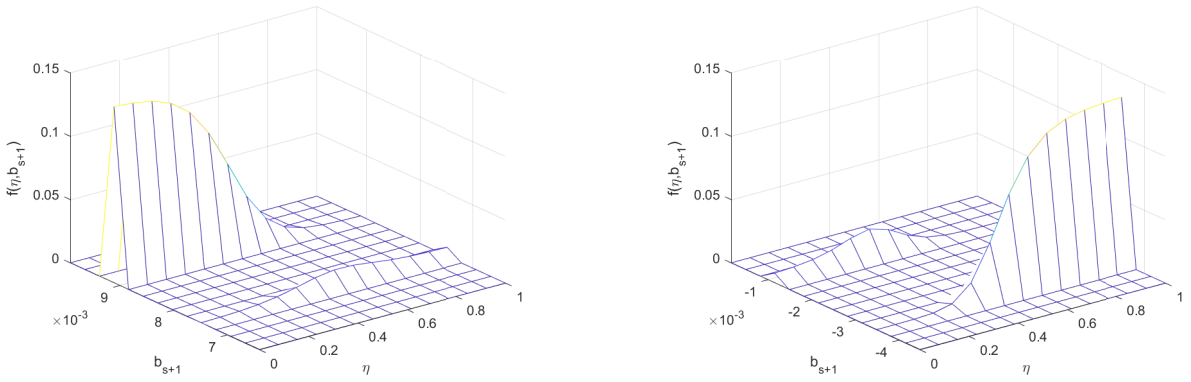


Figure C.3: Prior beliefs seen at steady state with $\psi = 0.0025$, $\rho = 0.9$

Figure C.4 plots the decision rules of a household with $b_s = 0$ with each prior.

[4]We iterate the RI problem until we obtain a set of prior beliefs such that each prior and subsequent information processing implies next-period priors which approximately equal another element of the set. It is in this sense that we obtain two 'groups' of priors rather than two exact priors. The priors in figure C.3 are examples from these groups.

Figure C.4: Decision rules for $\Pi = 1.005$, $b_s = 0$.



(a) Pessimistic prior belief.

(b) Optimistic prior belief.

If we left the model as it is here, wealthier households would process somewhat less information than poorer households as they face lower utility costs of mistakes. While all households with the wealth levels seen in steady state in our calibration still choose a signal with two possible realisations, the extent to which beliefs are updated from their priors, and so the shape of the prior beliefs in the next period, would therefore depend on the household's wealth history as well as their previous signals. In finding the steady states of the simple model in section 3.2 we used the result that all households receiving a high signal in one period had the same prior in the next period: that would no longer be true in this model.

To make progress in solving for the stable wealth and belief distributions, we therefore assume that wealth can only enter household decisions through the household *action strategy*, not through their *information strategy*. That is, all households choose signals as if they had zero assets. When they observe a realisation of the signal, however, they respond to their newly updated hiring rate expectations taking account of their wealth. Households that start a period with more wealth will choose to save more (in absolute terms) than a poorer household with the same beliefs.

Intuitively, we think of a household contracting their information processing to some outside agent. The information processor is given the household's prior beliefs, and the marginal cost of information, but not the household's wealth. The processor receives a signal and forms a posterior $\hat{\eta}$, which it reports back to the household. The household then decides how much to save given that posterior estimate. In this way wealth does not enter the information strategy (the processor's signal collection) but it does enter the action strategy (the household savings choice).

This does not mean that wealth and information are independent: households who receive a signal that η is low will save a lot, and so will become wealthier. They will

have prior beliefs in future periods which are biased towards low η. There is therefore feedback from signals and prior beliefs to wealth. The mechanism that our restriction removes is that wealth, in turn, affects the optimal amount of information processed, and so affects beliefs in future periods. This link from wealth to information is small: without the restriction on information strategies the expected hiring rate after receiving a high signal from a high prior, for example, is 0.780 for a household with a starting wealth of -0.03, and is 0.765 for a starting wealth of 0.03, implying that the signals themselves do not differ very much by wealth, even though actions in response to those signals do. This wealth range covers 99% of the asset holdings seen in equilibrium over the 10000 period simulation of the model with mutation shocks discussed in the following appendix C.3.6.

C.3.6 Transitions

At the individual level there are transitions between different prior beliefs every period, even when the aggregate economy is in steady state. As in section 3.2.8, in steady state households have either optimistic or pessimistic prior beliefs, and while the proportions of these are constant at each of the steady states individual households churn between the two 'types' of priors. The tables below give the transition probabilities between the two prior beliefs at each steady state:

Table C.1: Prior belief transition matrices

(a) Unemployment trap

	$\mathbf{P}_{s+1}$	$\mathbf{O}_{s+1}$
$\mathbf{P}_s$	0.983	0.017
$\mathbf{O}_s$	0.941	0.059

(b) Middle steady state

	$\mathbf{P}_{s+1}$	$\mathbf{O}_{s+1}$
$\mathbf{P}_s$	0.854	0.146
$\mathbf{O}_s$	0.277	0.723

(c) Employment trap

	$\mathbf{P}_{s+1}$	$\mathbf{O}_{s+1}$
$\mathbf{P}_s$	0.200	0.800
$\mathbf{O}_s$	0.002	0.998

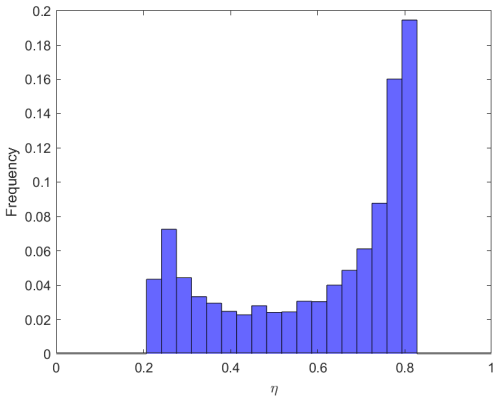
At the aggregate level, mutations to the distribution of prior beliefs, as studied for the simple model in section 3.2.9, can shift the economy between steady states. To see this, we run a simulation of the model with mutations. At the start of each simulation period, we hit the distribution of prior beliefs with two 'mutation' shocks. We take a fraction m_p of households with pessimistic priors and mutate their beliefs so they become optimistic (drawn at random, so over the large population of households we take the same proportion of the pessimistic households from each point in the wealth distribution). Simultaneously, we take a proportion m_o of households with optimistic priors and make them pessimistic.

We draw m_p, m_o from an i.i.d. truncated generalised Pareto distribution (location=0, scale=0.1, shape=25, truncated to be in the range $[0, 1]$). We use a fat-tailed distribution to ensure that, on occasion, the system does transition from one steady state to another

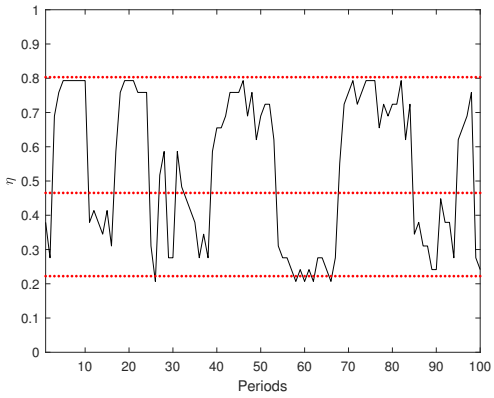
- but that most of the time shocks will be small enough for the economy to remain in the neighbourhood of the steady state.

Over many periods of random mutation shocks, the hiring rate displays a bimodal distribution, spending most of the time close to the high employment steady state, but occasionally transitioning and spending a substantial number of periods around the unemployment trap before a sufficiently large mutation shifts the economy back to the high steady state. This is shown in figures C.7a and C.7b .

Figure C.6: Results from hiring rate simulations (details in appendix C.3.6)



(a) Histogram of realisations of the hiring rate over a 10000 period simulation.



(b) The hiring rate over simulation periods 1-100. Red dotted lines show η steady states.