Cyclical Attention to Saving*

Alistair Macaulay†

September 3, 2021

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 20%. Policies that reduce the costs of comparing between financial products have substantial stabilization effects.

JEL codes: D83, E32, E71

*I thank Klaus Adam, Julio A. Blanco, Vasco Carvalho, Wei Cui, Martin Ellison, Jesus Fernandez-Villaverde, Ina Hajdini, Federico Kochen, Tobias König, John Leahy, Sang Lee, Filip Matějka, Michael McMahon, Christopher J. Palmer, Luigi Pistaferri, Ricardo Reis, Matthew Shapiro, Boromeus Wanengkirtyo, Mirko Wiederholt, and seminar and conference participants at the 3rd Annual NuCamp Conference, Applied Young Economist Webinar series, Bank of England, EEA Annual Congress 2020, Econometric Society World Congress 2020, Economics Graduate Student Conference 2020, Expectations in Dynamic Macroeconomic Models 2021, MMF Annual Conference 2021, Royal Economic Society Symposium of Junior Researchers 2021, University of Birmingham, University of Michigan, University of Oxford, and the Young Economists Symposium 2020 for helpful comments and suggestions. I am particularly grateful to Richard Harrison for kindly sharing the code for Harrison and Oomen (2010), and to Moneyfacts for permission to use their data.

†Department of Economics & St. Anne's College, University of Oxford. Email: alistair.macaulay@economics.ox.ac.uk

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 (Financial Conduct Authority, 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 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 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 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.[1]

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.[2]

The model predicts that savers should experience higher interest rates relative to this distribution of offered rates in contractions, as they increase 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, 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 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.

[1]The mechanism through which variable attention to savings products affects the business cycle does not necessarily apply in the same way to loans (see Appendix A.1 for a discussion).

[2]The UK financial 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).

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 (e.g. McKay, 2013; Kaplan and Menzio, 2016) mostly have households engaging in costly search following Burdett and Judd (1983), which outside of simple cases are not usually tractable enough to estimate. I retain many of the qualitative features of the Burdett-Judd model,[3] 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, as in the simple model.[4] This effect is substantial: the consumption response in the estimated model (cumulated over a year) to risk premium and TFP shocks is 31% and 20% 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 12%. 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 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. Compared with an estimated full-information version of the model, risk premium shocks in the baseline model account for 25% and 35% less of the variance of output and consumption respectively. The extra volatility is largely attributed to supply shocks, notably TFP and price markup shocks. There is also a greater role given to government spending shocks.

Related Literature. There is a large literature studying how information frictions affect the business cycle. Many of these papers study frictions in the information agents

[3]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.

[4]A small number of shocks have large effects of on interest rate dispersion that just outweigh the marginal utility effect, in which cases attention falls with consumption, weakening the shock. None of these shocks account for a substantial share of consumption and output volatility.

receive about continuously distributed exogenous shocks (e.g. Maćkowiak and Wiederholt, 2015),[5] or about the reaction functions of other agents and the relationships of endogenous variables to shocks (e.g. Eusepi and Preston, 2011).[6] Unlike these papers, the friction I study is over the discrete choice of which bank to use for saving each period.

Similar frictions have been studied in a wide range of papers on the role of information and inattention in portfolio choice. A literature starting with Arrow (1987) finds that information frictions are an important determinant of wealth inequality, as wealthier households optimally process more information about saving and investment choices, so make better choices and earn higher rates of return on average.[7] In a companion paper, I study the implications of this for the transmission of fiscal policy (Macaulay, 2021). Rational inattention can also account for several other observed features of portfolio choices, such as home bias and under-diversification (Van Nieuwerburgh and Veldkamp, 2009, 2010), and contagion between markets with unrelated fundamentals (Mondria and Quintana-Domeque, 2013). My focus is on cyclical changes in information processing, which also feature in Kacperczyk et al. (2016) and Rachedi (2018). I extend this literature by showing that cyclical changes in information processing can feed back into the business cycle, amplifying the effect of shocks to consumption.

Specifically, I model the information friction in deposit markets as a discrete choice rational inattention problem, drawing on Matějka and McKay (2015). This form of inattention has been used to study import decisions (Dasgupta and Mondria, 2018), hiring (Acharya and Wee, 2020), and capacity utilization (Sun, 2020).[8]

Another way of modelling the friction in financial product choice would be to use costly search or shopping effort. Coibion et al. (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, since unemployed households search harder for low goods prices, average search effort rises in recessions (Kaplan and Menzio, 2016). 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.[9]

[5]In most such models tracking exogenous or endogenous variables are equivalent as agents can perfectly map between them. For a review of several of these models see Hubert and Ricco (2018).

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

[7]Campanale (2007), Kacperczyk et al. (2019), Lei (2019) (among others) find that this is quantitatively important in explaining observed features of the wealth distribution over time.

[8]See Maćkowiak et al. (2020) for a review of this literature. The model in Sun (2020) is similar to mine, in that buyers are inattentive to their goods choices. However, in equilibrium there is no price dispersion, and so attention is always zero in his model. The variation in attention and the reaction of the equilibrium price distribution in this paper is novel, to the best of my knowledge.

[9]McKay (2013) also studies a model of search for higher interest rates, but he does not consider how

I also contribute to the literature on 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 et al. (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.

Yankov (2018) finds that search (or information) frictions can explain this limited pass-through in the US market for certificates of deposit, using a model based on Burdett and Judd (1983). While Yankov (2018) uses data similar to the Moneyfacts data used in this paper, I differ from his work in combining that with data on how savers choose between products, and how their attention behaviour interacts with the business cycle. Evidence of substantial inattention in retail financial markets can also be found in Martín-Oliver et al. (2009), Branzoli (2016), Deuflhard et al. (2019), and Adams et al. (2021) (among others). I extend this literature by studying how that inattention varies over the business cycle, and showing the macroeconomic consequences of that variation.

I also contribute to the literature on the drivers of the business cycle, by showing that countercyclical attention provides a structural interpretation for a portion of the risk premium shock that is commonly found to be important in estimated macroeconomic models (e.g. Smets and Wouters, 2007; Christiano et al., 2015).[10] Attention, however, is not exogenous, but is a response to other variables.

The rest of the paper is organised as follows: I develop a partial equilibrium model of rationally inattentive households interacting with heterogeneous banks in Section 2. In Section 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 4. In Section 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 2. Section 6 concludes.

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 infor-

this changes over the business cycle. In Appendix A.2 I show that a model of endogenous search effort gives the same qualitative implications as in the main body of the paper. The extension of persistent interest rates (Appendix C.1) would however be intractable in this alternative setup.

[10] Chari et al. (2009) argue that the lack of a clear structural explanation for this shock is a weakness of the Smets-Wouters model, though others (e.g. Fisher, 2015) have provided theoretical interpretations.

mation about which of a finite set of banks is offering the best interest rates 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.

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 5.

Bank n chooses the interest rate they offer to individuals i_t^n to maximise profits, taking into account that their market share will depend on how their interest rate compares with the distribution of rates offered by the other banks.[11] Denoting the probability a saver chooses bank n for a given interest rate distribution as $\Pr(n|i_t^n, i_t^{-n})$, the bank problem is:

$$i_t^n = \arg\max_{\hat{i}_t^n} \Pr(n|\hat{i}_t^n, i_t^{-n}) \cdot (i_t^{CB} - \hat{i}_t^n - \chi_t^n) \tag{1}$$

This gives the first order condition:[12]

$$\frac{d}{di_t^n} \Pr(n|i_t^n, i_t^{-n}) \cdot (i_t^{CB} - i_t^n - \chi_t^n) = \Pr(n|i_t^n, i_t^{-n}) \tag{2}$$

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. I assume that these costs are random variables, but for the moment do not place any structure on their distribution. I refer to a realized set of costs in period t as the state of the world s_t.

[11] All individuals are identical, and will choose one bank each per period, so the market share equals the probability a saver chooses that bank.

[12] The market share function $\Pr(n|i_t^n, i_t^{-n})$ is derived in Section 2.2. It is smooth as long as savers have less than full information, and is such that equation (2) is sufficient for profit maximisation.

2.2 Households

Each period households choose their consumption, and how much attention to pay to choosing between the different banks. Attention increases the expected 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 information processing, in this case about which banks are offering the highest interest rates that period. I assume the household is a net saver, so prefers to choose banks offering higher rates. All asset income is redistributed among individuals each period, so there is no inequality within the household. The household problem is:

$$\max_{c_t, b_t, \mathbb{E}_s i_t^e} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \big(u(c_t) - \mu \mathcal{I}(\mathbb{E}_s i_t^e)\big) \tag{3}$$

subject to

$$c_t + b_t = b_{t-1}(1 + i_{t-1}^e) + y_t \tag{4}$$

$$\mathcal{I}'(\mathbb{E}_s i_t^e) > 0, \mathcal{I}''(\mathbb{E}_s i_t^e) > 0 \tag{5}$$

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

The novel element of this problem is the term $\mathcal{I}(\mathbb{E}_s i_t^e)$, the amount of attention required for the household to earn an expected effective interest rate $\mathbb{E}_s i_t^e$ on assets bought in period t (which pay off in $t+1$), which will be derived below.[13] The key properties of this function are expressed in condition (5): if the household pays more attention they will earn a higher expected 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 e.g. Kamdar, 2019; Maćkowiak et al., 2020). This can be thought of as costly cognitive effort. In Appendix A.2 I show that a monetary cost leads to the same qualitative conclusions.

In the maximisation I allow the household to directly choose the expected effective

[13] The expectation is taken over states of the world s. For a given state of the world, the assumption of a large household of many individuals ensures that there is no uncertainty in the effective interest rate.

interest rate they face. This is equivalent to choosing the amount of attention to pay as there is a one-to-one mapping between the two variables (see Appendix B.2). The first order conditions comprise an Euler equation and a first order condition on the effective interest rate:

$$u'(c_t) = \beta \mathbb{E}_t(1 + i_t^e)u'(c_{t+1}) \tag{6}$$

$$\beta b_t \mathbb{E}_t u'(c_{t+1}) = \mu \mathcal{I}'(\mathbb{E}_s i_t^e) \tag{7}$$

The first order condition on effective interest rates (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 ($\mathcal{I}'(\mathbb{E}_s i_t^e)$) 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, so attention rises with rate dispersion.

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. This encourages further saving through the Euler equation (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 B.1).[14] Since households are net savers (government bonds are in positive net supply), $b_t > 0$ and the household always chooses to process some information.

I now turn to the derivation of $\mathcal{I}(\mathbb{E}_s i_t^e)$, from the decisions of individuals, who face a discrete choice Rational Inattention problem as studied in Matějka and McKay (2015).[15]

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.

[14]The interaction between attention and wealth implies that the model has two steady states, one with identical households 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 4 is only informative about average household choices, I study the model with identical households. See Macaulay (2021) for analysis of the two-agent steady state in a related model.

[15]The redistribution of asset income around the household renders individuals risk neutral with respect to interest rates, and so there is no incentive for them to diversify beyond a single bank. I therefore proceed within the discrete choice framework. Risk neutrality over interest rates also implies that the objective function in this Rational Inattention problem is simply the expected interest rate.

I assume that individuals share information on returning to the household at the end of the period, so all individuals have the same priors. Let $\mathcal{P}_{n,t}$ denote the probability that an individual chooses bank n if they process no information and rely only on their priors. Following Steiner et al. (2017) I refer to this 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 identify the best product in the market with certainty, 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 joint distribution of a noisy signal and the true distribution of interest rates among 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 entropy between the prior and posterior beliefs about the state of the world from observing a realisation of that signal.

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:

$$\max_{\Pr(n|s_t)} \mathbb{E}_{s_t}\Big(\sum_{n=1}^{N} i_{s_t}^{n} \Pr(n|s_t)\Big) \text{ subject to} \tag{8}$$

$$\mathcal{I}_t = -\sum_{n=1}^{N} \mathcal{P}_{n,t} \log(\mathcal{P}_{n,t}) + \mathbb{E}_{s_t} \sum_{n=1}^{N} \Pr(n|s_t) \log(\Pr(n|s_t)) \tag{9}$$

[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.

$\Pr(n|s_t)$ 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 positions of banks within it. The individual chooses a decision rule (a set of conditional choice probabilities for each possible s_t) to maximise their expected interest rate. They maximise subject to the constraint that $\Pr(n|s_t)$ cannot deviate too far from the predisposition $\mathcal{P}_{n,t}$. 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:

$$\Pr(n|i_t^n, i_t^{-n}) = \frac{\mathcal{P}_{n,t}\exp(\frac{i_t^n}{\lambda_t})}{\sum_{k=1}^N \mathcal{P}_{k,t}\exp(\frac{i_t^k}{\lambda_t})} \tag{10}$$

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 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.

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:

$$i_t^e = \sum_{k=1}^N i_t^k \Pr(k|i_t^k, i_t^{-k}) \tag{11}$$

Substituting out for the optimal conditional choice probabilities using equation (10), this becomes:

$$i_t^e = \frac{\sum_{k=1}^N i_t^k \mathcal{P}_{k,t}\exp(\frac{i_t^k}{\lambda_t})}{\sum_{k=1}^N \mathcal{P}_{k,t}\exp(\frac{i_t^k}{\lambda_t})} \tag{12}$$

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, which increases with attention $\mathcal{I}$. Therefore $\mathcal{I}'(\mathbb{E}_s i_t^e) > 0$. Diminishing returns to attention ensure that $\mathcal{I}''(\mathbb{E}_s i_t^e) > 0$.

In Appendix B.2 I show that $\mathcal{I}'(\mathbb{E}_s i_t^e) = \lambda_t^{-1}$. That is, the information required for a marginal increase in the expected effective interest rate is the inverse of the shadow value of information in the individual problem. This highlights the role of interest rate dispersion: if interest rates get more dispersed, then information is more valuable to the individual, as small improvements in the probability of choosing higher interest rate banks have a larger effect on the expected interest rate. If attention $\mathcal{I}$ is held constant, then the shadow value of information will rise to reflect this. That, in turn, leads to a lower $\mathcal{I}'(\mathbb{E}_s i_t^e)$, and so to an increase in attention (absent other changes in the attention first order condition).[17]

The predispositions are where this Rational Inattention model allows for more flexibility than the search-based models explored in Appendix A.2. 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 $\mathcal{P}_{n,t}$. 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 within the rate distribution, 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, which would be necessary to account for these situations, models of this type quickly become intractable.[18]

I study the case of persistence in bank costs, and so in the positions of banks within the interest rate distribution, in Appendix C.1. For the modelling in the main body of the paper however I assume for simplicity that no bank has more brand power than any other, and that bank costs have no persistence, which implies that there is no persistence in the ranking of banks within the rate distribution.[19] The predispositions then all equal $1/N$, and the conditional choice probabilities and effective interest rate become:

$$\Pr(n|i_t^n, i_t^{-n}) = \frac{\exp(\frac{i_t^n}{\lambda_t})}{\sum_{k=1}^N \exp(\frac{i_t^k}{\lambda_t})} \tag{13}$$

$$i_t^e = \frac{\sum_{k=1}^N i_t^k \exp(\frac{i_t^k}{\lambda_t})}{\sum_{k=1}^N \exp(\frac{i_t^k}{\lambda_t})} \tag{14}$$

[17] See Appendix B.2 for a proof that $\mathcal{I}''(\mathbb{E}_s i_t^e) > 0$. Appendix B.3 shows that higher dispersion implies a higher λ_t. Both proofs are for the case of uninformative priors used from this point on.

[18] See e.g. Menzio and Trachter (2015) for an example.

[19] There is in fact very little persistence in bank interest rate rankings for the products studied in Section 4 (see Appendix C.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.

Finally, I assume that the distribution of bank costs χ_t^n is the same each period, with the only variation being in which bank draws which cost. This ensures that the effective interest rate i_t^e is unaffected by which state of the world is realized.[20]

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 (6), the first order condition on attention (7), and the bank profit maximisation condition (2). The bank condition was left in Section 2.1 in terms of the probability of savers choosing each bank. Substituting in the conditional choice probabilities from equation (13) this becomes:

$$\big(1 - \Pr(n|s_t)\big) \cdot (i_t^{CB} - i_t^n - \chi_t^n) = \lambda_t \tag{15}$$

Similarly, we can write the household first order condition on attention as:

$$\beta b_t \mathbb{E}_t u'(c_{t+1}) = \mu \lambda_t^{-1} \tag{16}$$

The bank first order condition (15) implies that when attention rises, the distribution of interest rates shifts up (see Appendix B.4). Intuitively, higher 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.[21]

Therefore if a shock causes consumption to fall, households increase attention, which means they experience higher interest rates than if attention had remained constant

[20]This simplifies the household problem as the expectations operator within $\mathcal{I}(\mathbb{E}_s i_t^e)$ becomes redundant. We can think of this assumption as being that each period, a ranking of banks is drawn from an i.i.d. distribution, and then that bank costs are a deterministic function of these rankings.

[21]In Appendix B.4 I also show that interest rate dispersion falls when attention rises with $N = 2$ banks in the market. If we additionally impose an interest rate lower bound (akin to the reservation price in Burdett and Judd (1983)), then at very low levels of attention price dispersion is increasing in attention as rates begin to rise above the bound. Numerically, the same is true for $N > 2$ banks. Qualitatively, a rise in attention therefore has the same effect as a rise in search effort in Burdett and Judd (1983).

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 5 I find that this is rare, so on average variable attention amplifies the consumption effect of shocks.

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.[22] 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.

3.1 Data Sources

Each month Moneyfacts magazine publishes tables of the interest rates and product characteristics of the vast majority of saving and credit products on offer from retail financial institutions in the UK.[23] 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 remaining after controlling for 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

[22] The editions from January 2008, December 2008, and February 2009 were missing from the library collection at the University of Oxford when this research took place, so data from these months is missing. Where HP-filtered series are used below, I fill in the missing data by linearly interpolating between the months either side, then drop the interpolated observations after the series has been filtered.

[23] 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.

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. In addition, it is common for these products to come bundled with offers for current accounts and other financial services, so the headline interest rates may not accurately capture the value of each product. Further details on fixed interest rate savings products are given in Section 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 major dimensions of product heterogeneity identified in Moneyfacts, so it 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.[24] 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. In addition, this is one of the most common combinations of product features in the market, so my results in Section 4 are less affected by outliers than would be the case with a more niche combination of product features.

A limitation of the Quoted Household Interest Rate data is that the interest rates on qualifying products are weighted imperfectly. The ideal measure of the average rate achieved by households in these products would weight each bank's interest rate by the amount of new deposits that month in that product.[25] However, in the absence of deposit data by product, the Bank of England instead weights each interest rate by deposit inflows

[24]The only characteristics reported in Moneyfacts that differ among products in the Quoted Household Interest Rate are the penalty for withdrawing deposits before the end of the term, and whether the product is managed through a branch, by post, telephone or the internet (online-only products are excluded from the Quoted rate before 2009). The Financial Conduct Authority (2015) found that holders of fixed-rate bonds did not place much importance on these product characteristics, mostly valuing products based on their interest rate and term.

[25]If a bank has multiple products that qualify for the Quoted Household Interest Rate, the average only considers the one with the highest interest rate. I do the same when identifying the relevant set of products in the Moneyfacts data.

per bank and month across the somewhat broader set of all fixed-rate bonds with a term less than or equal to one year. While this implies that the Quoted Household Interest Rate is not a precisely quantity-weighted average, I show in Appendix D.1 that a bank's position in the distribution of interest rates on products qualifying for inclusion in the Quoted Household Interest Rate is very highly correlated with their position in the other market segments used in the weighting scheme. The countercyclical pattern of the Quoted Household Interest Rate relative to the distribution of interest rates in that set of products found in Section 4 therefore reflects a systematic shift towards banks that are more competitive across these market segments in recessions. Although imperfect, I therefore continue to refer to the Quoted Household Interest Rate as the average interest rate achieved by households. The measurement errors on the savings data in the quantitative model in Section 5 are included partly to reflect this limitation.

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.[26] 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 4.1.1). The largest four institutions had 74% of the market for current accounts in 2000, and the largest branch networks (Vickers, 2011). 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. In 2013, 12% of households held these products, and they accounted for 20% of all cash savings balances in the UK (Financial Conduct Authority, 2015). In the Moneyfacts data there are an average of 200 such products available each month in the sample. The mean number of products satisfying the criteria for inclusion in the Quoted Household Interest Rate is 34.

There are two other factors which aid analysis of choices in this particular market. First, product bundling is uncommon. In the median month, just 3.5% of products qualifying for inclusion in the Quoted Household Interest Rate are explicitly bundled with other products at the offering bank. 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

[26]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 (e.g. Hannah Maundrell, quoted in Hannah, 2017). As the degree of wholesale funding could be related to bank risk, I discuss this in Section 4.1.1.

Household Interest Rate data, but removing them does not substantially change the distribution of offered rates. Savers also do not appear to value having these accounts with the same institution as their other financial products, which might give rise to an implicit bundling of products. The Financial Conduct Authority (2015) found that 76% of savers using fixed rate bonds use an institution which is not their 'main provider of financial services'.

Second, the interest rate is the most important product feature for the large majority of savers in this market (Financial Conduct Authority, 2015). 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 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 (Financial Conduct Authority, 2015).

4 Empirical Results

In this section I explore household choice using the datasets described in Section 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 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.

4.1 Interest Rate Dispersion

Each month in the sample, households could achieve a wide range of interest rates by choosing different savings products from different providers. However, some of this dispersion is due to the fact that savings products differ on dimensions other than their interest rate, such as their duration and eligibility criteria. In Section 4.1.1 I show that in fact interest rate dispersion remains substantial even among products which are close substitutes. 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 4.1.2. 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.

4.1.1 Interest Rate Dispersion is not explained by Product Differentiation

To explore whether product heterogeneity can account for the differences in interest rates observed on fixed interest rate savings products, I study a set of products which are extremely close substitutes. If the market is perfectly competitive, and unobserved product heterogeneity is small, these products should all have similar interest rates. This is not what is observed: interest rates are substantially dispersed even among similar products.

To obtain this set of close substitutes I focus on the products used to compute the Quoted Household Interest Rate series, which have the same term, investment size, and interest rate payment frequency. This covers the major dimensions of product heterogeneity in this market (see Section 3.1), and yet the median within-month standard deviation of interest rates on these products is 45 basis points, on an average interest rate of 518 basis points. In October 2000, as an example, savers could earn annual rates of return between 450 and 680 basis points depending on their choice of bank (the standard deviation of rates that month is 44 basis points). The histogram of these rates is plotted in Figure 1. There is therefore substantial interest rate dispersion which cannot be explained by observable product differentiation.

Missing out on 45 basis points on the £5000 investments in these products only implies an annual loss of £22.50. However, it is not the magnitude of asset income that matters for intertemporal consumption decisions in standard macroeconomic models, and the models in Sections 2 and 5 of this paper. Rather it is the interest *rate*, and 45 basis points is large in terms of typical interest rate changes, for example stemming from monetary policy decisions. In fact, the small monetary loss helps explain why savers do not pay much attention to their product choice from this set.

This exercise, however, only controls for observable product heterogeneity. While I can discount many possible dimensions of unobserved heterogeneity (see Section 3.2), there could still be attributes known and valued by households that differentiate the products on offer.

Bank risk is potentially one such unobserved characteristic that could explain interest

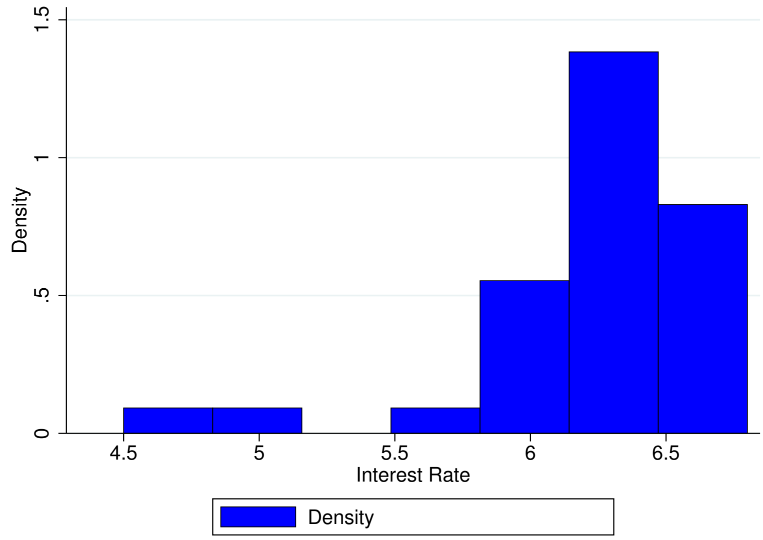


Figure 1: Histogram of annual interest rates on fixed interest rate bonds and term accounts on offer in October 2000.

rate dispersion, if riskier banks offer higher interest rates to compensate savers for their risk. 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, so as long as deposit insurance is credible risk should not affect pricing, as Ben-David et al. (2017) find for the US. Indeed, Chavaz and Slutzky (2020) find that deposit rates in the UK are on average uncorrelated with a variety of measures of bank risk, suggesting that risk is not the main driver of the dispersion found here. This is supported by the fact that regressing the panel of interest rates on bank and month fixed effects still leaves the mean and median unexplained within-month standard deviation of interest rates at 31 and 29 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

[27]This is an inferior way of capturing risk than that of Chavaz and Slutzky (2020), who use time-varying measures of bank risk from the Bank of England, as it ignores changes in bank risk over time. It also removes all variation which causes a bank to offer persistently high or low rates, whether that is driven by risk or not (see Appendix C.1 for an example where information costs imply rate persistence). The fixed-effects regression should therefore be taken as further suggestive evidence that the Chavaz and Slutzky (2020) results apply to the fixed-rate market specifically, as well as to retail deposits in general. While Chavaz and Slutzky (2020) 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 is only significant for variable-rate products.

why interest rate dispersion persists in equilibrium.

4.1.2 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 et al. (2006) for a review of the early literature). The existence of interest rate dispersion not accounted for by observed product 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 (Financial Services Authority, 2000; Financial Conduct Authority, 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 created to help "quickly and easily compare financial products" (Moneyfacts, 2021). 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 (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 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. In 2013, only 35% of savers in fixed-term products consulted a comparison website before choosing their product (Financial Conduct Authority, 2015).

Other authors have also concluded that inattention plays an important role in retail financial product markets. Martín-Oliver et al. (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. (2021) 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 2 can explain the time series variation in how households choose from among the set of offered rates.

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 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 (see Honka et al., 2017, for evidence that these both have large effects on consumer banking choices in the US). Specifically, I construct the benchmark rate by taking the average interest rate on offer from the ‘big four’ banks.[28]

[28]These are Barclays, HSBC, Lloyds, and Royal Bank of Scotland. In 1993 the big four had 48% of the bank branches in the UK. NatWest also has a large number of branches, but it is extremely rare for them to offer a product qualifying for the Quoted Household Interest Rate so I leave them out of the calculation of the benchmark rate.

Throughout the sample period these four banks hold most of the market share in many retail banking markets, and have many more branches than other banks (Office of Fair Trading, 2008). They are particularly dominant in current accounts, which are the key product from which banks cross-sell other services, such as savings accounts (Cruickshank, 2000). Using this as the benchmark interest rate assumes that households paying no attention to their choice 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 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 Rate, 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. 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 interest rate 50 basis points above the benchmark rate.

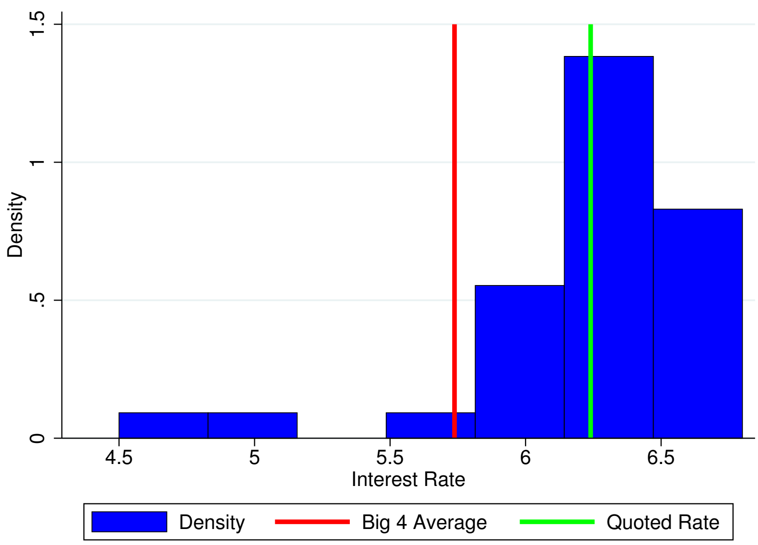


Figure 2: Histogram of annual interest rates on fixed interest rate bonds and term accounts on offer in October 2000, with the average interest rate at the big four (5.74%) in red and the Quoted Household Interest Rate (6.24%) in green. The highest rate on offer from a big four bank was 6.1%.

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:

$$\varphi_t := \frac{\mathbb{E}_h i_t - i_t^b}{\sigma(i_t)} \tag{17}$$

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. 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 to attention in the model in Sections 2 and 5.[29] In Section 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.

I will therefore interpret φ as a proxy for attention. To combat the concern that changes in φ could be driven purely by shifts in the position of the big four within the interest rate distribution rather than by household behaviour, the analysis in Section 4.3 and Appendix D uses the residual of φ after regressing it on a measure of the position of the big four within the rate distribution.[30] All of the results remain significant and of the same sign as those below if the raw values of φ are used instead. This is because the majority of the variation in φ does not come from movements in the position of the big four: the R^2 of the regression of φ on that position is 0.23.

4.3 Cyclicality of φ

Since φ can be measured each month, I can study choice behaviour at a high enough frequency to observe co-movements with aggregate variables over the business cycle. I find that φ is countercyclical.

[29] In Appendix B.5 I show that there is an exact correspondence between φ and attention in the model with two banks and uninformative priors, and that attention and φ remain closely related with more banks in the market. Numerically, φ and attention are also closely linked in the model with more general priors in Appendix C.1.

[30] Specifically, the big four position is measured analogously to the position of the Quoted Household Interest Rate: $pos_t^b = \frac{i_t^b - N^{-1}\sum_{n=1}^{N} i_t^n}{\sigma(i_t)}$

Figure 3 shows binned scatter plots of the (HP-filtered) cyclical component of φ against the cyclical components of the average interest rate and in unemployment.[31] Lower interest rates and higher unemployment are associated with higher φ. These relationships are strongly statistically significant.

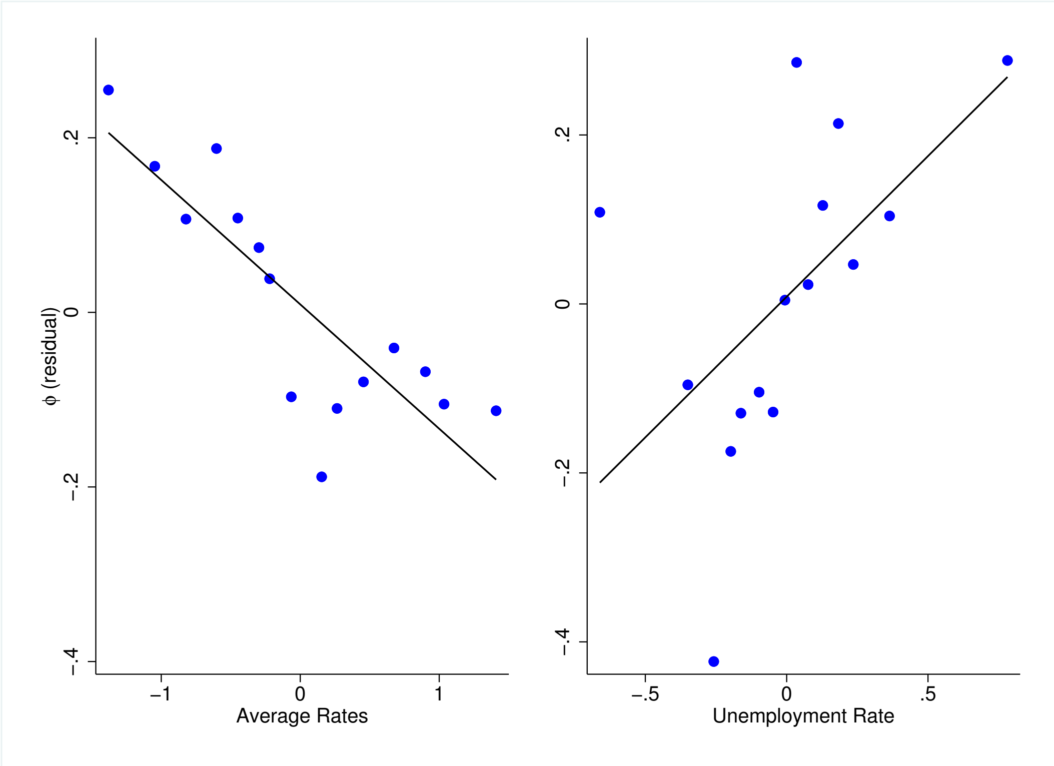


Figure 3: φ against (unweighted) average interest rates among products considered in the Quoted Household Interest Rate data and unemployment. All series are cyclical components after HP filtering. Black solid lines are from linear regressions, which give $\hat{\varphi} = -0.142\hat{\bar{i}}$ (t-statistic on slope coefficient -3.24) and $\hat{\varphi} = 0.333\hat{u}$ (t-statistic on slope coefficient 4.33). Blue circles are means of φ and the regressor of interest within groups of observations, grouped by their position within the distribution of the regressor. For the time series of φ, see Appendix D.2.

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. In Appendix D.3 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,

[31] I use the unweighted mean of the interest rates in the market studied, but this is highly correlated with many other interest rate measures. Using the benchmark rate, the Quoted Household Interest Rate, or the interest rate on one year UK treasury bills makes no difference to the qualitative conclusions, and little quantitative difference.

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 2. In recessions, consumption tends to be low, so the marginal utility of interest income is high, increasing the incentives to pay attention.[32] In addition, when average rates are low in this market the dispersion of interest rates tends to be high, increasing the benefits of attention.[33] 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.[34] In the model in Sections 2 and 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 D.4 I show that changes in the 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.

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 2 is estimated to be powerful. This amplification is substantial, increasing the consumption response to risk premium and TFP shocks (the two shocks explaining the largest fraction of consumption variance) by 31% and 20% 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

[32]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 many shocks consumption and labour supply co-move, in which case this opportunity cost of time channel is qualitatively the same as the marginal utility of income channel.

[33]$Corr(\bar{i}, \sigma(i)) = -0.272$ and is significantly different from 0 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 4.1.1). However, this correlation remains negative and significant if I exclude the crisis periods.

[34]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.

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 12%.

5.1 Model

Since the data in Sections 3 and 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). The log-linearised model equations are in Appendix E.

5.1.1 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 (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 F.

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, is that I use a risk premium shock rather than a discount factor shock, and I assume that the labour disutility shock is i.i.d. (Harrison and Oomen estimate its persistence at 0.001). For further details on the model setup please therefore see Harrison and Oomen (2010).

5.1.2 Attention Block

As in Section 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 bank's problem is as in Section 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 first order conditions is that I replace the nominal policy rate in the consumption Euler equation, the first order condition 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 $\chi_t^g = \zeta_t^\chi$. The other bank draws a high cost, and so I will refer to them as the 'bad' bank (superscript b). They face $\chi_t^b = \chi_1 + \chi_2(i_t^{CB} - \bar{i}^{CB}) + \zeta_t^\chi + \zeta_t^{\chi b}$.[35] 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) shocks ζ_t^χ and $\zeta_t^{\chi b}$ cause exogenous fluctuations in the level and dispersion of bank interest rates. 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 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 (16),

[35] In principle this formulation could allow for $\chi_t^b < \chi_t^g$, if a large enough negative value of $\zeta_t^{\chi b}$ is realised. The quantitative results have χ_1 sufficiently positive that the probability of this is negligible.

with the only differences being that μ is now subject to a mean-zero shock process ζ_t^μ, the stock of saving is set to 1, and the future marginal utility of income is affected by inflation:

$$\beta \mathbb{E}_t \frac{U'(c_{t+1})}{\Pi_{t+1}} = \mu e^{\zeta_t^\mu} \lambda_t^{-1} \tag{18}$$

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:

$$p_t^g = \frac{\exp(\frac{i_t^g}{\lambda_t})}{\exp(\frac{i_t^g}{\lambda_t}) + \exp(\frac{i_t^b}{\lambda_t})} \tag{19}$$

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

$$i_t^e = p_t^g i_t^g + (1 - p_t^g) i_t^b \tag{20}$$

Banks choose interest rates to maximise expected profits. Their first order condition is the same as equation (15) derived in Section 2, which for the good and bad bank respectively reduces to:

$$(1 - p_t^g) \cdot (i_t^{CB} - i_t^g - \zeta_t^\chi) = \lambda_t \tag{21}$$

$$p_t^g \cdot (i_t^{CB}(1 - \chi_2) - i_t^b - (\chi_1 - \chi_2 \bar{i}^{CB}) - \zeta_t^\chi - \zeta_t^{\chi b}) = \lambda_t \tag{22}$$

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 (18), the choice probability rule (19), the definition of i_t^e (20), and the two bank first order conditions ((21) and (22)). There are three new shocks, to attention (ζ_t^μ), the level of bank interest rates (ζ_t^χ) and their dispersion ($\zeta_t^{\chi b}$).

5.2 Estimation

I conduct a Bayesian Maximum Likelihood estimation of the model solved to a log-linear approximation around the zero-inflation steady state. 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 start point coincides with the beginning of the final UK monetary regime identified by Benati (2006).

In addition to the shocks to attention and the level and dispersion of bank costs, I allow for i.i.d. measurement error on each of the newly introduced observables.

I follow Harrison and Oomen in setting some parameters to match standard values or long-run features of UK data. I do the same for χ_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. The only new parameters to estimate are the cost of attention μ, the cyclicality of bank costs χ_2, and the persistence and volatility of the new shocks. μ must be greater than 0, but there are no such restrictions on χ_2. I choose relatively weak priors for both in the absence of strong evidence for the values they should take. For full details of the calibration and priors see Appendix F.

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 χ_2, which have estimated posterior means of 0.035 and -0.264. 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 the most important shocks through the economy. Variable attention therefore amplifies shocks.

Each row of Table 1 reports the magnitude of the cumulative response of consumption to a given 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 that shock in the fixed attention model than with variable attention. I list this for all shocks that explain more than 2% of the variance of consumption, ordered according to the share of consumption volatility they explain.

For most of the shocks, consumption is substantially less responsive when attention is held at its steady state. For risk premium and TFP shocks, which together explain 64% of consumption volatility in the baseline estimated model, attention variation amplifies the consumption response by 31% and 20% respectively. Overall, the variance of consumption is 17% larger with variable attention than if attention is held at steady state.

The intuition is as in Section 2: when a shock causes consumption to fall, the marginal utility of income rises, so attention goes up. More attention increases the effective interest

Table 1: Cumulative consumption response to shocks relative to variable attention baseline.

Shock	Fixed Attention
Risk premium	0.764
TFP	0.837
Govt. spending	0.726
Markup	1.132
Monetary policy	1.124
Bank costs (level)	0.734

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 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 χ_2 is estimated to be negative, this lower policy rate leads to greater interest rate dispersion, encouraging even more attention.

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 some shocks, however, shutting off variable attention leads to larger consumption responses, though none of them play a very large role in consumption fluctuations. The most important is the price markup shock, which accounts for 4% of consumption variance in the baseline model. Variable attention dampens these shocks 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 that cause consumption and interest rates to move in opposite directions, such as TFP, the marginal utility of income effect dominates and attention amplifies the shock. However, for markup and monetary policy shocks there is a large change in policy rates. Attention therefore comoves positively with consumption, dampening the shock.

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.

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.[36] 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.

With no information friction, the risk premium shock explains 53% of the variance of consumption, and 18% of the variance of output. Only TFP shocks explain a larger share of output variance. Moving to the baseline model with inattention the risk premium shock becomes substantially less important, explaining 35% and 13% of consumption and output variance respectively.

Cyclical attention can therefore plausibly explain 25%-35% 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 negligible fractions of consumption and output variance in the baseline model. This portion 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

[36]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.

variation. Full results are in Appendix F.

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 12%.

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.

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 a quarter 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.

References

Acharya, S. and Wee, S. L. (2020). Rational inattention in hiring decisions. *American Economic Journal: Macroeconomics*, 12(1):1–40.

Adams, P. D., Hunt, S., Palmer, C., and Zaliauskas, R. (2021). Testing the effectiveness of consumer financial disclosure: Experimental evidence from savings accounts. *Journal of Financial Economics*, 141(1):122–147.

Adolfson, M., Laséen, S., Lindé, J., and Villani, M. (2007). Bayesian estimation of an open economy DSGE model with incomplete pass-through. *Journal of International Economics*, 72(2):481–511.

Aegon (2017). Only 8% of people speak to an adviser about financial decisions (press release). Technical report.

Allen, J., Clark, R., and Houde, J.-F. (2014). Price Dispersion in Mortgage Markets. *The Journal of Industrial Economics*, 62(3):377–416.

Arrow, K. J. (1987). The demand for information and the distribution of income. *Probability in the Engineering and Informational Sciences*, 1(1):3–13.

Baye, M. R., Morgan, J., and Scholten, P. (2006). Information, Search, and Price Dispersion. In Hendershott, T., editor, *Handbook on Economics and Information Systems, Volume 1*. Elsevier, Amsterdam, Netherlands.

Ben-David, I., Palvia, A., and Spatt, C. (2017). Banks' Internal Capital Markets and Deposit Rates. *Journal of Financial and Quantitative Analysis*, 52(5):1797–1826.

Benati, L. (2006). UK Monetary Regimes and Macroeconomic Stylised Facts. *Bank of England working papers*, 290.

Bhutta, N., Fuster, A., and Hizmo, A. (2020). Paying Too Much? Price Dispersion in the US Mortgage Market. *CEPR Discussion Papers*, 14924.

Branzoli, N. (2016). Price dispersion and consumer inattention: evidence from the market of bank accounts. *Temi di discussione (Economic working papers), Banca d'Italia.*

Burdett, K. and Judd, K. L. (1983). Equilibrium price dispersion. *Econometrica*, 51(4):955–969.

Campanale, C. (2007). Increasing returns to savings and wealth inequality. *Review of Economic Dynamics*, 10(4):646–675.

Chari, V. V., Kehoe, P. J., and McGrattan, E. R. (2009). New Keynesian models: Not yet useful for policy analysis. *American Economic Journal: Macroeconomics*, 1(1):242–266.

Chavaz, M. and Slutzky, P. (2020). Retaining Worried Depositors: Evidence from Multi-Brand Banks.

Christiano, L. J., Eichenbaum, M. S., and Trabandt, M. (2015). Understanding the great recession. *American Economic Journal: Macroeconomics*, 7(1):110–167.

Coibion, O., Gorodnichenko, Y., and Hong, G. H. (2015). The cyclically of sales, regular and effective prices: Business cycle and policy implications. *American Economic Review*, 105(3):993–1029.

Connon, H. (2007). The man who made the tills ring at Moneysupermarket. *The Observer*.

Cook, M., Earley, F., Smith, S., and Ketteringham, J. (2002). Losing Interest: How Much Can Consumers Save by Shopping Around for Financial Products? *FSA Occasional Paper Series*, No. 19.

Cruickshank, D. (2000). *Competition in UK banking : a report to the Chancellor of the Exchequer*.

Dasgupta, K. and Mondria, J. (2018). Inattentive importers. *Journal of International Economics*, 112:150–165.

Deuflhard, F., Georgarakos, D., and Inderst, R. (2019). Financial literacy and savings account returns. *Journal of the European Economic Association*, 17(1):131–164.

Diebold, F. X. and Sharpe, S. A. (1990). Post-deregulation bank-deposit-rate pricing: The multivariate dynamics. *Journal of Business and Economic Statistics*, 8(3):281–291.

Drechsler, I., Savov, A., and Schnabl, P. (2017). The deposits channel of monetary policy. *Quarterly Journal of Economics*, 132(4):1819–1876.

Driscoll, J. C. and Judson, R. (2013). Sticky deposit rates. *Finance and Economics Discussion Series, Board of Governors of the Federal Reserve System*, 2013-80.

Eusepi, S. and Preston, B. (2011). Expectations, learning, and business cycle fluctuations. *American Economic Review*, 101(6):2844–2872.

Eusepi, S. and Preston, B. (2018). The science of monetary policy: An imperfect knowledge perspective. *Journal of Economic Literature*, 56(1):3–59.

Financial Conduct Authority (2015). Cash savings market study report. Technical report.

Financial Conduct Authority (2019). Mortgages Market Study Final Report. Technical Report March.

Financial Services Authority (2000). Better informed consumers. Assessing the implications for consumer education of research by BMRB. (April):1–86.

Finke, M. S., Howe, J. S., and Huston, S. J. (2017). Old Age and the Decline in Financial Literacy. *Management Science*, 63(1):1–278.

Fisher, J. D. (2015). On the Structural Interpretation of the Smets-Wouters "Risk Premium" Shock. *Journal of Money, Credit and Banking*, 47(2-3):511–516.

Hannah, F. (2017). Bank or building society? *The Independent.*

Harrison, R. and Oomen, O. (2010). Evaluating and Estimating a DSGE Model for the United Kingdom. *Bank of England working papers*, 380.

Honka, E., Hortaçsu, A., and Vitorino, M. A. (2017). Advertising, consumer awareness, and choice: evidence from the U.S. banking industry. *The RAND Journal of Economics*, 48(3):611–646.

Hubert, P. and Ricco, G. (2018). Imperfect Information in Macroeconomics. *Revue de l'OFCE*, 137:181–196.

Iscenko, Z. (2018). Choices of dominated mortgage products by UK consumers. *FCA occasional papers in financial regulation*, (33).

Kacperczyk, M., Nosal, J., and Stevens, L. (2019). Investor sophistication and capital income inequality. *Journal of Monetary Economics*, 107:18–31.

Kacperczyk, M., Van Nieuwerburgh, S., and Veldkamp, L. (2016). A Rational Theory of Mutual Funds' Attention Allocation. *Econometrica*, 84(2):571–626.

Kamdar, R. (2019). The Inattentive Consumer: Sentiment and Expectations.

Kaplan, G. and Menzio, G. (2016). Shopping externalities and self-fulfilling unemployment fluctuations. *Journal of Political Economy*, 124(3):771–825.

Kelly, J. (2019). How Monzo is banking on customer apathy. *Financial Times.*

Lei, X. (2019). Information and Inequality. *Journal of Economic Theory*, 184:104937.

Macaulay, A. (2021). The Attention Trap : Rational Inattention , Inequality , and Fiscal Policy. *European Economic Review*, 135:103716.

Maćkowiak, B., Matějka, F., and Wiederholt, M. (2020). Rational Inattention: A Review. *CEPR Discussion Papers*, no. 15408.

Maćkowiak, B. and Wiederholt, M. (2015). Business Cycle Dynamics under Rational Inattention. *Review of Economic Studies*, 82(4):1502–1532.

Martín-Oliver, A., Salas-Fumás, V., and Saurina, J. (2009). Informational differentiation interest rate dispersion and market power. *Applied Economics Letters*, 16(16):1645–1649.

Martinez-Miera, D. and Repullo, R. (2017). Search for Yield. *Econometrica*, 85(2):351–378.

Matějka, F. and McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–298.

McKay, A. (2013). Search for financial returns and social security privatization. *Review of Economic Dynamics*, 16(2):253–270.

Menzio, G. and Trachter, N. (2015). Equilibrium price dispersion with sequential search. *Journal of Economic Theory*, 160:188–215.

Mondria, J. and Quintana-Domeque, C. (2013). Financial Contagion and Attention Allocation. *The Economic Journal*, 123(568):429–454.

Moneyfacts (2021). About us - Info About Moneyfacts. moneyfacts.co.uk/about, accessed 2021-02-04.

Office of Fair Trading (2008). Current account research. Technical report.

Rachedi, O. (2018). Portfolio Rebalancing and Asset Pricing With Heterogeneous Inattention. *International Economic Review*, 59(2):699–726.

Ravn, M., Schmitt-Grohe, S., and Uribe, M. (2006). Deep Habits. *Review of Economic Studies*, 73(1):195–218.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690.

Smets, F. and Wouters, R. (2007). Shocks and frictions in US business cycles: A Bayesian DSGE approach. *American Economic Review*, 97(3):586–606.

Steiner, J., Stewart, C., and Matějka, F. (2017). Rational Inattention Dynamics: Inertia and Delay in Decision-Making. *Econometrica*, 85(2):521–553.

Stigler, G. J. (1961). The Economics of Information. *Journal of Political Economy*, 69(3):213–225.

Sun, T. (2020). Excess Capacity and Demand Driven Business Cycles.

Van Nieuwerburgh, S. and Veldkamp, L. (2009). Information immobility and the home bias puzzle. *Journal of Finance*, 64(3):1187–1215.

Van Nieuwerburgh, S. and Veldkamp, L. (2010). Information Acquisition and Under-Diversification. *Review of Economic Studies*, 77(2):779–805.

Vickers, J. (2011). *Independent Commission on Banking: Final Report: Recommendations.*

Yankov, V. (2018). In Search of a Risk-Free Asset: Search Costs and Sticky Deposit Rates.

A Alternative Mechanisms

A.1 Attention to Loans

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 in the UK market evidence suggests that interest rate dispersion is substantially lower there than it is in savings products. Cook et al. (2002) find that switching from the average to the cheapest 5-year fixed rate mortgage (one of the most common mortgage products in the UK) in December 2000 would have saved a borrower only 20 basis points, despite this category also including products covering a wide range of eligibility criteria and other product features, which Iscenko (2018) finds correlate strongly with headline mortgage interest rates. In contrast, going from the mean to the highest rate in the one-year fixed-rate bonds I study in this paper would have gained a saver 97 basis points that month. More recently, Iscenko (2018) finds that in 2015-2016 30% of those taking out a mortgage chose a product that was strictly dominated by another they were eligible for. The average overpayment within that group was approximately 24 basis points.[37] My data does not extend to 2015, but across my sample a saver could gain an average of 24 basis points by going from the 90th percentile to the best interest rate in the distribution of one-year fixed-rate bonds. The scope for attention to drive interest rate changes in mortgages 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. Consistent with this, the Financial Conduct Authority (2019) found 'high levels of consumer engagement' with mortgage decisions.

It is also extremely rare for a lender to negotiate with a borrower over mortgage rates in the UK (Iscenko, 2018). This may explain some of the difference with Canada and the

[37]The median borrower choosing a dominated product lost out on a 10% reduction in their annual mortgage payments, and the median annual percentage rate (APR) over that period was 2.44%.

US, where Allen et al. (2014) and Bhutta et al. (2020) find more substantial dispersion in mortgage rates, as both find heterogeneity in negotiation is a large driver of rate dispersion. However even in the US, Bhutta et al. (2020) find an average 90-10 mortgage rate gap of 54 basis points over 2015-2019. In my savings data the corresponding gap is (on average) 100 basis points, though this comparison is weak because of the difference in country and time period.

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. While the Moneyfacts data is insufficient to test this in the UK (see Section 3.2), Bhutta et al. (2020) find that willingness to negotiate a lower mortgage rate rises in the US when the level of the interest rate rises, consistent with borrowers being more willing to expend effort to secure a lower rate when their interest expenses are high. In this case, in a contraction we should expect savers to pay more attention and borrowers to pay less, and so for both to face higher interest rates than if attention was constant.

A.2 The core mechanism in alternative models

Here I show that the main mechanism of the inattention model of Section 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). I assume an exogenously fixed distribution of interest rates, since the difficult element of models based on Burdett and Judd (1983) is solving for the equilibrium price distribution. I show in Appendix B.4 that attention affects the equilibrium interest rate distribution in the model of Section 2 in (qualitatively) the same way as search effort affects the equilibrium price distribution in Burdett and Judd (1983).

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.

$$\max_{c_t, b_t, i_t^e} \mathbb{E} \sum_{t=0}^{\infty} \beta^t \big[u(c_t) - (1-\phi)C(i_t^e)\big] \tag{23}$$

subject to

$$c_t + b_t + \phi C(i_t^e) = y_t + b_{t-1}(1 + i_{t-1}^e) \tag{24}$$

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

$$u'(c_t) = \beta(1 + i_t^e)\mathbb{E}_t u'(c_{t+1}) \tag{25}$$

$$\beta b_t \mathbb{E}_t u'(c_{t+1}) = (1-\phi)C'(i_t^e) + \phi u'(c_t) C'(i_t^e) \tag{26}$$

The household problem in Section 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, which will cause current consumption to fall further through the consumption Euler equation, amplifying the shock.[38] This is the mechanism explored in Section 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 $F(i)$. Individ-

[38] 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 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.

uals 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:

$$i_t^e = (1 - \psi(e_t)) \int if(i)di + 2\psi(e_t) \int if(i)F(i)di \tag{27}$$

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 rate the household ends up facing:

$$e_t = \psi^{-1}\left(\frac{i_t^e - \int if(i)di}{2\int if(i)F(i)di - \int if(i)di}\right) \tag{28}$$

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:

$$C(i_t^e) = C_e\left(\psi^{-1}\left(\frac{i_t^e - \int if(i)di}{2\int if(i)F(i)di - \int if(i)di}\right)\right) \tag{29}$$

$$C''(i_t^e) > 0 \text{ if } C_e''(i_t^e) \geq 0 \text{ and } \psi''(e_t) \leq 0, \text{ one inequality strict} \tag{30}$$

B Proofs

B.1 The household first order conditions are sufficient for utility maximisation

Here I show that for plausible parameter values and CRRA utility, the household first order conditions are sufficient for utility maximisation in the simple model (Section 2), and in the quantitative model (Section 5).[39] First, write the household problem as an unconstrained maximisation by substituting out for consumption using the budget constraint:

$$\max_{b_t, i_t^e, X_t} U = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(u\left(\frac{b_{t-1}}{\Pi_t}(1 + i_{t-1}^e) + y_t(X_t) - b_t \right) - \mu \mathcal{I}(i_t^e) + v(X_t) \right) \tag{31}$$

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:

$$H = \left[\begin{array}{cccc} & & & 0 \\ & H_s & & \vdots \\ & & & 0 \\ & & & \frac{\partial^2 U}{\partial b_t \partial i_t^e} \\ 0 & \dots & 0 \; \frac{\partial^2 U}{\partial b_t \partial i_t^e} & \frac{\partial^2 U}{\partial {i_t^e}^2} \end{array} \right] \tag{32}$$

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:

$$xHx' = x_s H_s x_s' + 2yz \frac{\partial^2 U}{\partial b_t \partial i_t^e} + z^2 \frac{\partial^2 U}{\partial {i_t^e}^2} \leq 0 \tag{33}$$

[39] This proof relies on $\mathcal{I}''(\mathbb{E}_s i_t^e) > 0$, which I prove in Appendix B.2 for the model with uninformative priors. I therefore proceed with the assumption of uninformative priors, in which case $\mathbb{E}_s i_t^e = i_t^e$ is independent of the realized state s. Numerically, $\mathcal{I}''(\mathbb{E}_s i_t^e) > 0$ for more general priors, in which case this proof would also hold more generally.

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:

$$\frac{\partial^2 U}{\partial b_t^2} = u''(c_t) + \beta \mathbb{E}_t \frac{u''(c_{t+1})(1+i_t^e)^2}{\Pi_{t+1}^2} < 0 \tag{34}$$

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

$$y^2 \frac{\partial^2 U}{\partial b_t^2} + 2yz \frac{\partial^2 U}{\partial b_t \partial i_t^e} + z^2 \frac{\partial^2 U}{\partial i_t^{e2}} \leq 0 \tag{35}$$

Using the definition of U this condition becomes:

$$\begin{aligned} y^2 u''(c_t) + y^2 \beta \mathbb{E}_t u''(c_{t+1}) \frac{(1+i_t^e)^2}{\Pi_{t+1}^2} + 2yz\beta \mathbb{E}_t u''(c_{t+1}) \frac{(1+i_t^e) b_t}{\Pi_{t+1}^2} \\ + 2yz\beta \mathbb{E}_t u'(c_{t+1}) \frac{1}{\Pi_{t+1}} - z^2 \mu \mathcal{I}''(i_t^e) + z^2 \beta \mathbb{E}_t u''(c_{t+1}) \frac{b_t^2}{\Pi_{t+1}^2} \leq 0 \end{aligned} \tag{36}$$

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:

$$\begin{aligned} \beta \mathbb{E}_t \frac{u''(c_{t+1}) b_t^2}{\Pi_{t+1}^2} \left(\frac{y^2 (1+i_t^e)^2}{b_t^2} + z^2 - 2yz \left(\frac{c_{t+1}\Pi_{t+1} b_t - \gamma(1+i_t^e)}{b_t \gamma} \right) \right) \\ = \beta \mathbb{E}_t \frac{u''(c_{t+1}) b_t^2}{\Pi_{t+1}^2} \left(\frac{y^2 c_{t+1} \Pi_{t+1}}{b_t \gamma^2} \left(2\gamma(1+i_t^e) - c_{t+1}\Pi_{t+1} b_t \right) \right. \\ \left. + \left(z - y \left(\frac{c_{t+1}\Pi_{t+1} b_t - \gamma(1+i_t^e)}{b_t \gamma} \right) \right)^2 \right) \end{aligned} \tag{37}$$

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:

$$-\frac{\beta y^2 b_t}{\gamma} \mathbb{E}_t \frac{c_{t+1}^{-\gamma}}{\Pi_{t+1}} \left(2\gamma(1+i_t^e) - c_{t+1}\Pi_{t+1} b_t \right) \leq 0 \tag{38}$$

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 2 hold as long as this condition is satisfied. In the quantitative model this is easily the case for plausible parameters. There $b_t = 1$, and γ as the inverse of σ^c in Table 11 is estimated to be significantly greater than 1. Since steady state consumption and inflation are 0.662 and 1, and steady state $i_t^e = 1/\beta - 1 = 0.010$, in the region of the steady state this condition is comfortably satisfied.[40] 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.

B.2 Attention costs are increasing and convex

Here I show that $\mathcal{I}'(\mathbb{E}_s i^e) = \lambda^{-1} > 0$, and for the case of uninformative priors that $\mathcal{I}''(i^e) > 0$.

Substituting the optimal choice probabilities into the information constraint 9 gives (dropping time subscripts to simplify notation, as everything here is defined within the same period):

$$\mathcal{I} = \frac{\mathbb{E}_s i^e}{\lambda} - \sum_{s=1}^{S} \Pr(s) \log d_s \tag{39}$$

Where:

$$d_s = \sum_{k=1}^{N} \mathcal{P}_k \exp(\frac{i^{k,s}}{\lambda}) \tag{40}$$

Differentiate this with respect to $\mathbb{E}_s i^e$, holding the offered interest rates $i^{n,s}$ constant as individuals take them as given:

$$\frac{\partial \mathcal{I}}{\partial \mathbb{E}_s i^e} = \frac{1}{\lambda} - \frac{\mathbb{E}_s i^e}{\lambda^2} \frac{\partial \lambda}{\partial \mathbb{E}_s i^e} - \sum_{s=1}^{S} \frac{\Pr(s)}{d_s} \frac{\partial d_s}{\partial \mathbb{E}_s i^e} \tag{41}$$

Each term inside the sum is:

$$\begin{aligned}\frac{\Pr(s)}{d_s} \frac{\partial d_s}{\partial \mathbb{E}_s i^e} &= \frac{\Pr(s)}{d_s} \frac{\partial \lambda}{\partial \mathbb{E}_s i^e} \left[\left(\sum_{k=1}^{N} \exp(\frac{i^{k,s}}{\lambda}) \frac{\partial \mathcal{P}_k}{\partial \lambda} \right) - \frac{1}{\lambda^2} \left(\sum_{k=1}^{N} i^{k,s} \mathcal{P}_k \exp(\frac{i^{k,s}}{\lambda}) \right) \right] \\ &= \frac{\partial \lambda}{\partial \mathbb{E}_s i^e} \Pr(s) \left(\sum_{k=1}^{N} \frac{\Pr(k|s)}{\mathcal{P}_k} \frac{\partial \mathcal{P}_k}{\partial \lambda} \right) - \frac{\mathbb{E}_s i^e}{\lambda^2} \end{aligned} \tag{42}$$

[40] With c_{t+1} and Π_{t+1} at steady state and γ at its posterior mean in the estimation then the term in brackets is positive for all $i_t^e > -0.922$.

Substituting this back into equation (41) gives:

$$\frac{\partial \mathcal{I}}{\partial \mathbb{E}_s i^e} = \frac{1}{\lambda} - \frac{\partial \lambda}{\partial \mathbb{E}_s i^e} \sum_{s=1}^{S} \sum_{k=1}^{N} \frac{\Pr(s) \Pr(k|s)}{\mathcal{P}_k} \frac{\partial \mathcal{P}_k}{\partial \lambda} \tag{43}$$

Recall that $\mathcal{P}_k$ is defined as the unconditional probability of choosing bank k, so it can be written as $\sum_{s=1}^{S} \Pr(k|s) \Pr(s)$. Using this, equation (43) becomes:

$$\frac{\partial \mathcal{I}}{\partial \mathbb{E}_s i^e} = \frac{1}{\lambda} - \frac{\partial \lambda}{\partial \mathbb{E}_s i^e} \sum_{k=1}^{N} \frac{\partial \mathcal{P}_k}{\partial \lambda} \tag{44}$$

Since the sum of $\mathcal{P}_k$ over banks is always equal to 1, the sum of the derivatives of $\mathcal{P}_k$ must equal zero. We therefore have that:

$$\frac{\partial \mathcal{I}}{\partial \mathbb{E}_s i^e} = \frac{1}{\lambda} \tag{45}$$

Since λ is the Lagrange multiplier on the information constraint in the individual's problem, it is always strictly positive and $\mathcal{I}'(\mathbb{E}_s i^e) = \frac{\partial \mathcal{I}}{\partial \mathbb{E}_s i^e} > 0$.

Differentiating again with respect to $\mathbb{E}_s i^e$ we have:

$$\frac{\partial^2 \mathcal{I}}{\partial (\mathbb{E}_s i^e)^2} = -\frac{1}{\lambda^2} \frac{\partial \lambda}{\partial \mathbb{E}_s i^e} \tag{46}$$

$\mathcal{I}''(\mathbb{E}_s i^e)$ is therefore positive if $\frac{\partial \lambda}{\partial \mathbb{E}_s i^e} < 0$. I now proceed to prove that this is the case when individuals have uninformative priors and i^e is independent of the realized state s, as in all exercises in the paper except for Appendix C.1.

Differentiating the definition of i^e (12) with respect to i^e we have:

$$\frac{d\lambda}{di^e} = \frac{\lambda^2 \big(\sum_n \exp(\frac{i^n}{\lambda})\big)^2}{\big(\sum_n i^n \exp(\frac{i^n}{\lambda})\big)^2 - \big(\sum_n i^{n2} \exp(\frac{i^n}{\lambda})\big)\big(\sum_m \exp(\frac{i^m}{\lambda})\big)} \tag{47}$$

The numerator is always positive, so $\frac{d\lambda}{di^e}$ has the same sign as the denominator. After expanding the terms in brackets the denominator is:

$$\begin{aligned} \sum_n i^{n^2} \exp(\frac{2i^n}{\lambda}) + \sum_{m \neq n} i^n i^m \exp(\frac{i^n + i^m}{\lambda}) - \sum_n i^{n^2} \exp(\frac{2i^n}{\lambda}) - \sum_{m \neq n} i^{n^2} \exp(\frac{i^n + i^m}{\lambda}) \\ = -\sum_{m \neq n} (i^{n^2} - i^n i^m) \exp(\frac{i^n + i^m}{\lambda}) \end{aligned} \tag{48}$$

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:

$$\exp(\frac{i^j + i^k}{\lambda})(i^{j^2} - i^j i^k + i^{k^2} - i^k i^j) = \exp(\frac{i^j + i^k}{\lambda})(i^j - i^k)^2 > 0 \tag{49}$$

Each pair of terms inside the sum in equation (48) is therefore positive, and so $\frac{d\lambda}{di^e}$ is negative. That therefore implies that $\mathcal{I}''(i^e) = \frac{\partial^2 \mathcal{I}}{\partial (i^e)^2} > 0$.

B.3 The shadow value of information rises with rate dispersion

Here I show that if rate dispersion rises (specifically if all rates are subject to a mean-preserving spread) then λ falls, in the case of uninformative priors. With these uninformative priors we can write the probability of choosing bank n in state s as:

$$\Pr(n|s) = \frac{1}{1 + \sum_{j \neq n}^{N} \exp(\frac{i^j - i^n}{\lambda})} \tag{50}$$

Now consider a mean-preserving spread of interest rates, so replace each i^n with $\tilde{i}^n = k i^n - \bar{i}(k-1)$, where $\bar{i}$ is the unconditional mean of the pre-spread interest rates.

If choice probabilities are unchanged, and so attention $\mathcal{I}$ is unchanged, then it must be that for all n:

$$\sum_{j \neq n}^{N} \exp(\frac{i^j - i^n}{\lambda}) = \sum_{j \neq n}^{N} \exp(\frac{\tilde{i}^j - \tilde{i}^n}{\tilde{\lambda}}) = \sum_{j \neq n}^{N} \exp(\frac{k(i^j - i^n)}{\tilde{\lambda}}) \tag{51}$$

This is satisfied when $\tilde{\lambda} = k\lambda$. If $k > 1$ the mean-preserving spread increases the dispersion of interest rates, and correspondingly λ rises. Since $\mathcal{I}'(i^e) = \lambda^{-1}$, this reduces $\mathcal{I}'(i^e)$.

B.4 Equilibrium interest rates rise when attention increases

Here I show that as long as attention is not too high, when attention rises the interest rate distribution shifts up, just as it does in models based on Burdett and Judd (1983).

First, partially differentiate the first order condition for bank n (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:

$$-d_t^n \frac{\partial \mathcal{S}_t^n}{\partial \lambda_t} - (1 - \mathcal{S}_t^n) \frac{\partial i_t^n}{\partial \lambda_t} = 1 \tag{52}$$

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

$$\frac{\partial \mathcal{S}_t^n}{\partial \lambda_t} = \frac{\mathcal{S}_t^n(1-\mathcal{S}_t^n)}{\lambda_t}\frac{\partial i_t^n}{\partial \lambda_t} - \frac{\mathcal{S}_t^n(1-\mathcal{S}_t^n)i_t^n}{\lambda_t^2} + \mathcal{S}_t^n\left(\sum_{j\neq n}\frac{\mathcal{S}_t^j}{\lambda_t^2}(i_t^j - \lambda_t\frac{\partial i_t^j}{\partial \lambda_t})\right) \tag{53}$$

Substituting this in to equation (52) and rearranging we obtain:

$$\frac{\partial i_t^n}{\partial \lambda_t} = \frac{1}{\lambda_t(1-\mathcal{S}_t^n)(\lambda_t + d_t^n\mathcal{S}_t^n)}\left[i_t^n d_t^n \mathcal{S}_t^n(1-\mathcal{S}_t^n) - \lambda_t^2 - d_t^n\mathcal{S}_t^n\big(\sum_{j\neq n}\mathcal{S}_t^j(i_t^j - \lambda_t\frac{\partial i_t^j}{\partial \lambda_t})\big)\right] \tag{54}$$

From equation (15) we have $d_t^n = \lambda_t(1-\mathcal{S}_t^n)^{-1}$. Separately, we can write $\sum_{j\neq n}\mathcal{S}_t^j i_t^j = i_t^e - \mathcal{S}_t^n i_t^n$. Using these we obtain:

$$\frac{\partial i_t^n}{\partial \lambda_t} = \frac{\mathcal{S}_t^n}{\lambda_t(1-\mathcal{S}_t^n)}(i_t^n - i_t^e + \lambda_t\sum_{j\neq n}\mathcal{S}_t^j\frac{\partial i_t^j}{\partial \lambda_t}) - 1 \tag{55}$$

We now proceed with a guess-and-verify approach. Suppose that $\frac{\partial i_t^n}{\partial \lambda_t} < 0$ for all banks n, so every bank increases their interest rate when attention rises (λ falls). In that case we have that:

$$\frac{\partial i_t^n}{\partial \lambda_t} < \frac{\mathcal{S}_t^n}{\lambda_t(1-\mathcal{S}_t^n)}(i_t^n - i_t^e) - 1 \tag{56}$$

This means that a sufficient condition for $\frac{\partial i_t^n}{\partial \lambda_t} < 0$ is:

$$i_t^n < i_t^e + \frac{\lambda_t(1-\mathcal{S}_t^n)}{\mathcal{S}_t^n} \tag{57}$$

This is clearly true for all banks whose interest rate is below the effective interest rate. While it will not always be true for higher interest-rate banks, we can show that it will hold for sufficiently low attention (high λ_t).[41]

Recall that in this case with uninformative priors the effective interest rate rises monotonically with attention and falls monotonically with λ (see Appendix B.2), so $i_t^e \geq \bar{i}_t$, where $\bar{i}_t$ is the unweighted mean interest rate on offer in period t. Condition 57 is therefore satisfied if:

$$i_t^n < \bar{i}_t + \frac{\lambda_t(1-\mathcal{S}_t^n)}{\mathcal{S}_t^n} \tag{58}$$

Substituting out for i_t^n and $\bar{i}_t$ using the bank first order conditions, this becomes:

$$\lambda_t\left(\frac{1-\mathcal{S}_t^n + (\mathcal{S}_t^n)^2}{\mathcal{S}_t^n(1-\mathcal{S}_t^n)} - \frac{1}{N}\sum_{j=1}^{N}\frac{1}{1-\mathcal{S}_t^j}\right) > \bar{\chi}_t - \chi_t^n \tag{59}$$

[41] See Appendix B.5 for the proof that $\frac{\partial \lambda}{\partial \mathcal{I}} < 0$.

Consider the two fractions inside the brackets. The first is minimized at $\mathcal{S}_t^n = \frac{1}{2}$, at which point it equals 3:

$$\min_{\mathcal{S}_t^n} \frac{1 - \mathcal{S}_t^n + (\mathcal{S}_t^n)^2}{\mathcal{S}_t^n(1 - \mathcal{S}_t^n)} = 3 \tag{60}$$

The second is minimized when $\mathcal{S}_t^j = N^{-1}$ for all j, at which point:

$$\min_{\mathcal{S}_t^j} \left(-\frac{1}{N} \sum_{j=1}^{N} \frac{1}{1 - \mathcal{S}_t^j} \right) = -\frac{N}{N-1} \tag{61}$$

We therefore have:

$$\lambda_t \left(\frac{1 - \mathcal{S}_t^n + (\mathcal{S}_t^n)^2}{\mathcal{S}_t^n(1 - \mathcal{S}_t^n)} - \frac{1}{N} \sum_{j=1}^{N} \frac{1}{1 - \mathcal{S}_t^j} \right) > \lambda_t \left(\frac{2N-3}{N-1} \right) > \bar{\chi}_t - \chi_t^n \tag{62}$$

A sufficient condition for all banks to increase interest rates when λ_t falls is therefore:

$$\lambda_t > \frac{N-1}{2N-3} (\bar{\chi}_t - \chi_t^{min}) \tag{63}$$

Where χ_t^{min} is the lowest cost experienced by any bank in period t.

This condition (63) is sufficient rather than necessary, and may in fact be substantially more restrictive than necessary. In particular, it ignores the fact that interest rates are strategic complements ($\frac{\partial i_t^j}{\partial \lambda_t}$ enters equation (55) with a positive coefficient), so low-rate banks increasing their interest rates when λ falls will incentivize higher-rate banks to do the same. We can see this difference when $N = 2$, in which case the system of equations given by 55 has a straightforward analytic solution:

$$\frac{\partial i_t^n}{\partial \lambda_t} = \frac{1}{1 - \mathcal{S}_t^n + (\mathcal{S}_t^n)^2} \left[\frac{(\mathcal{S}_t^n)^2}{\lambda_t} (i_t^n - i_t^{-n} - \lambda_t) - 1 \right] \tag{64}$$

This is negative as long as (substituting out for i_t^n and i_t^{-n} using the bank first order condition):

$$\lambda_t > \frac{2(\mathcal{S}_t^n)^2(1 - \mathcal{S}_t^n)}{1 - 2\mathcal{S}_t^n + 3(\mathcal{S}_t^n)^2 - (\mathcal{S}_t^n)^3} (\bar{\chi}_t - \chi_t^{min}) \tag{65}$$

This is substantially less restrictive than condition (63). The right hand side of condition (65) is maximized at $\mathcal{S}_t^n = 0.603$, at which point the condition becomes $\lambda_t > 0.434(\bar{\chi}_t - \chi_t^{min})$, while condition (63) in the two-bank case is $\lambda_t > (\bar{\chi}_t - \chi_t^{min})$. In the estimated quantitative model condition (65) is easily satisfied at steady state, so all interest rates rise with attention in the region of the steady state.

In this case with two banks, we can also show that interest rate dispersion always falls when attention rises (λ_t falls). Using equation (64), we have that $\frac{\partial i_t^1}{\partial \lambda_t} > \frac{\partial i_t^2}{\partial \lambda_t}$ if:

$$i_t^1 - i_t^2 > \frac{\lambda_t(2\mathcal{S}_t^1 - 1)}{2(\mathcal{S}_t^1)^2 - 2\mathcal{S}_t^1 + 1} \tag{66}$$

Substituting out for i_t^1 and i_t^2 using the two bank first order conditions we obtain:

$$\chi_t^2 - \chi_t^1 > \lambda_t \left[\frac{2\mathcal{S}_t^1 - 1}{2(\mathcal{S}_t^1)^2 - 2\mathcal{S}_t^1 + 1} - \frac{1}{\mathcal{S}_t^1(1 - \mathcal{S}_t^1)} \right] = -\lambda_t \frac{(2(\mathcal{S}_t^1)^3 - (\mathcal{S}_t^1)^2 - \mathcal{S}_t^1 + 1)}{\mathcal{S}_t^1(1 - \mathcal{S}_t^1)(2(\mathcal{S}_t^1)^2 - 2\mathcal{S}_t^1 + 1)} \tag{67}$$

The fraction on the right hand side is positive for all $\mathcal{S}^1 \in (0, 1)$. We therefore have that in response to an attention rise, bank 1 raises interest rates by less than bank 2 ($\frac{\partial i_t^1}{\partial \lambda_t} > \frac{\partial i_t^2}{\partial \lambda_t}$) whenever bank 2 has higher costs - so whenever bank 2 offers lower rates. That gives us that dispersion falls when attention rises.

In general, search models based on Burdett and Judd (1983) have price dispersion initially rising in search effort, and then falling with search effort once effort is above some threshold. The reason for the difference with the inattention model is that Burdett-Judd models feature a reservation price, above which consumers do not buy. If we impose that interest rates cannot fall below some lower bound, then as attention approaches zero interest rates again converge on this lower bound, just as prices converge on the reservation price in Burdett and Judd (1983). In that case interest rate dispersion initially rises with attention as banks move away from the lower bound, then falls as found above, just as in Burdett-Judd models. Since there are two banks and no interest rate lower bound in the quantitative model in Section 5, this model behaves in a qualitatively similar way to a Burdett-Judd model in the region where more search effort reduces price dispersion. Numerically, when $N > 2$ interest rate dispersion is hump-shaped in attention even without an interest rate lower bound.

B.5 Relationship between attention and φ

In this section I demonstrate the close links between attention and φ, the empirical statistic constructed in Section 4. In the case of uninformative priors used in all model exercises except those in Appendix C.1, I first show that there is an exact correspondence between attention and φ in the model with $N = 2$ banks, as used in Section 5. I then show that with $N > 2$, if we hold the distribution of offered interest rates fixed then a rise in attention always implies a rise in φ. Numerically, this upward-sloping relationship also holds with $N > 2$ if we allow for the interest rate distribution to adjust, as long as attention is not too high. The same is true in the case studied in Appendix C.1, with

$N = 2$ and priors biased towards one bank over the other.[42]

B.5.1 N=2 banks, uninformative priors

As in Section 5, define p_t^g as the probability an individual chooses the high interest rate bank in period t:

$$p_t^g = \frac{\exp(\frac{i_t^g}{\lambda_t})}{\exp(\frac{i_t^g}{\lambda_t}) + \exp(\frac{i_t^b}{\lambda_t})} \tag{68}$$

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:

$$i_t^b = \mathcal{P}_1 i_t^1 + (1 - \mathcal{P}_1) i_t^2 = 0.5(i_t^1 + i_t^2) \tag{69}$$

With two banks and uninformative priors, the attention constraint 9 becomes:

$$\mathcal{I}_t = \log(2) + p_t^g \log p_t^g + (1 - p_t^g) \log(1 - p_t^g) \tag{70}$$

Attention is therefore a monotonically increasing function of p_t^g (as $p_t^g \geq 0.5$).

The empirical statistic φ is:

$$\varphi_t = \frac{p_t^g i_t^g + (1 - p_t^g) i_t^b - \frac{1}{2}(i_t^g + i_t^b)}{\frac{1}{2}(i_t^g - i_t^b)} \tag{71}$$

This simplifies to:

$$\varphi_t = \frac{p_t^g (i_t^g - i_t^b) - \frac{1}{2}(i_t^g - i_t^b)}{\frac{1}{2}(i_t^g - i_t^b)} = 2p_t^g - 1 \tag{72}$$

In this case φ_t is therefore a linear function of the probability an individual successfully chooses the higher interest rate bank, which itself is an increasing concave function of attention. This case also highlights the importance of normalizing the spread $i_t^e - i_t^b$ by the standard deviation of interest rates to obtain φ_t: without that, φ_t would be increasing in $i_t^g - i_t^b$, even if p_t^g and so attention are held constant. The normalization therefore prevents changes in rate dispersion from mechanically affecting φ_t.

The normalization only exactly removes all dependence on the shape of the rate distribution in this case of $N = 2$ and uninformative priors, but still helps mitigate the dependence of $i_t^e - i_t^b$ on the spread of interest rates more generally. In particular, it ensures that φ_t is homogeneous of degree 0 in interest rates, so a mean-preserving spread

[42] In this case φ will depend on which state of the world is realized, since it is based on the realized effective and benchmark rates, and realized rate dispersion. Attention, in contrast, is determined before the state realization is known. Simulations of the model in Appendix C.1 show that higher attention increases φ in both possible states of the world, for all possible levels of attention.

of the interest rate distribution (as studied in Appendix B.3) leaves φ_t unchanged unless attention, and so choice probabilities, change.

B.5.2 N>2 banks

Since all variables here are defined within the same period I drop all time subscripts to simplify notation. 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:

$$\varphi = \frac{\sum_n i^n \Pr(\text{choose } n) - \bar{i}}{\sigma(i)} = \frac{\frac{\sum_n i^n \exp(\frac{i^n}{\lambda})}{\sum_m \exp(\frac{i^m}{\lambda})} - \bar{i}}{\sigma(i)} \tag{73}$$

First, note that as $\mathcal{I}$ approaches 0, λ tends to infinity,[43] and so $\varphi = 0$ when attention is 0:

$$\lim_{\lambda \to \infty} \varphi = \frac{\frac{1}{N}\sum_n i^n - \bar{i}}{\sigma(i)} = 0 \tag{74}$$

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 $\varphi > 0$:

$$\varphi(\mathcal{I} = \log(N)) = \frac{i^1 - \bar{i}}{\sigma(i)} = \frac{\frac{1}{N}\sum_n (i^1 - i^n)}{\sigma(i)} > 0 \tag{75}$$

Since φ is continuous in attention for $\mathcal{I} \in (0, \log(N))$, the statements above guarantee that $\mathcal{I}$ and φ are positively related at least in some portions of this range.

To make further progress, I now consider how φ changes in the model assuming that interest rates are held fixed. We use the chain rule to write:

$$\frac{\partial \varphi}{\partial \mathcal{I}} = \frac{\partial \varphi}{\partial \lambda}\frac{\partial \lambda}{\partial \mathcal{I}} \tag{76}$$

I start with $\frac{\partial \lambda}{\partial \mathcal{I}}$. Again using the chain rule we have:

$$\frac{\partial \lambda}{\partial \mathcal{I}} = \frac{\partial \lambda}{\partial i^e}\frac{\partial i^e}{\partial \mathcal{I}} < 0 \tag{77}$$

To see why this is negative, recall that in Appendix B.2 I showed that $\frac{\partial \lambda}{\partial i^e} < 0$ and $\frac{\partial \mathcal{I}}{\partial i^e} = \lambda^{-1} > 0$. If attention rises, then holding the distribution of interest rates constant the shadow price of attention falls.

[43] To see this, recall that when $\mathcal{I} = 0$ all choice probabilities must equal N^{-1}. From equation (13), $\Pr(n|s) = \frac{\exp(i^n/\lambda)}{\sum_{j=1}^{N}\exp(i^j/\lambda)}$. If there is any interest rate dispersion then this only approaches N^{-1} for all banks if $\lambda \to \infty$.

Now consider $\frac{\partial \varphi}{\partial \lambda}$. Since $\frac{\partial \lambda}{\partial i^e} < 0$ we have:

$$\frac{\partial \varphi}{\partial \lambda} = \frac{1}{\sigma(i)} \frac{\partial i^e}{\partial \lambda} < 0 \tag{78}$$

Together, equations (77) and (78) imply that $\frac{\partial \varphi}{\partial \mathcal{I}} > 0$. Holding the distribution of interest rates constant, φ monotonically increases with attention.

This, however, is only the direct effect of a change in attention on φ. As shown in Appendix B.4, a change in attention also implies a change in the interest rate distribution, which when $N > 2$ will have an indirect effect on φ. Numerically, these indirect effects are small, such that attention and φ are positively related as long as attention is not extremely high.

If attention is very high, then φ can fall as attention increases, because an increase in attention causes the highest rate bank to lower their rates, or only raise them a small amount (see Appendix B.3). Since attention is very high, individuals choose this bank with a very high probability, and so their effective interest rate only increases a small amount with attention. The increase in attention does, however, cause lower-rate banks to increase their interest rates, and so the benchmark rate increases more strongly than i^e. With $N = 2$ this is counteracted in φ by the normalization by $\sigma(i)$, but with a larger number of banks this adjustment is incomplete because the $N - 1$ lowest rate banks do not converge on each other at the same rate as they converge on the best bank. This breakdown of the link between φ and $\mathcal{I}$, however, only occurs at extreme levels of attention outside of plausible parameter ranges.[44]

C Persistent bank costs

C.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.

Suppose that, as in Section 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 5, assume that there is persistence in the state. Specifically, the state of the world , denoted

[44] If χ^n are spaced equally on $[0, \chi_1]$, where χ_1 is the highest bank cost in the steady state of the quantitative model, and i^{CB} is at the steady state value from that model, then with $N = 3$ the peak of φ occurs when attention is such that individuals choose the highest rate bank with probability 0.87. As N rises the $\Pr(1|1)$ associated with the threshold level of attention does fall, but only gradually. With $N = 25$, φ is increasing in $\mathcal{I}$ as long as $\Pr(1|1) < 0.85$.

s_t, follows a two-state Markov process, in which $\Pr(s_{t+1} = s | s_t = s) = g$, where $g \geq 0.5$.

C.1.1 Savers

Assume that savers know the previous state of the world: they observe whether they chose correctly or not when the interest rate payouts occur.[45] Their choice problem in period t therefore remains a static problem. The persistence in s_t shows up as a prior belief biased towards the previous period's realised state, which I assume without loss of generality to be state 1. 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, where again I have dropped time subscripts since the saver problem is static (the same will also be true of the bank problem):

Table 2: Payoff matrix, 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 if state s is realized. 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 (10):

$$P(n|i^{n,s}, i^{-n,s}, s) = \frac{\mathcal{P}_n \exp(\frac{i^{n,s}}{\lambda})}{\mathcal{P}_n \exp(\frac{i^{n,s}}{\lambda}) + (1-\mathcal{P}_n)\exp(\frac{i^{-n,s}}{\lambda})} \tag{79}$$

The unconditional choice probabilities (predispositions) are found as the solution to two normalization conditions (following Matějka and McKay, 2015):

$$\frac{\exp(\frac{i^{1,1}}{\lambda})g}{\mathcal{P}_1 \exp(\frac{i^{1,1}}{\lambda}) + (1-\mathcal{P}_1)\exp(\frac{i^{2,1}}{\lambda})} + \frac{\exp(\frac{i^{1,2}}{\lambda})(1-g)}{\mathcal{P}_1 \exp(\frac{i^{1,2}}{\lambda}) + (1-\mathcal{P}_1)\exp(\frac{i^{2,2}}{\lambda})} = 1 \tag{80}$$

$$\frac{\exp(\frac{i^{2,1}}{\lambda})g}{\mathcal{P}_1 \exp(\frac{i^{1,1}}{\lambda}) + (1-\mathcal{P}_1)\exp(\frac{i^{2,1}}{\lambda})} + \frac{\exp(\frac{i^{2,2}}{\lambda})(1-g)}{\mathcal{P}_1 \exp(\frac{i^{1,2}}{\lambda}) + (1-\mathcal{P}_1)\exp(\frac{i^{2,2}}{\lambda})} = 1 \tag{81}$$

[45] 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 Steiner et al. (2017).

The $\mathcal{P}_1$ that satisfies these conditions is:

$$\mathcal{P}_1 = \frac{e^{\frac{i^{21}}{\lambda}} e^{\frac{i^{22}}{\lambda}} - (1-g) e^{\frac{i^{21}}{\lambda}} e^{\frac{i^{12}}{\lambda}} - g e^{\frac{i^{11}}{\lambda}} e^{\frac{i^{22}}{\lambda}}}{e^{\frac{i^{11}}{\lambda}} e^{\frac{i^{12}}{\lambda}} - e^{\frac{i^{21}}{\lambda}} e^{\frac{i^{12}}{\lambda}} - e^{\frac{i^{11}}{\lambda}} e^{\frac{i^{22}}{\lambda}} + e^{\frac{i^{21}}{\lambda}} e^{\frac{i^{22}}{\lambda}}} \tag{82}$$

C.1.2 Banks

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 2.1 (again dropping time subscripts):

$$\frac{d}{di^n} P(n|s) \cdot (i^{CB} - i^n - \chi^n) = P(n|s) \tag{83}$$

I assume that banks take saver predispositions as given when deciding their interest rates. Intuitively, predispositions reflect household knowledge of the exogenous law of motion for the state of the world, and of bank policy functions. If households learn about how banks respond to different costs over time, then a bank changing its policy will not have any effect on predispositions until households learn about the change over many periods. The assumption can therefore be seen as assuming that banks are myopic, and don't take into account the future benefits of manipulating predispositions.[46] 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 2:

$$\big(1 - P(n|s)\big) \cdot (i^{CB} - i^n - \chi^n) = \lambda \tag{84}$$

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.

C.1.3 Equilibrium

To find equilibrium, take equation (84) and equation (79) for each of the four combinations of bank and state, and equation (82) to give 9 equations in 9 variables: the four interest rates, four conditional choice probabilities, and the predisposition towards bank 1. Since this allows $\mathcal{P}_1$ to vary in response to interest rates, this equilibrium can be taken as the

[46]This is similar to the assumptions in the deep habits model of Ravn et al. (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. This assumption avoids counterintuitive equilibria in which a fall in attention implies fierce competition for predispositions as households lean more heavily on these in their decisions.

steady state of the system after predispositions have had time to adjust. I solve this system numerically for an example calibration, and study how the resulting equilibrium varies with λ and g. The qualitative results are robust to a wide variety of calibrations.

All of the results from the static cost model still hold: as attention rises interest rate dispersion falls and average rates rise. The highest rate in the market rises as λ falls as long as λ is above some threshold level. Graphs showing this with some example parameters are in Figure 4 below.

On top of those results, we have two new results. 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.

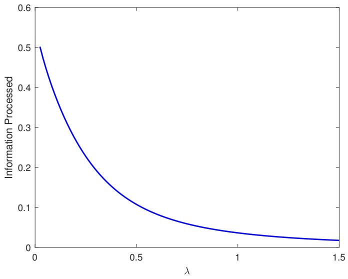
The second result is that the effective interest rate faced by the large household depends on the state of the world. Bank 1 is more likely to be the low cost bank, so savers are predisposed to choose them. Bank 1 responds to this predisposition by offering lower interest rates. This only partially offsets the prior belief effect, so savers have $\mathcal{P}_1 > 0.5$ in equilibrium. This means that if the state stays at s_1 (bank 1 is low cost), savers are more likely to correctly identify the low cost bank than they are if the state changes to s_2. This increases the effective interest rate in s_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 s_2, which increases the effective interest rate in s_2 relative to s_1. Which effect dominates depends on the calibration, but in either case there are two possible effective interest rates each period, and whenever there is a transition from one state to the other the effective interest rate will change even if all other variables are at steady state. 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 shocks to the household effective interest rate, with $i^{e,1}$ realized with probability g and $i^{e,2}$ realized with probability $1-g$. These shocks are the key qualitative difference between this model and the static cost model in Sections 2 and 5.

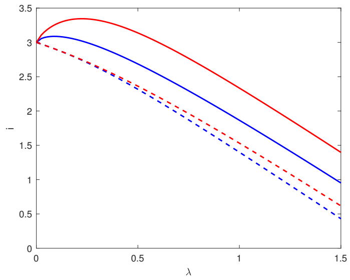
C.2 Persistence of interest rate rankings in the data

In Sections 2 and 5 I assume that the ranking of a bank in the interest rate distribution has no persistence. Table 3 shows the bank transition probabilities between quintiles of the interest rate distribution of the products studied in Section 4 over a month and a year. The length of a period in Section 5 is one month, but the annual transition probabilities

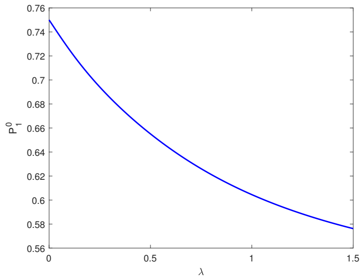
Figure 4: Long run equilibrium varies with λ



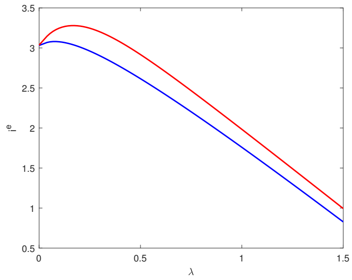
(a) Amount of information processed



(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).

are also relevant since these products have a term of one year, so individual savers buying these products have to revisit their decision a year later (or exit the market).

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 37% probability of that bank still being in the top quintile. This explains why adding bank fixed effects do not account for much of the dispersion of interest rates, as discussed in Section 4.1.1.

Table 3: 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.

	1	2	3	4	5
1	0.59	0.16	0.10	0.07	0.07
2	0.19	0.51	0.19	0.06	0.04
3	0.03	0.28	0.43	0.20	0.07
4	0.01	0.08	0.30	0.41	0.20
5	0.01	0.03	0.08	0.23	0.65

(a) Monthly

	1	2	3	4	5
1	0.36	0.23	0.15	0.13	0.13
2	0.25	0.31	0.22	0.13	0.09
3	0.15	0.25	0.25	0.21	0.14
4	0.09	0.19	0.21	0.28	0.23
5	0.06	0.15	0.19	0.25	0.36

(b) Annual

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:

$$-2\ln\left(\frac{\prod_{n=1}^{5}\prod_{m=1}^{5}p_{n,m}}{\prod_{n=1}^{5}\prod_{m=1}^{5}0.2}\right)\sim\chi^2_{19} \tag{85}$$

The critical value of the test statistic for 5% significance is 30.1. The monthly and annual transition matrices give test statistics of 24.6 and 4.3 respectively. We therefore cannot reject the hypothesis of no persistence at either an annual or a monthly frequency.

D Further results and robustness for Section 4

D.1 Relationship between bank positions in different market segments

To calculate the Quoted Household Interest Rate used to construct φ in Section 4, the Bank of England computes a weighted average of the interest rates in the set of products detailed in Section 3.1. The weights are the quantities of new deposits per bank across a broader set of products than those from which the interest rates are taken. Here I show that a bank's position in the distribution of interest rates qualifying for inclusion in the Quoted Household Interest Rate is extremely closely related to their position in the other market segments included when the weights are calculated. As argued in Section 3.1, this implies that the cyclical patterns in φ found in Section 4.3 reflect a systematic shift towards banks at the top of all of these market segments when unemployment is high and interest rates are low. φ is therefore informative about the position of household choices within the distribution of available rates despite this data limitation.

The weights for the Quoted Household Interest Rate are constructed using new deposits in all fixed interest rate bonds with terms up to one year. This is broader than the set of products from which the interest rates are taken: 30% of the products in the broader set qualify for inclusion in the Quoted Household Interest Rate. Taking all products in the broader set from the Moneyfacts data, I divide them into market segments based on their term, investment size, and interest payment frequencies. The set of such characteristics is given in Table 4.

Table 4: Bank product characteristics used for subdividing the fixed rate bond market

Characteristic	**Division**
Term length (months)	{1-3, 4-6, 7-9, 10-12}
Investment size (£000s)	$\{1, 2.5, 5, 10, 25, 50\}$
Interest payment frequency	{Monthly, Quarterly, On maturity}

Taking all combinations of these characteristics yields 72 market segments. Many products are included in multiple segments because an investment of £10000, for example, is often eligible for products with lower minimum investments.

For each segment each month, I rank the banks that compete in that segment-month by their interest rate in that segment-month. I do the same for the set of products included in the Quoted Household Interest Rate (the Q segment). I then compute the correlation between these ranks each month, then finally for each market segment I take the mean of these rank correlations over the months, weighting by the number of banks competing in both that segment and the Q segment that month (i.e. weighting by the number of observations used to construct that month's correlation). This gives an average interest rate rank correlation between the Q segment and every other market segment used in constructing the Quoted Household Interest Rate weights.

For 30 of the market segments, there are either no products with that set of characteristics, or there are no occasions where more than one bank simultaneously competes in that segment and the Q segment. This leaves 42 segments for which the rank correlation with the Q segment can be computed.

In these remaining market segments, bank rankings are extremely highly correlated with the rankings in the Q segment. The mean rank correlation across the segments is 0.70, and this is distorted by a small number of market segments which very few banks ever compete in. Of the six segments with rank correlations with the Q segment below 0.5, four are 7-9 month bonds with a monthly payment frequency, which contain less than 1 product per month on average. The other two are also very small segments, with an average of 1.02 and 1.18 banks competing simultaneously in them and the Q segment each month. These correlations are therefore based off very few observations, and the small number of banks competing there each month suggests that they are not large market segments, making them unlikely to play a big role in the weights used to calculate the Quoted Household Interest Rate.

Other market segments are much larger. In the ten largest market segments, the average number of banks competing in those segments and the Q segment each month is greater than 11. For the largest five segments, it exceeds 25.

The mean rank correlation across the segments rises to 0.84 when segments are weighted by this mean number of banks competing there and in the Q segment each month. If we take the number of banks competing in a segment as indicative of the size of that market segment, this shows that bank positions within the interest rate distribution analysed in Section 4 (in the Q segment), are very highly correlated with bank positions in the other substantial market segments that are included in the weights behind the Quoted Household Interest Rate data.

D.2 Time series behaviour of φ

In Section 4.3 I showed evidence that φ is countercyclical. The time series of φ is plotted in Figure 5. Consistent with the finding that φ is countercyclical, the largest falls in

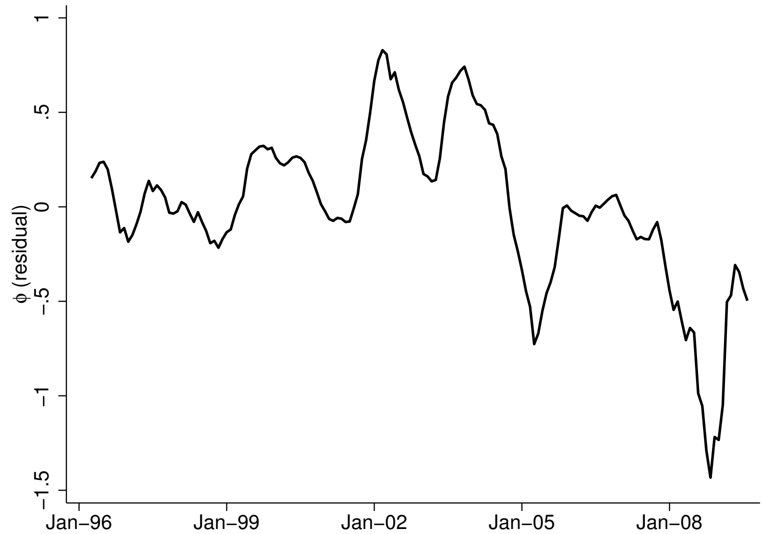


Figure 5: Time series of the residual of φ after regressing on pos_t^b (the position of the average big four interest rate within the interest rate distribution), 6 month moving averages.

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.

Given its otherwise strongly countercyclical nature, it is notable that φ does not start rising earlier in the Great Recession. This is possibly because late 2008 was a tumultuous period in the UK retail banking market. There were several large mergers and bailouts as the financial crisis hit the market. At this time the big four banks increased their interest rates relative to the rest of the market. The initial lack of increase in φ could therefore be explained by heightened awareness of bank risk causing savers to stay away from the larger banks who were more exposed to international financial markets.[47]

D.3 Alternative measures of φ

Here I present two alternatives to the household choice statistic φ, which corroborate the evidence in Section 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.

$$\varphi_{best} = \frac{\mathbb{E}_h i_t - i_t^{best}}{\sigma(i_t)} \tag{86}$$

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.

$$\varphi_{pct} = \Pr(i_t^n < \mathbb{E}_h i_t) \tag{87}$$

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 5 below. As in Section 4, all correlations are between the cyclical components of each variable, extracted with a HP filter.

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 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.

[47] This large movement of the big four within the rate distribution is rare. The variance of the percentile of the benchmark interest rate within the interest rate distribution is 4x smaller than that of the Quoted Household Interest Rate.

Table 5: Pairwise contemporaneous correlations of attention proxies, the unemployment rate, and within-month mean interest rates.

	φ	φ_{best}	φ_{pct}	U	$\bar{i}$
φ	1				
φ_{best}	-0.627^{***}	1			
φ_{pct}	0.713^{***}	-0.556^{***}	1		
U	0.273^{***}	-0.549^{***}	0.342^{***}	1	
$\bar{i}$	-0.277^{***}	0.454^{***}	-0.365^{***}	-0.795^{***}	1

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

D.4 Market composition cannot explain fluctuations in φ_t

In this appendix I show that the composition of households holding fixed term savings bonds does not vary significantly through the Great Recession, suggesting that such compositional changes are unlikely to explain the cyclical variation in φ_t found in Section 4.

Drechsler et al. (2017) 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. In principle, this kind of switching 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 compositional change 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 term savings bonds.[48] As the three waves span the Great Recession, if a composition effect is driving the cyclicality of φ_t we should find that characteristics associated with being more attentive to financial decisions become relatively more common over the recession, among the people who hold fixed-term bonds.

Bhutta et al. (2020) and Iscenko (2018) find that households are more likely to be attentive to mortgage decisions if they have high incomes and high levels of education. Iscenko (2018) also finds that age matters, but in a non-linear way. I therefore explore compositional changes among fixed-term bond-holders along these lines. Specifically, I consider household income by decile of the overall income distribution,[49] indicators for any

[48] Note that I cannot distinguish fixed rate from variable rate products with fixed terms in this data.

[49] Raw income would be inappropriate as the aggregate income distribution was changing over this time. Deciles of the income distribution are computed from the self-reported labour income plus self-employed

educational qualifications and for degree-level qualifications, and an indicator for whether the household is aged 45-54, the age identified by Finke et al. (2017) as corresponding to peak financial knowledge. Table 6 reports the results of regressing each of these on indicators for the wave in which the person was surveyed, using the subset of households who hold a fixed-term bond. All regressions use robust standard errors and are weighted using the survey weights provided in the WAS.

Table 6: Regressions on variables related to financial literacy. Wave 1 (2006) is the baseline. Waves 2 and 3 took place in 2008 and 2010.

	(1) Income decile	(2) Some qualification	(3) Degree qualification	(4) Aged 45-54
Wave=2	-0.0336	0.00734	0.0156	-0.0172
	(-0.34)	(0.59)	(1.01)	(-1.49)
Wave=3	-0.322**	-0.0197	0.00135	-0.0116
	(-3.29)	(-1.54)	(0.09)	(-0.98)
Constant	4.897***	0.833***	0.305***	0.140***
	(78.06)	(102.50)	(30.69)	(18.36)
Observations	6860	6856	6856	6860

t statistics in parentheses

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

The only composition change that is significantly different from zero is that the income of fixed term bond-holders declined slightly relative to the overall income distribution between waves 1 and 3. This is the opposite direction to the compositional change that would be required to explain the cyclical patterns of φ_t. All other compositional changes are not significantly different from zero. It is therefore unlikely that compositional changes explain the cyclicality of φ_t.

E Quantitative model equations

Table 7 lists the (endogenous and exogenous) variables of the quantitative model, and Table 8 lists the log-linearised model equations. For a complete derivation see Harrison and Oomen (2010). In the tables below, $\overline{X}$ denotes the steady state of the variable X, and e_{Xt} is an i.i.d.-normal innovation. Notice that along with the monetary policy and labour disutility shocks, the attention shock ζ^{μ} is assumed to be i.i.d. This is because the estimation finds the shock has a negligible effect on all of the observables, so cannot identify the shock's persistence. As the shock is so small, calibrating the persistence to any other value $[0, 1)$ makes no difference to the results.

income within each survey wave.

Table 7: Description of variables in the quantitative model

Variable	Description	Variable	Description
b^f	Net foreign assets	π^m	Inflation: imports
c	Consumption: total	π^w	Inflation: wage
c^h	Consumption: domestic goods	π^{xvf}	Inflation: producer price of exports
c^m	Consumption: imports	q	Real exchange rate
h	Hours	r	Rental rate on capital
inv	Investment	r^b	Gross bad bank interest rate
k	Capital	r^{CB}	Gross policy interest rate
λ	Shadow value of information	r^e	Gross effective interest rate
p^g	Probability of choosing the good bank	r^g	Gross good bank interest rate
p^h	Relative price of domestic final goods	u_c	Marginal utility of c
p^{hv}	Relative price of domestic intermediate goods	w	Real wage
p^m	Relative price of imported goods	x	Exports
p^x	Relative price of exported goods	y^h	Output: used domestically
p^{xv}	Relative producer price of exported final goods	y^v	Output: total
π	Inflation: total	z	Capital utilisation
π^{hv}	Inflation: domestic intermediates		
cf	Foreign demand	ζ^χ	Bank interest rate level shock
g	Government spending	$\zeta^{\chi b}$	Bank interest rate dispersion shock
πf	Foreign inflation	ζ^{hb}	Markup shock
pxf	Foreign relative export prices	ζ^k	Capital adjustment cost shock
rf	Foreign interest rate	$\zeta^{\kappa h}$	Labour disutility shock
tfp	TFP	ζ^{rg}	Monetary policy shock
ζ^c	Risk premium shock	ζ^μ	Attention shock

Table 8: Log-linearised quantitative model equations

Name	**Equation**
Wage inflation definition	$\pi_t^w = w_t - w_{t-1} + \pi_t$
Wage Phillips Curve	$(1+\beta\epsilon^w)\pi_t^w - \epsilon^w \pi_{t-1}^w$
	$= \beta \mathbb{E}_t \pi_{t+1}^w + \frac{\psi^w(1-\beta(1-\psi^w))}{1-\psi^w}\big(\frac{\sigma^h}{\sigma^h+\sigma^w}\big)\big(\frac{1}{\sigma^h}h_t - u_{ct} - w_t + \zeta_t^{\kappa h}\big)$
Marginal Utility of c	$u_{ct} = -\frac{1}{\sigma^c}c_t + \psi^{hab}\big(\frac{1}{\sigma^c} - 1\big)c_{t-1}$
Consumption Euler equation	$u_{ct} = \mathbb{E}_t(u_{ct+1} + r_t^e - \pi_{t+1}) + \zeta_t^c$
k first order condition	$p_t^h + \chi^k(k_t - k_{t-1} - \epsilon^k(k_{t-1} - k_{t-2})) + \zeta_t^k + r_t^e - \mathbb{E}_t\pi_{t+1}$
	$= \beta\mathbb{E}_t(\chi^k(k_{t+1} - k_t - \epsilon^k(k_t - k_{t-1})) + \chi^z r_{t+1} + (1-\delta)p_{t+1}^h + \zeta_{t+1}^k)$
z first order condition	$r_t = \sigma^z z_t + p_t^h$
Relative import demand	$c_t^m = -\sigma^m p_t^m + c_t$
Relative home good demand	$c_t^h = -\sigma^m p_t^h + c_t$
Consumption basket	$c_t = \frac{\overline{c}^h}{\overline{c}}(p_t^h + c_t^h) + \frac{\overline{c}^m}{\overline{c}}(p_t^m + c_t^m)$
Attention first order condition	$u_{ct} - \mathbb{E}_t\pi_{t+1} = -\lambda_t + \zeta_{\mu t}$
Optimal bank choice probability	$p_t^g = \frac{1-\overline{p}^g}{\overline{\lambda}}(\overline{r}^g r_t^g - \overline{r}^b r_t^b - (\overline{r}^g - \overline{r}^b)\lambda_t)$
Effective rate definition	$\frac{1}{\beta}r_t^e = \overline{p}^g(\overline{r}^g - \overline{r}^b)p_t^g + \overline{p}^g\overline{r}^g r_t^g + (1-\overline{p}^g)\overline{r}^b r_t^b$
Production function	$y_t^v = tfp_t + \frac{(1-\alpha)\overline{h}^{\frac{\sigma^y-1}{\sigma^y}}}{(1-\alpha)\overline{h}^{\frac{\sigma^y-1}{\sigma^y}} + \alpha\overline{k}^{\frac{\sigma^y-1}{\sigma^y}}}h_t + \frac{(1-\alpha)\overline{h}^{\frac{\sigma^y-1}{\sigma^y}}}{\alpha\overline{k}^{\frac{\sigma^y-1}{\sigma^y}} + \alpha\overline{k}^{\frac{\sigma^y-1}{\sigma^y}}}(k_{t-1} + z_t)$
Domestically consumed inflation definition	$\pi_t^{hv} = p_t^{hv} - p_{t-1}^{hv} + \pi_t$
Export inflation definition	$\pi_t^{xvf} = p_t^{xv} - p_{t-1}^{xv} + \pi f_t + q_t - q_{t-1}$
Domestic good Phillips Curve	$(1+\beta\epsilon^{hv})\pi_t^{hv} - \epsilon^{hv}\pi_{t-1}^{hv}$
	$= \beta\mathbb{E}_t\pi_{t+1}^{hv} - \frac{\sigma^{hb}-1}{\chi^{hv}}(p_t^{hv} + \frac{1}{\sigma^y}(y_t^v - h_t) - w_t + \frac{\sigma^y-1}{\sigma^y}tfp_t) + \zeta_t^{hb}$
Export good Phillips Curve	$(1+\beta\epsilon^{xv})\pi_t^{xvf} - \epsilon^{xv}\pi_{t-1}^{xvf}$
	$= \beta\mathbb{E}_t\pi_{t+1}^{xvf} - \frac{\sigma^{xb}-1}{\chi^{xv}}(p_t^{xv} + \frac{1}{\sigma^y}(y_t^v - h_t) - w_t + \frac{\sigma^y-1}{\sigma^y}tfp_t) + \zeta_t^{hb}$
Optimal k-h ratio	$z_t + k_{t-1} - h_t = \sigma^y(w_t - r_t)$
Good bank profit maximisation	$\lambda_t = \frac{1}{\overline{r}^{CB} - \overline{r}^g}(\overline{r}^{CB}r_t^{CB} - \overline{r}^g r_t^g - \zeta_{\chi t}) - \frac{\overline{p}^g}{1-\overline{p}^g}p_t^g$
Bad bank profit maximisation	$\lambda_t = \frac{1}{\overline{r}^{CB} - \overline{r}^b - \chi_1}(\overline{r}^{CB}(1-\chi_2)r_t^{CB} - \overline{r}^b r_t^b - \zeta_{\chi t} - \zeta_{\chi b t}) + p_t^g$
Taylor rule	$r_t^{CB} = \zeta_t^{rg} + \theta^{rg}r_{t-1}^{CB} + (1-\theta^{rg})(\theta^p\pi_t + \theta^y(y_t^v - tfp_t))$
Export demand	$x_t = cf_t - \sigma^x(q_t + p_t^x - pxf_t)$
Import inflation definition	$\pi_t^m = p_t^m - p_{t-1}^m + \pi_t$
Import Phillips Curve	$(1+\beta\epsilon^m)\pi_t^m - \epsilon^m\pi_{t-1}^m$
	$= \beta\mathbb{E}_t\pi_{t+1}^m + \frac{\psi^m(1-\beta(1-\psi^m))}{1-\psi^m}(pxf_t - q_t - p_t^m)$
Price of domestic consumption basket	$p_t^h = \kappa^{hv}p_t^{hv} + (1-\kappa^{hv})p_t^m$
Price of export consumption basket	$p_t^x = \kappa^{xv}p_t^{xv} + (1-\kappa^{xv})p_t^m$
k law of motion	$\delta inv_t = k_t - (1-\delta)k_{t-1} + \chi^z z_t$
Goods market clearing	$y_t^v = \kappa^{hv}(\overline{c}^h c_t^h + \overline{inv}inv_t + \overline{g}g_t) + \kappa^{xv}\overline{x}x_t$
Domestic goods market clearing	$y_t^h = \frac{\overline{c}^h}{\overline{y}^h}c_t^h + \frac{\overline{c}^h}{\overline{y}^h}inv_t + \frac{\overline{inv}}{\overline{g}}g_t$
b^f law of motion	$b_t^f = \beta^{-1}b_{t-1}^f + \beta^{-1}\overline{b}^f(rf_{t-1} - \pi f_t - q_t + q_{t-1}) + \overline{x}(p_t^x + x_t) - \overline{c}^m c_t^m$
	$-(1-\kappa^{hv})\overline{y}^h y_t^h - (1-\kappa^{xv})\overline{x}x_t - (\overline{c}^m + (1-\kappa^{hv})\overline{y}^h + (1-\kappa^{xv})\overline{x})p_t^m$
Real UIP	$\mathbb{E}_t q_{t+1} - q_t + \chi^{bf}b_t^f = rf_t - r_t^{CB} - \mathbb{E}_t(\pi f_{t+1} - \pi_{t+1})$
TFP	$tfp_t = \rho_{tfp}tfp_{t-1} + e_{tfpt}$
Government spending	$g_t = \rho_g g_{t-1} + e_{gt}$
Risk premium shock	$\zeta_t^c = \rho_{\zeta^c}\zeta_{t-1}^c + (1-\rho_{\zeta^c}^2)^{\frac{1}{2}}e_{\zeta^c t}$
Markup shock	$\zeta_t^{hb} = \rho_{\zeta^{hb}}\zeta_{t-1}^{hb} + (1-\rho_{\zeta^{hb}}^2)^{\frac{1}{2}}e_{\zeta^{hb}t}$
Capital adjustment cost shock	$\zeta_t^k = \rho_{\zeta^k}\zeta_{t-1}^k + (1-\rho_{\zeta^k}^2)^{\frac{1}{2}}e_{\zeta^k t}$
Labour disutility shock	$\zeta_t^{\kappa h} = e_{\zeta^{\kappa h}t}$
Monetary policy shock	$\zeta_t^{rg} = e_{\zeta^{rg}t}$
Bank rate level shock	$\zeta_t^{\chi} = \rho_{\zeta^{\chi}}\zeta_{t-1}^{\chi} + (1-\rho_{\zeta^{\chi}}^2)^{\frac{1}{2}}e_{\zeta^{\chi}t}$
Bank rate dispersion shock	$\zeta_t^{\chi b} = \rho_{\zeta^{\chi b}}\zeta_{t-1}^{\chi b} + (1-\rho_{\zeta^{\chi b}}^2)^{\frac{1}{2}}e_{\zeta^{\chi b}t}$
Attention shock	$\zeta_t^{\mu} = e_{\zeta^{\mu}t}$
Foreign variables	VAR(4) in Appendix F.2

F Estimation Details

F.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 the standard observables. See their paper for details of the data series. The exception to their method is that I use industrial production for all foreign countries, where they use a mix of industrial production and GDP.

I take log first differences of all domestic real variables, and transform inflation and interest rates into quarterly gross rates before taking logs and de-meaning. For the foreign real variables, I take logs and then extract the cyclical components using a one-sided HP filter. For the average and standard deviation of interest rates in Moneyfacts I follow the same procedure used for the treasury bill rate, averaging across months within each quarter before taking logs, and leaving a quarter as missing when a month of data is missing. I include φ in levels to avoid losing more observations after the quarters with missing months through first-differencing. I therefore use a one-sided HP filter to extract the cyclical component of φ.[50] I choose χ_1 to match the average gap between the highest and the (unweighted) mean interest rate available in the data.

Using $N = 2$ banks in the quantitative model keeps the equations simple, but it also means that the model-implied φ is always in the range $[0, 1]$. The observed data has larger numbers of banks, so to map that into suitable data for the model I measure the maximum possible φ in the data each period, that would be achieved if the Quoted Household Interest Rate was equal to the highest rate available that month. I divide the observed φ by the mean of these values (2.99) before HP-filtering to give an approximate mapping into the $\varphi \in [0, 1]$ range seen in the model.

F.2 Foreign VAR

Foreign variables are assumed to follow a VAR(4) process estimated outside of the model, as in Adolfson et al. (2007). Denoting the vector of foreign variables as Y_t, the structural

[50] I do not take logs of φ as on several occasions it is close to zero. This is therefore a measure of linearised, not log-linearised, φ. The observation equation is adjusted accordingly.

VAR process is:

$$F_0 Y_t = F_1 Y_{t-1} + F_2 Y_{t-2} + F_3 Y_{t-3} + F_4 Y_{t-4} + u_t \tag{88}$$

To identify the parameters, I start with the Adolfson et al. (2007) restrictions: output and inflation are assumed to be unaffected by contemporaneous shocks to anything other than themselves, but interest rates respond to both. As I have an extra variable not in Adolfson et al. (2007) (relative export prices), I add that inflation and output also do not respond contemporaneously to shocks to relative export prices. Furthermore, I assume that the foreign interest 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:

$$F_0 = \begin{bmatrix} 1, & 0, & 0, & 0 \\ 0, & 1, & 0, & 0 \\ -\gamma_\pi, & -\gamma_y, & 1, & 0 \\ -\gamma_\pi^p, & -\gamma_y^p, & \gamma_r^p, & 1 \end{bmatrix} \tag{89}$$

Where the order of variables in Y_t is inflation, output, interest rates, relative export prices. The model is over-identified. We cannot reject the over-identifying restrictions (p-value 0.87).

F.3 Calibration, priors, and estimation results

Table 9 gives descriptions of each calibrated parameter and its calibrated value. Table 10 gives descriptions of each estimated parameter and its prior. See Harrison and Oomen (2010) for the reasoning behind each calibrated value and prior except those specific to the attention block, which are discussed in Section 5.2.

Tables 11 and 12 show the estimation results for the baseline model and the full information model in Section 5.4 respectively. The variance decomposition in the full information and inattention models discussed in Section 5.4 is shown in Figure 6.

Table 9: Description of calibrated parameters

Parameter	Description	Value
α	Capital income share	0.3
β	Discount factor	0.99
δ	Depreciation rate	0.025
χ^{bf}	Net foreign asset adjustment cost	0.01
χ^{z}	Capital utilisation cost	$\beta^{-1}-1+\delta$
κ^{hv}	Share of domestic value added in home goods	0.935
κ^{xv}	Share of domestic value added in export goods	0.748
ψ^{m}	Expenditure weight of imports in consumption	0.248
ψ^{pm}	Imports Calvo parameter	0.4
σ^{hb}	Elasticity of substitution: goods varieties	9.668
σ^{m}	Elasticity of substitution: home vs. foreign goods	1.77
σ^{w}	Elasticity of substitution: labour varieties	8.3
σ^{x}	Elasticity of substitution: exports	1.5
σ^{xb}	Elasticity of substitution: export varieties	9.668
σ^{y}	Elasticity of substitution: labour vs. capital in production	0.5
$\bar{\sigma}(i^{n})$	Steady state standard deviation of interest rates	0.002*

*The steady state bad bank cost χ_1 is the parameter that adjusts to ensure this target is met.

Table 10: Description of estimated parameters

Parameter	Description	Prior Distribution
σ^{c}	Intertemporal elasticity of substitution	$N(0.66, 0.198)$
ψ^{hab}	Consumption habit parameter	$\text{Beta}(0.69, 0.05)$
σ^{h}	Labour supply elasticity	$N(0.43, 0.108)$
χ^{k}	Capital adjustment cost constant	$N(201, 60.3)$
ϵ^{k}	Indexation to past capital adjustment in capital adjustment cost	$\text{Beta}(0.5, 0.25)$
σ^{z}	Capital utilization cost elasticity	$N(0.56, 0.168)$
χ^{hv}	Domestic goods price adjustment cost	$N(326, 97.8)$
ϵ^{hv}	Domestic goods inflation indexation	$\text{Beta}(0.26, 0.1)$
χ^{xv}	Export goods price adjustment cost	$N(43, 12.5)$
ϵ^{xv}	Export goods inflation indexation	$\text{Beta}(0.14, 0.05)$
ψ^{pm}	Imported goods Calvo parameter	$\text{Beta}(0.40, 0.15)$
ϵ^{m}	Imported goods inflation indexation	$\text{Beta}(0.17, 0.05)$
ψ^{w}	Wage Calvo parameter	$\text{Beta}(0.21, 0.05)$
ϵ^{w}	Wage inflation indexation	$\text{Beta}(0.58, 0.145)$
θ^{p}	Taylor Rule inflation weight	$N(1.87, 0.131)$
θ^{y}	Taylor Rule output weight	$N(0.11, 0.028)$
θ^{rg}	Taylor Rule persistence	$\text{Beta}(0.87, 0.05)$
μ	Marginal cost of information	$\text{InvGamma}(0.005, 0.5)$
χ_2	Elasticity of inefficient bank costs to the policy rate	$N(0, 0.2)$
ρ_{tfp}	Persistence of TFP shock	$\text{Beta}(0.89, 0.05)$
σ_{tfp}	s.d. TFP shock	$\text{InvGamma}(0.006, 2)$
ρ_{g}	Persistence of government spending shock	$\text{Beta}(0.96, 0.025)$
σ_{g}	s.d. government spending shock	$\text{InvGamma}(0.009, 2)$
ρ_{x}	Persistence of shock x	$U(0.5, 0.289)$
$\sigma_{\zeta^{\kappa h}}$	s.d. labour disutility shock	$\text{InvGamma}(0.01, 2)$
$\sigma_{\zeta^{c}}$	s.d. monetary policy shock	$\text{InvGamma}(0.025, 2)$
$\sigma_{\zeta^{h}b}$	s.d. price markup shock	$\text{InvGamma}(0.006, 2)$
$\sigma_{\zeta^{k}}$	s.d. capital adjustment cost shock	$\text{InvGamma}(0.06, 2)$
σ_{y}	s.d. shock y	$\text{InvGamma}(0.001, 2)$
$\sigma_{\nu z}$	s.d. measurement error on z	$\text{InvGamma}(0.01, 2)$

$x = \zeta^{c}, \zeta^{hb}, \zeta^{k}, \zeta^{\mu}, \zeta^{\chi}, \zeta^{\chi b}$ refers to the shock to the risk premium, price markups, capital adjustment costs, information costs, interest rate level and dispersion. All other shocks are assumed i.i.d.
$y = \zeta^{rg}, \zeta^{\mu}, \zeta^{\chi}, \zeta^{\chi b}$. z contains the mean and standard deviation of bank deposit rates, and φ.

Table 11: Estimated posteriors in baseline model

Parameter	Mean	5%	95%	Parameter	Mean	5%	95%
σ^c	0.235	0.174	0.294	$\rho_{\zeta^{hb}}$	0.276	0.048	0.489
ψ^{hab}	0.744	0.679	0.814	ρ_{ζ^k}	0.797	0.601	0.976
σ^h	0.448	0.282	0.601	ρ_{ζ^χ}	0.921	0.862	0.978
χ^k	148.258	46.917	236.721	$\rho_{\zeta^{\chi b}}$	0.785	0.700	0.877
ϵ^k	0.132	0.001	0.269	μ	0.035	0.025	0.044
σ^z	0.549	0.280	0.829	χ_2	-0.264	-0.476	-0.066
χ^{hv}	415.793	277.100	553.532	σ_g	0.033	0.028	0.038
ϵ^{hv}	0.228	0.088	0.375	$\sigma_{\zeta^{\kappa h}}$	1.743	0.764	2.817
χ^{xv}	34.691	10.374	57.552	$\sigma_{\zeta^{rg}}$	0.001	0.001	0.002
ϵ^{xv}	0.135	0.058	0.206	σ_{tfp}	0.007	0.006	0.008
ψ^{pm}	0.639	0.362	0.882	σ_{ζ^c}	0.009	0.006	0.011
ϵ^m	0.162	0.075	0.238	$\sigma_{\zeta^{hb}}$	0.007	0.005	0.008
ψ^w	0.260	0.193	0.328	σ_{ζ^k}	0.248	0.056	0.503
ϵ^w	0.339	0.185	0.510	σ_{ζ^μ}	0.002	0.000	0.004
θ^p	1.812	1.586	2.021	σ_{ζ^χ}	0.003	0.002	0.004
θ^y	0.149	0.109	0.190	$\sigma_{\zeta^{\chi b}}$	0.003	0.002	0.004
θ^{rg}	0.910	0.890	0.931	$\sigma_{\nu\varphi}$	0.094	0.076	0.109
ρ_{tfp}	0.957	0.934	0.980	$\sigma_{\nu s}$	0.007	0.003	0.013
ρ_g	0.957	0.928	0.984	$\sigma_{\nu m}$	0.002	0.001	0.002
ρ_{ζ^c}	0.908	0.853	0.963				

Table 12: Estimated posteriors in full information model

Parameter	Mean	5%	95%	Parameter	Mean	5%	95%
σ^c	0.186	0.105	0.268	$\rho_{\zeta^{hb}}$	0.272	0.050	0.487
ψ^{hab}	0.725	0.654	0.793	ρ_{ζ^k}	0.703	0.500	0.921
σ^h	0.440	0.289	0.616	ρ_{ζ^χ}	NA	NA	NA
χ^k	170.175	66.864	268.994	$\rho_{\zeta^{\chi b}}$	NA	NA	NA
ϵ^k	0.175	0.001	0.370	μ	NA	NA	NA
σ^z	0.511	0.219	0.777	χ_2	NA	NA	NA
χ^{hv}	414.783	272.223	554.419	σ_g	0.033	0.028	0.038
ϵ^{hv}	0.217	0.065	0.360	$\sigma_{\zeta^{\kappa h}}$	2.127	0.746	3.430
χ^{xv}	33.942	3.837	56.226	$\sigma_{\zeta^{rg}}$	0.001	0.001	0.002
ϵ^{xv}	0.140	0.056	0.221	σ_{tfp}	0.007	0.006	0.008
ψ^{pm}	0.656	0.422	0.901	σ_{ζ^c}	0.013	0.007	0.018
ϵ^m	0.165	0.083	0.240	$\sigma_{\zeta^{hb}}$	0.007	0.005	0.008
ψ^w	0.241	0.163	0.307	σ_{ζ^k}	0.193	0.073	0.300
ϵ^w	0.321	0.178	0.471	σ_{ζ^μ}	NA	NA	NA
θ^p	1.851	1.656	2.077	σ_{ζ^χ}	NA	NA	NA
θ^y	0.148	0.107	0.188	$\sigma_{\zeta^{\chi b}}$	NA	NA	NA
θ^{rg}	0.913	0.892	0.934	$\sigma_{\nu\varphi}$	NA	NA	NA
ρ_{tfp}	0.962	0.938	0.986	$\sigma_{\nu s}$	NA	NA	NA
ρ_g	0.951	0.914	0.988	$\sigma_{\nu m}$	NA	NA	NA
ρ_{ζ^c}	0.889	0.834	0.951				

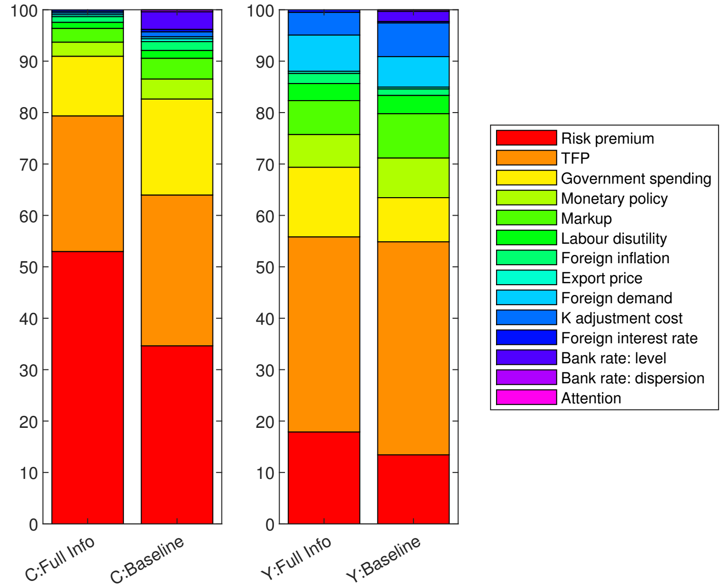


Figure 6: Percentage of the variance of consumption and output due to each shock in the full information and variable attention models.

The risk premium shock is the bottom segment of each bar (displayed in red). As described in Section 5.4, without information frictions the risk premium shock explains 53% of the variance of consumption. It explains the second largest share of output variance (18%), after TFP shocks (38%).

In moving to the baseline model with inattention the risk premium shock becomes less important, explaining 35% of consumption variance and 13%. of output variance. This is not picked up by shocks to attention, which explain a negligible fraction 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 38% and 7%, but with inattention they explain 41% and 9%. For the variance of consumption, government spending, TFP and price markup shocks explain 12%, 26% and 3% with full information, and 19%, 29% and 4% in the baseline model. Shocks to the level of bank interest rates (ζ^{χ}) are the only new shocks to play a non-negligible role in fluctuations, explaining 2% of output variance and 3% of consumption variance.