Department of Economics
Discussion Paper Series

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

Alistair Macaulay

Number 931
January, 2021

Department of Economics
Manor Road Building, Manor Road
Oxford, OX1 3UQ

ISSN 1471-0498

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

Alistair Macaulay†

January 22, 2021

Abstract

I show that if it is costly for households to process information about asset returns, a model with ex-ante identical households features persistent inequality. The steady state has a two-agent structure, with inequality maintained by a complementarity between attention and wealth: wealthy households have stronger incentives to pay attention to asset choices, and so earn higher returns than asset-poor households. Fiscal expansions are less powerful in this model than in a standard model with heterogeneous discount factors, because when an expansion causes poor households to start saving, they also increase their attention. They therefore earn higher interest rates, and so save even more, smoothing the windfall from the policy over a longer time period. I provide evidence for this increase in attention using cross-state variation in uncertainty about savings interest rates in the aftermath of the 2017 Tax Cuts and Jobs Act in the US. The effects of fiscal policy therefore depend not just on the existence of inequality, but also on the cause of that inequality.
Keywords: Rational Inattention, Inequality, TANK, Fiscal Policy

*I thank Klaus Adam, Isaac Baley, Paul Beaudry, Florin Bilbiie, Martin Ellison, Cosmin Ilut, Filip Matějka, Matthew McKernan, Michael McMahon, Michalis Rousakis, Veronika Selezneva, Christopher Sims, Ctirad Slavik, and seminar participants at the University of Oxford and CERGE-EI for helpful comments and suggestions.
†Department of Economics & St. Anne's College, University of Oxford. Email: alistair.macaulay@economics.ox.ac.uk

1 Introduction

A recent literature has demonstrated that the effects of aggregate shocks are crucially dependent on their distributional consequences (see Kaplan and Violante (2018) for a review). Models in this literature often assume that inequality is driven by exogenous factors[1], but an empirical literature has highlighted a variety of forces contributing to wealth inequality which plausibly respond to aggregate shocks, with heterogeneity in asset returns a prominent example (e.g. Benhabib et al. (2019), Fagereng et al. (2020)). This raises a question: we know that the *existence* of wealth inequality matters for the transmission of aggregate shocks, but does the *source* of that inequality matter? In particular, does incorporating returns heterogeneity affect shock transmission in these models?

In this paper I show that the source of inequality does matter. I develop a two-agent New Keynesian (TANK) model in which wealth inequality is entirely due to returns heterogeneity, through an 'attention trap': a complementarity between wealth and asset returns that emerges from optimal household decisions of how much attention to pay to saving choices. Focusing on fiscal policy shocks, often the point of greatest difference between heterogeneous-agent and representative-agent models (Kaplan and Violante, 2018), I show that debt-funded tax cuts have smaller and more persistent effects in this model than in a comparable TANK model where inequality is driven by discount factor heterogeneity (as in Bilbiie et al., 2013).

This difference arises because the attention decisions that drive returns heterogeneity are endogenous to fiscal policy. After a tax cut, low-wealth households increase their saving. This leads them to pay more attention to their choice of savings product, as they now have more to gain from discovering which products are offering the highest interest rates (Arrow, 1987) - they have more 'skin in the game'. The subsequent increase in the interest rate faced by poor households pushes them to save even more, and so they smooth the tax cut windfall over more periods than in the heterogeneous discount factor model. The marginal propensity to consume (MPC) of poor households is therefore lower in the attention trap model. I provide empirical evidence that attention does indeed respond to fiscal policy in this way using cross-state variation in a survey-based measure of uncertainty about savings interest rates in the aftermath of the 2017 Tax Cuts and Jobs Act (TCJA) in the US.

[1]Some households are at their borrowing constraint in these models because they cannot access financial markets (e.g. Bilbiie (2008)), because they are less patient than 'saver' households (e.g. Bilbiie et al. (2013)), or because they have received negative productivity shocks (e.g. Kaplan et al. (2018)). In two-agent models such as Bilbiie (2008) all households have zero wealth, but only some are on their Euler equation. For simplicity I also refer to this heterogeneity as 'wealth inequality', as in this paper and the heterogeneous-agent literature these households are unconstrained because of their liquid wealth.

The attention trap model differs from most TANK models in that households are identical ex-ante, with the same preferences and access to asset markets. Households must decide how much information to process before choosing between different savings products, some of which offer higher interest rates than others. As in the literature on rational inattention (e.g. Sims, 2003), I refer to this information processing as attention. Attention is costly, but with more attention the household is more likely to identify which savings products are offering the best interest rates.

This information problem generates persistent inequality because attention choices depend on wealth. Wealthier households have more to gain from increasing their interest rate because that higher rate operates over a larger pool of assets, and so they pay more attention and earn higher rates (as in Arrow, 1987). If two households are not on exactly equal wealth, the wealthier household will therefore experience higher interest rates, and so will save more, than the less wealthy household. The difference in interest rates acts just like the heterogeneous discount factors in Bilbiie et al. (2013), and as in their model it leads to a steady state with two types of households, wealthy households on their Euler equation and poor households against a borrowing constraint. The poor households do not save, even in steady state, because they optimally pay no attention to choosing between saving products and so face very low interest rates.

In fact, the steady state of this model is identical to that in a model where inequality is maintained by heterogeneous discount factors. The two models also give the same response to any shocks that are sufficiently small that all poor households remain at their borrowing constraint. In that case, poor household attention remains at zero and they continue to consume all of their income.

If, however, the fiscal expansion is sufficiently large that poor households temporarily leave the borrowing constraint, the attention trap and heterogeneous discount factor models diverge because the poor households in the attention trap model start to process information. This reduces their MPC, and so the impact of the policy on output.

In addition, this interaction between fiscal policy and attention means that any features of the environment that affect the costs and benefits of attention will also impact the power of fiscal policy, through poor household attention decisions. Lower costs of information about savings choices increase the attention poor households pay once their saving is positive after a fiscal expansion. This further increases their interest rates and reduces their MPCs, and accordingly reduces the impact of the fiscal expansion. Lower taxes on asset income have a similar effect, for the same reason.

Finally, if the interest rates on offer become less dispersed, tax cuts will have even smaller and more persistent effects on aggregate consumption. The highest interest rate in the market in steady state is unchanged, because it is pinned down by the wealthy household Euler equation. Lower dispersion therefore implies that the interest rate faced by poor households is higher for any given level of attention, so there are stronger incentives for poor households to save. They will start to save after a smaller fiscal stimulus, and for a given stimulus they will save more than if rate dispersion is high. Despite the fact that lower interest rate dispersion reduces the incentive to pay attention to savings choices, the MPC of poor households is therefore lower when interest rate dispersion is low, and fiscal policy will be weaker as a result.

The key mechanism of the paper is therefore that the attention of poor households rises after large fiscal expansions. I provide evidence for this using cross-state heterogeneity in the effects of the 2017 Tax Cuts and Jobs Act. Liu and Williams (2019) find that GDP responds significantly to corporate tax shocks in fewer than 50% of states. In the model, fiscal expansions affect attention through household income, so tax cuts should have a greater impact on poor household attention in states where economic activity responds positively to such cuts. I therefore take the states that Liu and Williams (2019) find respond significantly to corporate tax cuts to be the treatment group[2] in a difference-in-differences model, and estimate the effect of the TCJA on a measure of uncertainty about the interest rate on savings accounts from the Survey of Consumer Expectations.

As predicted by the attention trap model, uncertainty about savings interest rates falls more after the tax cut in the responsive states. This reaction is concentrated in low income households who are not at risk of being unable to pay their bills. This is consistent with the model, where the attention response comes from poor households who start saving on receipt of a tax cut. There is no significant difference in the paths of uncertainty before the tax cut.

The TCJA highlights that large fiscal expansions are not theoretical curiosities[3], and it is not the only recent example. Fiscal stimulus packages in the US during the Great Recession and Covid-19 pandemic were substantially larger as a portion of GDP than any expansion considered in this paper[4]. In addition, in a less simplified model with a richer wealth distribution, there would always be some households close to lifting off their borrowing constraints. Even small expansions will therefore lead to some households starting to save, and the analysis in this paper suggests that they play an important role in the effects of shocks on consumption.

---

[2]The corporate tax cut was the largest single tax change in the TCJA, but other taxes changed as well. The 19 responsive states in my treatment group contain 4 of the 5 states most exposed to the income tax changes in the act as estimated by Kumar (2020), using the method of Zidar (2019).

[3]It is estimated that the TCJA will have added $1-$2 trillion to the federal debt by 2027 (Joint Committee on Taxation, 2017).

[4]The 2009 American Recovery and Reinvestment Act added $787 billion to US government debt (Dupor, 2017). The 2020 CARES Act was even larger, amounting to more that 10% of US GDP and involving $500 billion of direct payments to households, which are similar to the lump sum tax cuts studied in this paper.

The results are particularly pertinent in the current environment of interest rates close to the zero lower bound, especially if interest rates remain low for a prolonged period. The lack of room for conventional monetary easing may mean that fiscal policy has to play a greater role in stabilising future recessions (Blanchard and Summers, 2019)[5]. However, extremely low interest rates may also lead to lower interest rate dispersion as retail banks are unwilling to pass on negative interest rates to households (Eggertsson et al., 2019). The attention trap model suggests that this may weaken the expansionary power of fiscal policy precisely when it is most needed.

Related literature: This paper contributes principally to the literature on the role of household heterogeneity in macroeconomic policy. Several recent papers have shown that heterogeneity in wealth and access to financial markets can have substantial effects on the power of different policies[6]. While many of these papers study models with a rich and realistic distribution of wealth, Debortoli and Galí (2018) show that a simpler two-agent (TANK) model based on Bilbiie (2008), with households divided into standard intertemporal optimisers and hand-to-mouth consumers, can match many of the effects of heterogeneity on policy[7]. The model I build has a similar two-agent structure, but in contrast to Bilbiie (2008) and others in the TANK literature the hand-to-mouth households are at their borrowing constraint because of endogenous attention decisions. To the best of my knowledge, I am the first to incorporate the insight that heterogeneous asset returns can explain a significant fraction of wealth inequality (e.g. Benhabib et al. (2019)) into a TANK model, and I show that modeling inequality in this way does alter the transmission of fiscal policies compared to a model with exogenously-driven inequality[8].

The link between attention to saving choices and wealth at the heart of the attention trap was first proposed by Arrow (1987). Since then, a number of papers have assessed the quantitative importance of this mechanism for wealth inequality in rich heterogeneous agent models[9]. McKay (2013) also uses a model of this kind to quantitatively study the welfare effects of social security privatisation. This paper extends this literature by

[5]See the calls from several central bankers for fiscal expansion in the recession caused by the Covid-19 pandemic (e.g. Powell (2020), Lagarde (2020)).

[6]For monetary policy, see Gornemann et al. (2016), McKay et al. (2016), Kaplan et al. (2018), among others. For fiscal policy, see Auclert et al. (2018), Hagedorn et al. (2019), Bayer et al. (2020), among others. Werning (2015) and Acharya and Dogra (2020) show that the cyclicality of idiosyncratic income risk is critical in determining whether household heterogeneity amplifies or dampens shocks. Relatedly, Bilbiie (2019a) shows in an analytic HANK model that amplification occurs only if the ratio of the income of hand-to-mouth households to total income is countercyclical. For empirical evidence on the effects of household heterogeneity on monetary policy, see Almgren et al. (2019) and Holm et al. (2020).

[7]Other papers on TANK models include Galí et al. (2007), Giambattista and Pennings (2017), Bilbiie (2019b), and Cantore and Freund (2020).

[8]This comparison model, and the consideration of the effects of debt-funded tax cuts, is most closely related to Bilbiie et al. (2013), where inequality is driven by heterogeneity in household discount factors.

[9]See e.g. Campanale (2007), Lusardi et al. (2017), Kacperczyk et al. (2019), Lei (2019).

showing that the Arrow (1987) mechanism delivers a two-agent structure in an otherwise standard New Keynesian model, and that fiscal policy in such a model is less powerful than in an equivalent two-agent model with exogenous wealth heterogeneity.

In support of this mechanism, Fagereng et al. (2020) find in Norwegian data that wealth is correlated with asset returns even within asset classes. In Sweden, Briggs et al. (2020) find that stock market participation increases after lottery wins, and that a model with plausible levels of financial participation costs over-predicts the size of this response. This indicates that other costs, such as the cost of paying attention, are substantial. I contribute to this empirical literature in finding that the Tax Cuts and Jobs Act (TCJA) led to a fall in uncertainty about savings markets among poor households, as predicted by the model. This is the first evidence directly linking information about savings to fiscal policy. In studying this, I also contribute to the literature on the heterogeneity of the effects of tax shocks across states in the US, and the effects of the TCJA specifically[10].

A closely related paper is Best and Dogra (2019). To my knowledge, theirs is the only other paper in which endogenous returns heterogeneity leads to ex-ante identical households being divided into wealthy and poor in steady state. Their focus, however, is on the stability of different equilibria, so they build a Ramsey model which is sufficiently tractable to solve analytically. Since my focus is on policy and shock transmission, I extend their work by showing that a two-agent steady state exists in an otherwise standard New Keynesian model, and I explore how the implied wealth inequality responds to and affects the transmission of fiscal policies.

I also contribute to the literature on rational inattention in macroeconomics. Beginning with Sims (2003), many papers have explored how costly attention affects macroeconomic models when agents process information about normally distributed random variables[11]. While this work has been instructive and valuable, Sims (2006) notes that rational inattention also has potential applications in many situations where the assumption of normality is not appropriate. This paper develops one such application, using the discrete choice framework of Matějka and McKay (2015) to embed rational inattention to savings product choices into a New Keynesian model. This discrete choice framework has previously been applied to importing (Dasgupta and Mondria, 2018) and hiring decisions (Acharya and Wee, 2020). This is the first paper to use it to model saving choices[12].

Finally, it is often argued that fiscal policy must have a greater role in stabilising shocks as the natural rate of interest falls and policy rates become more likely to hit their lower bound (e.g. Blanchard and Summers (2019)). My model highlights a potential

[10]See e.g. Liu and Williams (2019), Occhino (2019), Zidar (2019), Kumar (2020).

[11]See e.g. Adam (2009), Maćkowiak and Wiederholt (2009, 2015), Afrouzi (2020), among others.

[12]I use a similar framework in Macaulay (2020) to study the cyclicality of the attention choices of a representative household, rather than the interaction between attention and wealth that I analyse here.

limiting force on the power of fiscal policy in such an environment if low real interest rates are associated with low interest rate heterogeneity. Low rate heterogeneity implies that a given fiscal policy will raise poor households further away from their borrowing constraint, which means they smooth the gains from the expansion more over time, reducing the impact of the policy. Fiscal policy will continue to have an effect at the zero lower bound, and as long as there is some interest rate dispersion this effect will be stronger than in a representative agent model, but the model does suggest that fiscal policy may not be as powerful in a low interest rate environment as implied by models with exogenous wealth heterogeneity.

The rest of the paper is organised as follows: section 2 sets out the attention trap model and demonstrates how the feedback from wealth to attention leads to a two-agent structure. Section 3 compares the effects of debt-funded tax shocks in this model with the effects in a comparison model with discount factor heterogeneity. Section 4 provides empirical evidence for the link from tax cuts to attention, and section 5 concludes.

2 Attention Trap Model

The model is a standard New Keynesian model in many respects, except that there are two savings products available to households each period, with different interest rates. Households must process costly information to discover which offers the better deal. In this way attention (i.e. information processing) will lead households to achieve higher interest rates on their assets (as in Arrow, 1987).

I restrict households to be net savers, so they always prefer higher interest rates. This simplifies the household information problem, but is not necessary for the results. I discuss this further in section 2.5. The firms and monetary authority are the same as in many standard New Keynesian models, so I leave their problems for appendix A.1.

2.1 Savings Products

Each period, many financial institutions buy one-period government bonds at a price of $(1 + i_t)^{-1}$, then sell them to households at marginal cost. Every period, the institutions draw their marginal costs from an i.i.d. distribution[13]. With probability g, they have no marginal cost beyond the bond price paid to the government. I shall refer to these as the 'good' savings products. With probability $1 - g$, the financial institution faces an additional marginal cost of $\tau > 0$. The 'bad' products from these institutions are

[13] Assuming i.i.d. costs simplifies the rational inattention problem facing households, as it means that information processed about dealers in one period has no value in future periods. See Steiner et al. (2017) for results on information processing problems in dynamic settings.

therefore sold to individuals at price:

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

The 'bad' products have a higher price, and so a lower interest rate ($i_t^b < i_t$). As net savers, all households prefer the higher interest rates of the good products to the bad[14].

2.2 Households

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

The household sets the (real) quantity of bonds b_t that each individual will buy before knowing which of the two available bond prices they face, and there is perfect consumption sharing around the household. Each individual chooses one savings product[15]. They choose a 'good' product with probability q_t, which is a function of information gathered $\mathcal{I}$. I refer to this information processing as attention, as is common in the rational inattention literature. There is a constant marginal cost μ to increasing $\mathcal{I}$.

These assumptions imply that the household faces the average bond price:

$$\frac{1}{1+\tilde{i}_t} = \frac{q_t}{1+i_t} + \frac{1-q_t}{1+i_t^b} \tag{2}$$

Each individual faces a discrete choice rational inattention problem (see Matějka and McKay, 2015). As they always prefer the good savings product, they seek to maximise the success probability q. Using the entropy-based measure of information processing common in the rational inattention literature (e.g. Sims, 2003), we have[16] for $q \in [g, 1)$:

$$\mathcal{I}_t = q_t \log\left(\frac{q_t}{g}\right) + (1-q_t)\log\left(\frac{1-q_t}{1-g}\right) \tag{3}$$

With no information processing, individuals choose savings products at random and so

[14] The borrowing constraint is set to zero, so households at the borrowing constraint are indifferent between the two products, but any household saving a positive amount strictly prefers the good products.

[15] Redistributing consumption around the household means that individuals are risk neutral with respect to the bond price, seeking only to minimise their expected bond price. Their risk-neutrality means there is nothing to gain from spreading their savings across multiple banks, so they choose a single bank for their portion of the household's saving.

[16] This function is derived in appendix B.1, where it is shown that the relationship between information processing and the success probability q is unaffected by the number of products available. The proportion of products which are good (g) is all that matters.

choose a good product with probability $q = g$. As attention increases, the individual processes more information before deciding on a product, and so chooses successfully with a greater probability. The probability of choosing a good product is concave in information processing, so there are diminishing returns to attention[17].

The household problem is to choose consumption c_t, labour supply n_t, real bond purchases b_t, and the effective interest rate $\tilde{i}_t$ (reflecting attention choice) to solve:

$$\max_{c_t,n_t,b_t,\tilde{i}_t} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{c_t^{1-\sigma}}{1-\sigma} - \frac{n_t^{1+\varphi}}{1+\varphi} - \mu \mathcal{I}(\tilde{i}_t) \right) \tag{4}$$

subject to

$$c_t + \frac{b_t}{1+\tilde{i}_t} = \frac{b_{t-1}}{\Pi_t} + w_t n_t + \mathcal{D}_t - T_t \tag{5}$$

$$b_t \geq 0 \tag{6}$$

Equation 5 is the budget constraint in real terms. Π_t is gross inflation, w_t is the real wage, $\mathcal{D}_t$ is profits from firms, and T_t is a lump sum tax. The effective interest rate is an increasing function of the probability of successfully choosing a good savings product q_t, and information processing $\mathcal{I}$ is an increasing function of q_t. Therefore $\mathcal{I}'(\tilde{i}_t) > 0$. Diminishing returns to attention ensure that $\mathcal{I}''(\tilde{i}_t) > 0$.

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

$$c_t^{-\sigma} = \beta(1+\tilde{i}_t)\mathbb{E}_t \left(\frac{c_{t+1}^{-\sigma}}{\Pi_{t+1}} \right) + (1+\tilde{i}_t)\lambda_t \tag{7}$$

$$c_t^{\sigma} n_t^{\varphi} = w_t \tag{8}$$

$$b_t c_t^{-\sigma} = \mu(1+\tilde{i}_t)^2 \mathcal{I}'(\tilde{i}_t) \tag{9}$$

λ_t is the Lagrange multiplier on the borrowing constraint, which is positive when the constraint binds and zero otherwise. The first order condition on attention (equation 9) implies that a wealthier household will choose to process more information, and so will experience a higher interest rate. This encourages further saving through the Euler equation (equation 7), but the non-concavity this implies is small enough at plausible parameter values that the first order conditions remain sufficient for utility maximisation (proof in appendix B.2). This interaction between attention and wealth is the reason this model has a steady state with two types of households, some wealthy and others at the

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

borrowing constraint, despite those households having identical preferences and access to asset markets. In the remaining sections I denote a variable pertaining to wealthy or poor households only with a superscript w or p respectively, and let ω be the proportion of wealthy households[18].

2.3 Government

Government issues a positive amount of short term bonds each period and funds its interest expenditures with lump sum taxes. This could be (for example) because the government wants to supply some bonds to ensure that households have access to a risk-free asset if they want it.

Let $\bar{b}_t$ be the real supply of bonds issued in period t. The real government budget constraint is:

$$\frac{\bar{b}_{t-1}}{\Pi_t} = \frac{\bar{b}_t}{1+i_t} + (\omega T_t^w + (1-\omega)T_t^p) \tag{10}$$

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

The lump sum taxes on the two types of households are always equal, except in the event of the i.i.d. redistributive tax shocks considered in appendix C.

Bond market clearing implies:

$$\bar{b}_t = \omega b_t^w + (1-\omega)b_t^p \tag{11}$$

2.4 Closing the Model

The firm problem gives an aggregate production function and a Phillips curve (derived in appendix A.1), and the interest rate on government debt is set by a monetary authority according to a Taylor rule:

$$Y_t = A_t(\omega n_t^w + (1-\omega)n_t^p)^{1-\alpha} \tag{12}$$

$$Y_t(1-\epsilon) + \frac{\epsilon}{1-\alpha} w_t \left(\frac{Y_t}{A_t}\right)^{\frac{1}{1-\alpha}} - \psi(\Pi_t - 1)\Pi_t Y_t + \psi\beta\mathbb{E}_t(\Pi_{t+1}-1)\Pi_{t+1}Y_{t+1}\left(\frac{\bar{c}_t}{\bar{c}_{t+1}}\right)^{\sigma} = 0 \tag{13}$$

$$1 + i_t = \bar{i} + \left(\frac{\Pi_t}{\bar{\Pi}}\right)^{\delta_\Pi}\left(\frac{Y_t}{\bar{Y}}\right)^{\delta_Y} \tag{14}$$

[18] As noted in section 2.5 below, which households end up poor is a function of past idiosyncratic shocks, so in principle ω could vary over time. I abstract from that here, assuming that in some previous period idiosyncratic shocks divided the households, but that there are no further idiosyncratic shocks.

Where A_t is TFP, α is the degree of diminishing returns to scale in the production function, ϵ is the elasticity of substitution between intermediate goods varieties, ψ is the Rotemberg (1982) price adjustment cost parameter, $\bar{c}_t$ is aggregate consumption, and $\bar{i}$ is the nominal interest rate target.

All output Y_t must be consumed in period t by households, be used to pay price adjustment costs in firms, or be lost to transaction costs in bad financial institutions. The market clearing condition is:

$$Y_t = \omega c_t^w + (1-\omega)c_t^p + \frac{\psi}{2}(\Pi_t - 1)^2 Y_t + \left(\frac{1}{1+i_t^b} - \frac{1}{1+i_t}\right)\left(\omega(1-q_t^w)b_t^w + (1-\omega)(1-q_t^p)b_t^p\right) \quad (15)$$

This completes the model.

2.5 Steady States

There are two steady states of this model. As well as a representative-agent steady state, there is also a steady state with households separated into two types, wealthy (w) households on their Euler equation and poor (p) agents up against their borrowing constraint. This is the same form as the two-agent models in Bilbiie (2008) and others, even though (unlike those existing models) the households have identical preferences and access to asset markets. Poor households do not converge to the positive wealth level of the rich households because their lack of wealth means they do not have an incentive to pay attention to choosing good savings products, and so they face low interest rates, which encourages them not to save.

To see this, consider the model with no costs of attention. All households earn the same interest rate on any savings, and the Euler equations for the wealthy and poor households respectively in steady state are:

$$1 + \bar{i} = \frac{1}{\beta} \quad (16)$$

$$1 + \bar{i} = \frac{1}{\beta + \lambda^p \bar{c}^{p^\sigma}} \quad (17)$$

These can only both be true if $\lambda^p = 0$, which implies that the poor household is not against the borrowing constraint, and therefore that the households are identical. Existing two agent New Keynesian models obtain steady state inequality either by assuming that the Euler equation doesn't hold for poor households because they are barred from asset markets (e.g. Bilbiie, 2008) or by assuming that the two households have different β

parameters (e.g. Bilbiie et al., 2013).

With costly attention, the steady state Euler equations become:

$$\frac{1}{1+\bar{\tilde{i}}^w} = \frac{\bar{q}^w}{1+\bar{i}} + \frac{1-\bar{q}^w}{1+\bar{i}^b} = \beta \tag{18}$$

$$\frac{1}{1+\bar{\tilde{i}}^p} = \frac{\bar{q}^p}{1+\bar{i}} + \frac{1-\bar{q}^p}{1+\bar{i}^b} = \beta + \lambda^p \bar{c}^{p^\sigma} \tag{19}$$

Substituting out for the bad interest rate in terms of the good (using equation 1) and combining we obtain:

$$\tau(\bar{q}^w - \bar{q}^p) = \lambda^p \bar{c}^{p^\sigma} \tag{20}$$

If wealthy households pay more attention to choosing between bond dealers than poor households (i.e. if $\bar{q}^w > \bar{q}^p$) then $\lambda^p > 0$ in steady state, which implies that the poor household is at their borrowing constraint. In steady state, the first order condition on attention is:

$$\bar{b}^h \bar{c}^{h^{-\sigma}} = \mu(1+\bar{\tilde{i}}^h)^2 \mathcal{I}'(\bar{\tilde{i}}^h), \quad h \in \{w,p\} \tag{21}$$

The increase in information processing required for a marginal increase in the effective interest rate, $\mathcal{I}'(\bar{\tilde{i}}^h)$, is positive and convex. The marginal information is zero when no information is processed:

$$\mathcal{I}'(\tilde{i}) = 0 \iff \mathcal{I}(\tilde{i}) = 0 \iff q = g \tag{22}$$

A poor household at the borrowing constraint has zero assets, so $\bar{b}^p = 0$, which in equation 21 implies that they optimally pay no attention to their choice of bond dealer. As household wealth $\bar{b}^h$ rises the left hand side of equation 21 rises, and so therefore does attention[19]. This is because the costs of information do not depend on wealth, but the benefits of being more informed are higher interest rates on the entire stock of wealth saved by the household. If more is saved, the benefits of information are larger[20].

There is therefore a steady state in which wealthy households pay more attention to choosing between bond dealers than poor households who are against their borrowing constraint. Poor households have no 'skin in the game', and so pay no attention. This means that they face lower interest rates on average, which reduces their incentive to save. This is the 'attention trap', and it gives a two-agent steady state without hetero-

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

[20]This mechanism is the same as in Arrow (1987) and the literature that followed. Evidence of this can be found in Briggs et al. (2020), Fagereng et al. (2020), and section 4 of this paper.

geneity in household preferences or access to asset markets. In fact, the feedback between attention and wealth renders the representative-agent steady state unstable with respect to idiosyncratic shocks: if a shock leaves a household poorer than others, that poorer household pays less attention, earns a lower return on their saving, and so saves even less, causing the wealth difference to grow until the poor household hits the borrowing constraint (see appendix D).

A further implication of equation 20 is that if wealthy households pay less attention to their savings product choices (for example, if the cost μ rises), or if τ falls, then $\lambda^p \bar{c}^{p^\sigma}$ must fall. λ^p is the marginal utility of a loosening of the borrowing constraint to a poor household, so a fall in this implies that poor households must be closer to lifting off that constraint, which will be important in section 3.3. As long as wealthy households process some positive amount of information (as long as μ remains finite and consumption is positive), and there remains some interest rate heterogeneity among bond dealers ($\tau > 0$), then the poor households will remain at the borrowing constraint in steady state ($\lambda^p > 0$).

Equation 20 also highlights that the two-agent steady state does not rely on the borrowing constraint being precisely at zero, or on there being a positive supply of government bonds. The key requirement is that the probability of choosing low-price bonds is higher for wealthy households than for poor households against their borrowing constraint. If the model had a zero net supply of bonds and a negative minimum asset limit (i.e. some borrowing is allowed) the results would be the same. Poor households would borrow by selling bonds to bond dealers, and wealthy households would save by buying them. Borrowers would process information to try and find the dealers offering the highest price (the lowest interest rate), while wealthy households would aim for a low bond price as in the model above. Just as in this model, poor households would therefore face lower interest rates than wealthy households, which would sustain the wealth inequality.

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

3 Fiscal Policy

In this section I show that large fiscal expansions have smaller and more persistent effects on aggregate consumption in the model developed in section 2 than in a model where inequality is due to heterogeneous discount factors. In both models, large expansions cause poor households to temporarily lift off the borrowing constraint, but in the attention trap model those poor households also then pay more attention to their choice of savings product. They therefore earn higher interest rates, and so they save even more of the windfall from the policy, reducing the impact of that policy on aggregate consumption. This also prolongs the period over which the poor households spend the windfall and return to the borrowing constraint, which makes the effects of the policy more persistent.

If we instead consider small fiscal expansions, which leave poor households at the borrowing constraint every period (as in Bilbiie et al., 2013), then the only difference between the attention trap model and the heterogeneous discount factor model comes from the way the attention of wealthy households reacts to policy. As this is less clear than the attention response of poor households I abstract from this issue by considering the case where information costs are sufficiently low that wealthy households pay enough attention that the probability of choosing a bad product $1-q^w$ is negligible in all periods[21]. For small fiscal expansions the attention trap and heterogeneous β models are therefore identical, but this is not the case for larger fiscal expansions that cause poor households to temporarily save some of their income.

The specific fiscal expansion studied here is an unanticipated one-period reduction in the lump sum taxes faced by each household, funded by an increase in government debt[22]. The level of real debt decays back to steady state following an AR(1) process with persistence ρ.

3.1 The Comparison Model

To assess the importance of the source of inequality for the effects of fiscal policy I compare the TANK model from section 2 with a model in which there is no attention problem, and steady state wealth inequality is driven by discount rate heterogeneity, as in Bilbiie et al. (2013). I choose this, rather than the model of limited asset market participation in Bilbiie (2008), to allow the poor households in the comparison model to save positive amounts if the tax windfall is sufficiently large, as they do in the attention trap model.

The model is identical to that in section 2, except that the cost of information is

[21]For wealthy households, small changes in wealth after a fiscal policy shock are outweighed in the attention decision by changes in the interest rate distribution for most calibrations of the model. In this simplified model, the dispersion of interest rates rises mechanically with the policy rate. However, in other work (Macaulay, 2020) I show that this is not how interest rate dispersion behaves in the data. I therefore leave discussion of that result, and cyclical patterns in the attention of wealthy households, to that paper. There I also discuss the interaction of attention with monetary policy shocks, which differs from the case of fiscal policy because of the response of interest rate dispersion. In the calibration used for simulations q^w remains within 10^{-40} of 1 in all periods of the simulations.

[22]This is one of the policies studied in Bilbiie et al. (2013). I show in appendix C that the qualitative differences between the attention trap and heterogeneous β model are the same after the other kind of expansion they consider, an intratemporal redistribution from wealthy to poor households.

$\mu = 0$, and a proportion $1 - \omega$ of households have a lower discount factor than other households, $\beta^p < \beta$. The Euler equations of the two household types in steady state together imply:

$$\beta = \beta^p + \lambda^p \bar{c}^{p^\sigma} \tag{23}$$

At any given interest rate, poor households save less than wealthy (high β) households. The only steady state therefore has a two-agent structure similar to that in the model in section 2, with poor households at the borrowing constraint ($\lambda^p > 0$). The key difference is that here wealth inequality is driven by preferences, which are exogenous to fiscal policy, whereas in the attention trap model inequality is maintained by endogenous attention choices. This means that the identity of those at the borrowing constraint is random in the attention trap model, depending only on initial endowments, while here the poor households are poor because they are less patient than others.

The full model equations are given in appendix A.2. For the comparisons I calibrate this model and the attention trap model using the same parameters where possible, which means they have the same steady state. I choose β^p and the parameters of the attention problem such that λ^p, the shadow value of relaxing the borrowing constraint for poor households, is the same in the two models.

3.2 Debt-funded Tax Cuts

Tax cuts to all households funded by increases in government debt offer a clear view of the differences between the attention trap and heterogeneous β models, as well as being an important part of the way the US government responded to both the Great Recession and the Covid-19 pandemic[23]. Such tax cuts are expansionary in both models, but if they are sufficiently large that poor households begin to save the aggregate consumption response is smaller on impact and more persistent in the attention trap model.

3.2.1 Analytic results on MPCs

The differences between the models are driven by the MPC of poor households. I therefore proceed by showing that this MPC is lower in the attention trap model than the hetero-

[23] A lump sum tax cut in the model is a close approximation to the mailing of checks carried out in both circumstances, though the real policies did taper the transfer at very high levels of income. While the model exercise is an expansionary fiscal policy shock at the steady state, the real fiscal expansions were intended to mitigate large recessions. I abstract from the differential effects of fiscal policy across the business cycle for simplicity, but note that the mechanism differentiating the attention trap and heterogeneous β models is still potentially important in a deep recession. The stimulus packages mentioned were not targeted exclusively at those at borrowing constraints before the recessions began: some of those receiving the stimulus money had only just been pushed to their constraints by the recession, so it is plausible that the stimulus relaxed their constraints.

geneous β model after large tax cuts. Specifically, I find the partial equilibrium response of poor household consumption to an unanticipated tax cut in period s, holding all aggregate variables constant[24]. Since the two models have the same steady state, I assume that aggregate variables are the same in the two models for this exercise. Household-level variables will differ across models, and the subscript $x \in \{A, B\}$ indicates the attention trap and heterogeneous β models respectively.

Let $\mathcal{B}^p_{x,s}$ be the amount poor households spend on bonds in period s, so $\mathcal{B}^p_{A,s} = \frac{b^p_{A,s}}{1+\tilde{i}^p_s}$ and $\mathcal{B}^p_{B,s} = \frac{b^p_{B,s}}{1+i_s}$. Differentiating the poor household period s budget constraint with respect to taxes gives:

$$\frac{dc^p_{x,s}}{dT^p_s} = -\frac{1}{1 + \frac{\sigma}{\varphi} w_s^{\frac{1+\varphi}{\varphi}} c^{p\ -\frac{\sigma+\varphi}{\varphi}}_{x,s}} \left(1 + \frac{d\mathcal{B}^p_{x,s}}{dT^p_s}\right) \tag{24}$$

First, consider the case of a tax cut that is too small to lift poor households off the borrowing constraint, so $\frac{d\mathcal{B}^p_{x,s}}{dT^p_s} = 0$ in both models. The derivative $\frac{dc^p_{x,s}}{dT^p_s}$ is negative, which implies that the MPC of poor households is positive (a tax cut is a fall in T^p_s). It is also the same across both models.

Small tax cuts are therefore expansionary in these models. If all households were on their Euler equation, the MPC out of tax cuts would be zero as Ricardian equivalence would hold, but borrowing constraints mean that poor households are not Ricardian. Their consumption rises substantially, and so therefore does aggregate consumption. Two general equilibrium effects amplify the expansion. First, higher consumption means that labour demand rises, and so therefore do wages, further increasing the income and consumption of poor households. Second, in future periods taxes will be high to pay off the increased government debt. Since poor households have a high MPC, their consumption will be below steady state in these periods, which reduces inflation through the Phillips curve. The monetary authority will respond to this by reducing nominal interest rates such that real interest rates are low in those future periods (they obey the Taylor principle). Anticipating this, when wealthy households observe the tax cut in period s they expect lower future real interest rates, and so they also consume more in period s (see Bilbiie et al. (2013) for a detailed discussion of these channels).

Now turn to the case where the tax cut does lift poor households off their borrowing constraint in period s, so $\frac{d\mathcal{B}^p_{x,s}}{dT^p_s} < 0$. In both models, this reduces the MPC out of the tax cut because poor households are no longer consuming all of the extra income from the

[24] Strictly, when I consider large tax cuts below the response of consumption is no longer the marginal propensity to consume, as I am considering large changes not the gradient of consumption. I maintain the MPC language, however, as it shows the link from this discussion to the wider literature on MPC heterogeneity.

tax cut[25]. The models are not, however, identical here as poor households will increase saving by different amounts.

To see this, consider the simple example where poor households return to the borrowing constraint one period after the tax cut, so $b^p_{x,s+1} = 0$. The Euler equations between periods s and $s+1$ in the two models are:

$$c^{p^{-\sigma}}_{A,s} = \frac{\beta(1+\tilde{i}^p_s)}{\Pi_{s+1}} c^{p^{-\sigma}}_{A,s+1}, \quad c^{p^{-\sigma}}_{B,s} = \frac{\beta^p(1+i_s)}{\Pi_{s+1}} c^{p^{-\sigma}}_{B,s+1} \tag{25}$$

Recall that to compare the models, β^p and the parameters of the information problem are set so that the steady state shadow value of relaxing the borrowing constraint (λ^p) is the same in the two models. At steady state, this implies that $\beta(1+\bar{\tilde{i}}^p) = \beta^p(1+\bar{i})$. All aggregate variables are the same in the two models and the tax cut hits at the steady state, so if poor households in the attention trap model did not change their attention we would also have $\beta(1+\tilde{i}^p_s) = \beta^p(1+i_s)$.

This will not, however, be the case. We know from equation 9 that when saving rises so does attention, which implies that $\beta(1+\tilde{i}^p_s) > \beta^p(1+i_s)$. This is sufficient to ensure that the consumption response of poor households is lower in the attention trap model:

Proposition 1 *If the path of taxes is such that $b_{x,s} > 0$ and $b_{x,s+1} = 0$ for $x \in \{A, B\}$, then $c_{A,s} < c_{B,s}$.*

Proof. The proof is by contradiction. Suppose that $c^p_{A,s} \geq c^p_{B,s}$. From the period s budget constraints, this implies that $\mathcal{B}^p_{A,s} \leq \mathcal{B}^p_{B,s}$. Since $\tilde{i}^p_s < i_s$ by definition, this implies that $b^p_{A,s} < b^p_{B,s}$.

In period $s+1$, poor households therefore have lower asset income in the attention trap model, and so $c^p_{A,s+1} < c^p_{B,s+1}$. Using this and the fact that $\beta(1+\tilde{i}^p_s) > \beta^p(1+i_s)$ in the Euler equations (25) implies that $c^p_{A,s} < c^p_{B,s}$, contradicting the initial assumption. The only solution therefore has $c^p_{A,s} < c^p_{B,s}$. ■

The intuition is straightforward. Before the tax cut hits, poor households in the two models have the same desire to save. When the tax cut comes, they both start to save, which induces the poor household in the attention trap model to pay more attention to their choice of savings product. In doing so, the interest rate they face rises relative to the policy rate, increasing their desire to save further. They therefore save more, and consume less, in the period of the tax cut than the poor households in the heterogeneous β model. This lower MPC in the attention trap model implies that the fiscal expansion is less powerful in affecting aggregate consumption and output, partly because the direct effect

[25] Positive shocks therefore imply smaller impact MPCs than negative shocks in both of these models, consistent with the evidence in Bracha and Cooper (2014), Bunn et al. (2018), and Christelis et al. (2019). The smaller impact of fiscal expansions in the attention trap model increases this asymmetry, as the two models are identical for contractionary shocks just as for small positive shocks.

on consumption is smaller, and partly because the general equilibrium effects discussed above are also correspondingly weaker.

The tax cut will have more persistent effects in the attention trap model as long as poor households buy more bonds there (if $b^p_{A,s} > b^p_{B,s}$). In that case, poor household income is higher in period $s+1$ in the attention trap model due to the extra asset income, and so their consumption is higher than the poor households in the heterogeneous β model ($c^p_{A,s+1} > c^p_{B,s+1}$). The intertemporal-MPC of Auclert et al. (2018) is therefore higher in the attention trap model[26]. If they are consuming more in period $s+1$, this implies a lower λ_{s+1} in equation 7: poor households are closer to staying above their borrowing constraint. In this way larger tax cuts can imply that poor households remain above the borrowing constraint for longer in the attention trap model than in the heterogeneous β model. Intuitively, when poor households save more because they have increased attention, they have more asset income in the next period, which increases their desire to save in that future period. If they do maintain positive savings for multiple periods, they will also maintain positive attention in those future periods, which further supports their saving.

This is not, however, a guaranteed result. Proposition 1 only shows that consumption will be lower in the attention trap model in the period of the tax cut. This implies that the amount spent on bonds will be higher ($\mathcal{B}^p_{A,s} > \mathcal{B}^p_{B,s}$), but since bonds are more expensive for poor households in the attention trap model this is not sufficient to guarantee that the quantity of bonds bought will be higher. For that, we require that the difference in saving budgets $\mathcal{B}^p_{x,s}$ is large relative to the difference in bond prices. Using the definitions of $\mathcal{B}^p_{x,s}$ and $\tilde{i}^p_s$ it can be shown that $b^p_{A,s} > b^p_{B,s}$ if:

$$\mathcal{B}^p_{A,s} > \mathcal{B}^p_{B,s}(1 + \tau(1 - q^p_s)(1 + i_s)) \tag{26}$$

The probability an individual in a poor household chooses a good savings product, q^p_s, is increasing in the amount saved through the first order condition on attention (equation 9). Condition 26 is therefore more easily satisfied for larger tax cuts, which imply larger poor household saving. Even for very large interest rate dispersion (large τ), a sufficiently large tax cut will ensure that $b^p_{A,s} > b^p_{B,s}$, as q^p_s approaches 1 as $b^p_{A,s}$ gets large.

3.2.2 Simulation

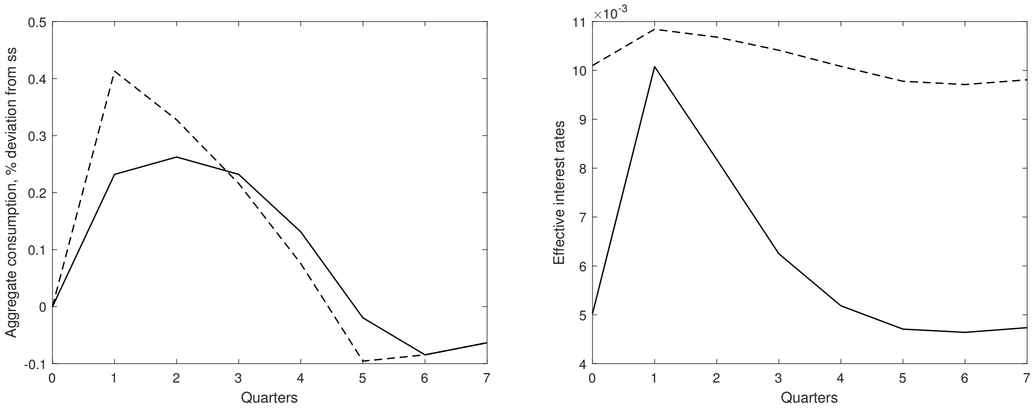
The MPC of poor households in the attention trap model is therefore smaller, and for large enough shocks the intertemporal-MPC is larger, than in the heterogeneous β model. Here I simulate a debt-funded aggregate tax cut in the two models, to show the combination

[26] More generally, this shows that TANK models where poor households leave the borrowing constraint after large positive shocks can predict the positive intertemporal-MPCs found empirically by Auclert et al. (2018), but only for positive income shocks.

of these direct effects and the general equilibrium effects discussed above, in the more general case where poor households continue to save for several periods.

The results are in figure 1. The shock is chosen to be approximately one third of the size of the 2009 fiscal stimulus measures in the US[27], and occurs in quarter 1. Government debt returns to steady state following an AR(1) process with persistence 0.75. Details of the calibration are given in appendix A.3.

Figure 1: Effects of a debt-funded aggregate tax cut



(a) Aggregate consumption in the attention trap model (solid) and the heterogeneous β model (dashed).

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

As shown analytically above, the response of aggregate consumption (figure 1a) is substantially lower on impact in the attention trap model (solid line) than in the heterogeneous β model (dashed line). With this calibration the rise in aggregate consumption on impact is 44% lower in the attention trap model. Aggregate consumption in the attention trap model is also more persistent, which means that the cumulative aggregate consumption deviation over quarters 1-7 (when the models coincide once more) is only 13% lower in the attention trap model. Poor household saving returns to 0 a quarter later in the attention trap model.

The mechanism can be seen in figure 1b, which shows the effective interest rates earned by wealthy (dashed line) and poor (solid line) households in the attention trap model. As poor households lift off their borrowing constraint they increase attention to choosing between savings products. That causes their effective interest rate to rise substantially relative to the rate experienced by the wealthy households, which is the policy rate. The gap between the effective interest rates of the two households widens as debt is brought back to steady state and poor households reduce their saving back to the

[27]The \$787 billion fiscal stimulus in the American Recovery and Reinvestment Act (Dupor, 2017) amounted to 21.6% of the previous quarter's GDP (in real terms), though this did not all enter the economy in one quarter. I set the tax shock such that government debt-to-GDP rises by 7.2% in the period that the shock hits.

borrowing constraint, causing them to reduce their attention. This would happen even without further tax changes through the logic of the attention trap, but it is accelerated here by the higher taxes used to repay the government debt, which reduce the income, and so saving, of all households.

While the tax cut causes interest rate heterogeneity to fall in the attention trap model[28], it is not eliminated entirely. The first order condition on attention (equation 9) implies that the two households only face the same effective interest rate if they save the same amount (in marginal utility units), so no aggregate tax cut of any size will reduce interest rate heterogeneity to zero. In fact, aggregate tax cuts increase the heterogeneity in saving in marginal utility units, because wealthy households absorb the bulk of the extra government debt, and they do so by reducing consumption.

Redistributive taxation, in contrast, can permanently remove interest rate and wealth heterogeneity, if the transfers are such that all households have the same income in some period. This, however, is a knife-edge result, as even if the incomes of the households differ by a very small amount the attention trap will kick in, taking the economy back to the two-agent steady state. This instability of the representative-agent steady state is discussed in appendix D.

3.3 Factors Affecting the Power of Fiscal Policy

As well as implying that fiscal expansions have smaller and more persistent effects, the attention trap model differs from the heterogeneous β model in that a number of features of the economic environment affect the power of fiscal policy through novel channels not present in that comparison model. In general, anything that affects the incentives of poor households to pay attention to their choice between savings products will affect their consumption response to tax cuts, and so the aggregate impact of such a policy shock. I analyse three such factors: the cost of information, asset income taxation, and interest rate dispersion.

3.3.1 The cost of information

The cost of information μ only enters the equilibrium conditions of the model in the attention first order condition (equation 9). Using the definitions of i^b (equation 1) and

[28] This is not specific to the case studied here where $q^w \approx 1$ in all periods. Diminishing marginal returns to attention imply convex costs of higher effective interest rates ($\mathcal{I}''(\tilde{i}) > 0$), and so wealthy households increase their effective rates by less after an aggregate expansionary shock than poor households, because the marginal cost of doing so is more sharply increasing for wealthy households.

$\mathcal{I}$ (equation 3), this can be written as:

$$b_t^h c_t^{h^{-\sigma}} = \frac{\mu}{\tau} \log\left(\frac{q_t^h(1-g)}{(1-q_t^h)g}\right) \tag{27}$$

Taking the same approach as in section 3.2.1, consider the response of the probability of choosing a good savings product to a change in taxes in period s, holding all aggregate variables constant:

$$\frac{dq_s^h}{dT_s^h} = \frac{c_s^{h^{-\sigma}} q_s^h(1-q_s^h)\tau}{\mu}\left(\frac{db_s^h}{dT_s^h} - \frac{\sigma b_s^h}{c_s^h}\frac{dc_s^h}{dT_s^h}\right) \tag{28}$$

For wealthy households $q_s^w \approx 1$, so $\frac{dq_s^w}{dT_s^w} \approx 0$. For poor households, however, equation 28 shows that the elasticity of q_s^p to tax cuts is higher when the cost of attention μ is low.

Intuitively, poor households pay no attention in steady state if the cost is positive, as they derive no benefit from information while they are not saving. When they start to save, they increase attention until the (now positive) marginal benefits of doing so equal the marginal costs. With a lower marginal cost of information, poor households therefore increase their attention by more after a tax cut. Their effective interest rate rises by more, and so they save more, which further reduces their MPC out of the tax cut and increases the intertemporal-MPC, making the consumption response to the tax cut even smaller and more persistent[29]. These effects are amplified by the fact that when poor households save more and consume less in the period of the shocks, that appears in equation 28 as a more negative $\frac{db_s^p}{dT_s^p}$ and a less negative $\frac{dc_s^p}{dT_s^p}$: the extra saving adds to the incentives to increase attention.

3.3.2 Asset income taxes

Asset income taxation can be incorporated into the attention trap model by rewriting the budget constraint as[30]:

$$c_t^h + \frac{b_t^h}{1+\tilde{i}_t^h} = \frac{b_{t-1}^h}{\Pi_t}(1-T_b) + w_t n_t^h + \mathcal{D}_t - T_t^h \tag{29}$$

[29]This, of course, is only one effect of lower information costs. Lower μ also implies that when poor households do save, less output is wasted in the transaction costs of bad bond dealers as poor households choose them with a lower probability. I study the implications of policies that lower information costs in a more general setting with profit maximising banks and cyclical variation in the attention of all households in Macaulay (2020).

[30]The revenue from the asset income tax also enters the government budget constraint. This means that a debt-funded tax cut is cheaper for the government than without the asset tax, because with more asset supply they increase the revenue from this tax. This does affect the impacts of a debt-funded lump sum tax cut, but here I focus on the channels that are novel to the attention trap model.

The asset income tax T_b does not directly alter the attention first order condition, but the Euler equation becomes:

$$c_t^{h^{-\sigma}} = \beta(1+\tilde{i}_t^h)(1-T_b)\mathbb{E}_t\left(\frac{c_{t+1}^{h^{-\sigma}}}{\Pi_{t+1}}\right) + (1+\tilde{i}_t^h)\lambda_t^h \tag{30}$$

Any rise in the effective interest rate $\tilde{i}_t^h$ therefore has a smaller effect on the intertemporal path of consumption when asset income taxes are higher. This means that after a debt-funded lump sum tax cut, a given rise in poor household attention will lead to less of a rise in poor household saving. With less saving, poor households increase their attention by less, so the asset tax indirectly affects the attention first order condition. Through both of these channels, asset income taxation therefore implies that lump sum debt-funded tax cuts will have larger and less persistent effects.

This suggests that cuts to asset income taxes, as occurred for example in the UK in 2016, could reduce the power of fiscal policy to stimulate consumption. This analysis, however, abstracts from the effect of asset income taxation on wealthy households, in part because I assume $q^w \approx 1$, and also because the tax is levied on the payoff from bonds while attention acts to lower their price. An alternative setup would have all bonds sold at the same price, with attention increasing the interest rate paid out in the next period. In that case, lower taxes on interest income directly increase the returns to attention, which leads to an additional effect of lower T_b. If wealthy households are not already paying so much attention that $q^w \approx 1$, then lower T_b leads them to pay more attention in every period, which reduces the output wasted on transaction costs at bad bond dealers[31]. In this way reducing asset income taxes may be beneficial for welfare, though that effect is absent in the model used here.

3.3.3 Interest rate dispersion

Interest rate dispersion is controlled in the model by the parameter τ. The direct effects of dispersion on attention can be seen in equations 27 and 28: lower dispersion (lower τ) implies a lower level of attention and a smaller response of attention to taxes. This is because lower interest rate dispersion reduces the interest rate gain from choosing a good savings product rather than a bad, and so reduces the utility benefit of attention. Just as lower marginal costs of attention increase the level and response of attention to tax cuts for poor households, the lower marginal benefit of attention brought about by lower dispersion reduces both q_t^p and $\frac{dq_s^p}{dT_s^p}$, if everything else is held constant.

[31] The model in Macaulay (2020) uses such a setup. The qualitative results in this paper are the same if these alternative assumptions are used, but since the proofs are cleaner when attention affects the current bond price I use that here.

This is not, however, the only effect of a fall in interest rate dispersion. Recall that from equation 20, the steady state shadow value of relaxing the borrowing constraint for poor households (λ^p) falls with τ. The savings poor households would have chosen without that constraint rise closer to zero.

To see why, notice that since wealthy households pay full attention ($q^w \approx 1$) the steady state policy rate, determined by their Euler equation, is unaffected by τ: $1+\bar{i} = \frac{1}{\beta}$. Lower dispersion must translate in steady state to a higher interest rate on the bad products $\bar{i}^b$. Lower dispersion therefore leads to a higher effective interest rate for any level of attention that implies bad products are chosen with positive probability. Poor households face higher effective interest rates in steady state (when $\mathcal{I}^p = 0$ so $\bar{q}^p = g$), and this increases their incentives to save.

As discussed in section 2.5, this does not lead to poor households saving in steady state unless τ reaches 0. An intermediate fall in τ does however imply that poor households begin to save after smaller tax cuts. Again consider the partial equilibrium effects of a tax cut in period s: substituting the budget constraint and the definition of $\tilde{i}^p_s$ into the Euler equation of a poor household beginning to save gives:

$$T^p_s + \mathcal{B}^p_s = w_s n^p_s + \mathcal{D}_s - \left(\frac{\Pi_{s+1}}{\beta(1+i_s)} + \frac{\tau(1-q^p_s)\Pi_{s+1}}{\beta}\right)c^p_{s+1} \qquad (31)$$

To find the maximum level of taxes such that poor households will save a positive amount, take this and set $\mathcal{B}^p_s = 0$, which implies $q^p_s = g$. A fall in τ causes this maximum tax to rise, implying that the tax cut required to lift poor households from their borrowing constraint is smaller. If a tax cut is just small enough to leave households at the borrowing constraint with some high level of τ, a fall in dispersion will therefore reduce the power of such a tax cut as the poor household will start to save.

For a given tax cut that is large enough to cause poor households to start saving even with a high τ, however, the effects of lower dispersion are more complicated. In equation 31, a lower τ directly increases the amount saved for a given T^p_s. However, lower τ also implies less attention for any given level of saving (equation 27), reducing q^p_s, which in equation 31 acts to reduce saving.

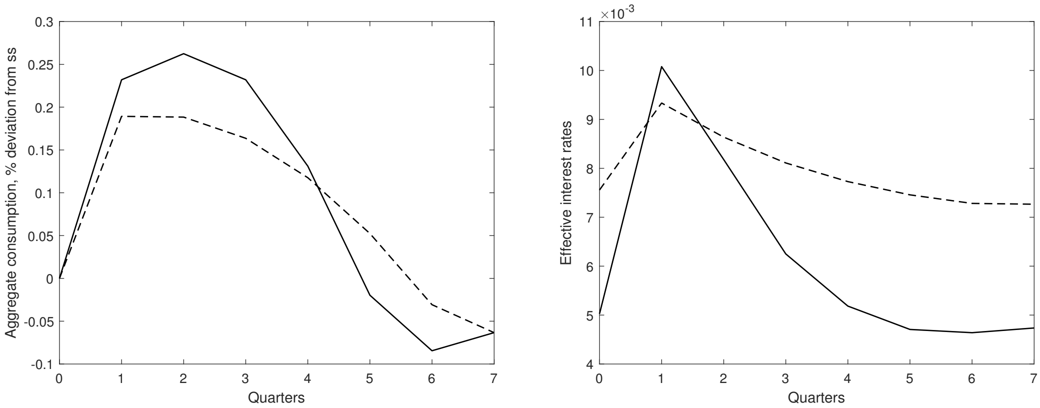
With the calibration used in section 3.2.2, reducing τ increases $\tau(1-q^p_s)$. The effective interest rate faced by poor households in the period of the tax cut, $\tilde{i}^p_s$, is therefore lower when τ is lower, even though for any given level of attention the effective interest rate would be higher. However, this does not lead to a greater MPC when interest rate dispersion is lower, because period s consumption is determined by the discounted sum of future effective interest rates until the poor household hits the borrowing constraint again, not just the effective interest rate in the period of the tax cut.

As taxes rise to pay off the government debt issued to fund the tax cut, poor households reduce their saving, and so reduce their attention to choosing between savings products. This fall in attention has a smaller effect on the effective interest rate when interest rate dispersion is low, and so in some period after the tax cut the effective interest rate in the low-dispersion environment overtakes that in a high-dispersion setting. The discounted sum of effective interest rates can therefore be higher when dispersion is low even if the initial interest rate is lower, reducing the MPC of poor households and the impact of the tax cut.

Moreover, low interest rate dispersion implies that poor households remain above the borrowing constraint for longer, because the minimum income required for them to save is lower (equation 31). Positive saving in some future period reduces expected consumption in that period, which also causes poor households to reduce initial consumption and increase saving in order to smooth consumption.

These effects are shown in figure 2, which repeats the baseline simulation of figure 1 (solid line) and compares it to a simulation with a 50% lower τ (dashed line). With this calibration, the effective interest rate faced by poor households is higher in the low-τ case in all periods except period 1, when the tax cut hits. Poor households therefore increase their consumption by less when they receive the tax cut in the low-τ case, so the tax cut has a 23% smaller effect on aggregate consumption on impact than in the baseline with higher interest rate dispersion[32].

Figure 2: Effects of a debt-funded aggregate tax cut, high vs low τ



(a) Aggregate consumption in the attention trap model with high (solid) and low (dashed) τ.

(b) Effective interest rates facing poor households with high (solid) and low (dashed) τ.

These findings are of particular relevance in the current environment of persistently

[32]The qualitative result that lower τ reduces the impact of fiscal policy is robust to various different calibrations. In particular, as long as government debt decays with a persistence $\rho > 0.05$, b^p remains above 0 for at least one more quarter in the low-τ environment, so consumption responds less to the tax cut when τ is low.

low natural rates of interest (Holston et al., 2017), since low nominal interest rates are likely to coincide with low interest rate dispersion, as banks are unwilling to pass on negative interest rates to households (Eggertsson et al., 2019)[33]. Low interest rate dispersion may therefore reduce the power of fiscal policy precisely when it is needed to play a greater role in stabilising demand in recessions, that is when monetary policy is constrained by the effective lower bound on policy rates.

4 Evidence for Attention Rising After a Tax Cut

Fiscal policy has smaller and more persistent effects in the attention trap model because poor households increase their attention to choosing between savings products when their income rises after a tax cut. In this section I provide evidence for such an attention response, using cross-state variation to identify the effect of the Tax Cuts and Jobs Act (TCJA) in the US on a survey-based measure of uncertainty about savings interest rates. The TCJA was signed on 22nd December 2017 and came into effect on 1st January 2018, and the largest single tax change was the cut in the federal corporate income tax rate from 35% to 21%. I find that this significantly reduced uncertainty about savings among poorer households, consistent with the model prediction that these households increased their attention.

4.1 Data and Identification

The key data source is the Survey of Consumer Expectations (SCE). This large panel survey asks respondents for the probability that the interest rate on savings will rise in the next 12 months. To construct a measure of uncertainty about savings, I note that since "interest rates will rise" is a binary statement, the variance in individual i's belief about the statement is given by $p_i(1-p_i)$, where p_i is their subjective probability of the statement being true, and their survey response. I use this variance as a measure of how uncertain individual i is about the savings market. Intuitively, the further an individual's response is from 50/50 the less uncertain they are about the path of future savings rates.

In the model, tax cuts cause poor households to pay more attention to their choice of savings product, reducing their uncertainty about which product to choose for their saving in that period. This is not the same as the uncertainty measured using the SCE, which is uncertainty about the path of future savings interest rates rather than current interest

[33]While Eggertsson et al. (2019) study interest rates in Sweden, this reduced rate dispersion is not confined to economies with negative policy rates. For the UK, Barrett (2018) notes in the Financial Times that "as interest rates continue to bump along the bottom, the difference between the best and the worst easy-access savings rates are not that great".

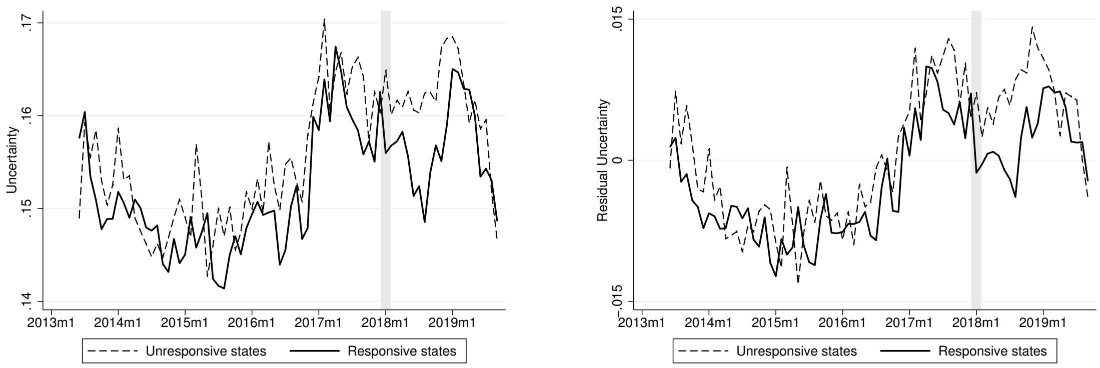
rates. I assume that the two forms of uncertainty are correlated. Intuitively, while doing research on where to save today, households also learn more about the savings market overall, and so become less uncertain about its future[34].

To identify the effect of the TCJA on this measure of uncertainty, I make use of the results on the state-level effects of corporate tax cuts in Liu and Williams (2019) (LW). They find that corporate tax cuts only lead to a significant increase in GDP in 19 of the states of the US, while most states see no significant effect. This heterogeneity is explained by differences in the industry make-up of different states. In particular, less capital intensive states are more likely to expand after a federal corporate tax cut.

I split states following LW's results. I refer to the states that see a significant expansion in GDP after corporate tax cuts as 'responsive'[35], and the other states as 'unresponsive'. Since it is income that determines attention in the model, the model predicts that uncertainty will fall by more in states where the tax cut causes an expansion in economic activity, that is in the responsive states. Industry and capital intensity are slow-moving variables, so should not change substantially in the months immediately before and after the tax cut. The LW estimation uses data from 1950-2017, stopping before the TCJA.

Figure 3 plots the average uncertainty in the responsive and unresponsive states over time, with and without controls[36]. In both graphs, there is no clear difference in the trends in uncertainty between the two groups before the tax cut. After the tax cut, uncertainty falls temporarily in the responsive states relative to the unresponsive states. The difference closes up after just over a year.

Figure 3: Uncertainty in unresponsive and responsive states. The tax cut occurs in 2018m1.



(a) Raw series

(b) With controls

[34]More specifically, the model only features saving in one-period bonds, but if longer term bonds are also available then to compare between asset maturities savers will need to collect information on the likely path of future short term savings interest rates. Greater attention to the current asset choice will therefore reduce uncertainty about future rates.

[35]These are CO, FL, GA, HI, ID, IL, MA, MD, ME, MI, MO, NE, NH, NJ, NV, RI, SC, WA, WV.

[36]Specifically, the series in figure 3b is the residual uncertainty after controlling for month-of-the-year, state and individual fixed effects, income and financial fragility, where financial fragility is defined as the self-reported probability of the household being unable to pay their bills in the next 3 months.

In section 4.2 I formally test for this fall in uncertainty in responsive states. I find that it is significant, and is concentrated among households with low incomes but who are not in immediate danger of being unable to pay their bills. This is consistent with the model, in which a tax cut causes a temporary rise in attention from poor households who increase their saving a little after the cut. I also test, and fail to reject, the assumption of parallel trends in uncertainty before the tax cut.

4.2 Testing for the effect of the TCJA on Uncertainty

Figure 3 suggests that uncertainty about interest on savings fell in responsive states relative to unresponsive states after the tax cut. To test this, I run the following regression:

$$unc_{ist} = D_{st} \sum_{c=1}^{4} \beta_c \mathbb{1}_{C_{ist}=c} + \gamma_i + \delta_s + \zeta_t + \theta X_{ist} + \varepsilon_{ist} \tag{32}$$

The dependent variable unc_{ist} is the uncertainty in the survey response of individual i in state s at time t. D_{st} is the treatment dummy: it equals 1 if the individual lives in a state that is responsive to corporate tax cuts and the date of the survey is on or after the date of the tax cut coming into force (January 2018)[37].

I allow the tax cut to have different effects depending on whether the individual has high or low income (above or below the median in the sample), and if they are financially fragile, which I define as reporting a greater than 5% chance of being unable to pay their bills in the next 3 months. C_{ist} is a categorical variable for the combination of these two characteristics. The model predicts that the households who change their attention after the tax cut are those who are close to their borrowing constraint, so if the model is a good representation of attention decisions in this market uncertainty should fall only for those with low income who are not financially fragile.

The remaining terms in equation 32 are individual, state, and time fixed effects, and time-varying individual-level controls (income and financial fragility). Individual fixed effects account for any demographic variables that do not change over the time the individual remains in the survey, though there is little demographic difference between responsive and unresponsive states (see appendix E.1). I estimate equation 32 both over the whole survey with heteroskedasticity-robust standard errors, and over the slightly smaller set of survey respondents who do not move state while in the survey, for which I

[37] I do not include leads and lags of the treatment variable in the regression because individuals only remain in the survey for an average of 7 months, so leads and lags would substantially reduce the number of observations available. I therefore leave the testing of the parallel trend assumption to another specification which does not make as much use of the panel dimension of the survey (see figure 4).

can cluster standard errors at the state level[38]. The results are displayed in table 1.

Table 1: Regression on the uncertainty in forecasts of the interest rate on savings.

	(1) Uncertainty	(2) Uncertainty
Low income non-fragile	-0.00874**	-0.00859**
	(-2.45)	(-2.27)
High income non-fragile	-0.000277	0.000690
	(-0.07)	(0.22)
Low income fragile	0.00151	0.00212
	(0.43)	(0.63)
High income fragile	-0.00151	-0.000511
	(-0.37)	(-0.13)
Individual FE	Yes	Yes
State FE	Yes	No
Time FE	Yes	Yes
Clustered SE	No	Yes
Observations	99748	97702

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

As predicted by the model, low income non-fragile households become less uncertain after the tax cut in responsive states relative to unresponsive states, and there is no effect on the uncertainty of high income or fragile households. The fall in the uncertainty of low income non-fragile households is non-negligible: it is 6.4% of the mean uncertainty for these households across the survey, or 9.8% of the standard deviation.

To test for parallel trends, I first regress savings interest rate uncertainty on individual fixed effects and the time-varying individual characteristics X_{ist}. I then take the average of the residual uncertainty by state, high/low income, and financial fragility. This gives four continuous series of average uncertainty for each state, one for each combination of income and financial fragility. Using these average uncertainty series, I run:

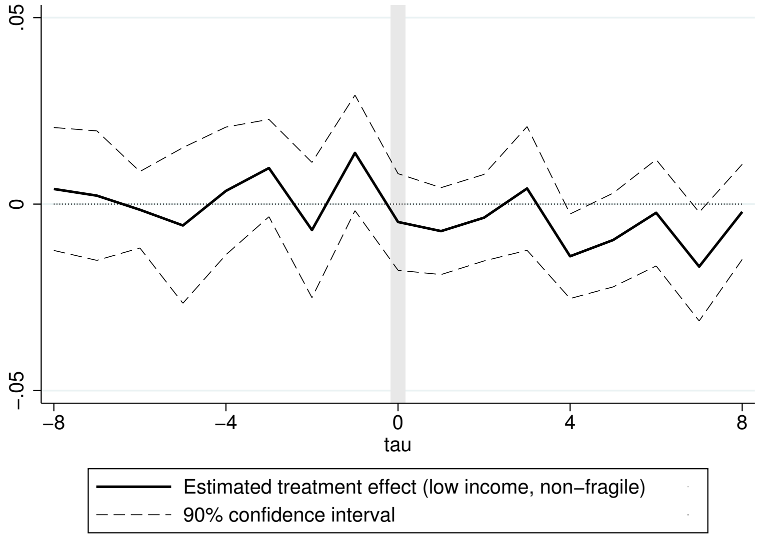
$$u\bar{n}c_{cst} = \sum_{\tau=-8}^{8} \sum_{c=1}^{4} \tilde{D}_{st+\tau} \beta_{c\tau} \mathbb{1}_{C_{cst+\tau}=c} + \gamma_c + \delta_s + \zeta_t + \varepsilon_{cst} \tag{33}$$

Here $u\bar{n}c_{cst}$ is the average residual uncertainty for individuals with income and fragility characteristics c in state s at time t. $\tilde{D}_{st}$ equals 1 in states that are responsive to corporate tax cuts on the date of the cut, and equals zero for all states before and after[39].

[38]Without movers individual fixed effects ensure all state fixed effects are zero in this specification.
[39]This differs from the treatment dummy in equation 32, D_{st}, which is 1 for all periods on or after the

The index τ therefore denotes leads and lags of the treatment. C_{cst} denotes the combination of income and financial fragility characteristics as before. The remaining terms are characteristic, state and time fixed effects. The results are plotted[40] in figure 4.

Figure 4: Estimated coefficients $\beta_{t+\tau}$ in equation 33.



We do not reject the parallel trend assumption. The effect of leads of the tax cut is insignificantly different from zero. After the tax cut, the estimated effects of the policy on uncertainty become negative, as in the individual-level model in equation 32. The magnitudes are similar to the individual-level model, with the effects on low income non-fragile individuals 4 and 7 months after the tax cut estimated at -0.014 and -0.017 respectively. These lags of the policy are significant at the 10% level (p-values 0.048 and 0.064), despite the effects of the tax cut being estimated much less precisely in this specification than in equation 32, since aggregating to the state-characteristic level reduces the number of observations by more than a factor of 10.

Taken together, this evidence suggests that the tax cut was associated with lower uncertainty about savings markets for individuals with low income but who are not financially fragile, in states that are responsive to corporate tax cuts. This is consistent with the model, and provides evidence for its key mechanism. Further tests in appendix E.2 also support the model: the tax cut is associated with an increase in income for low income non-fragile households, but is not associated with a change in the subjective probability of an interest rate rise. The results here could therefore be driven by income

date of the tax cut. This change aids the interpretation of the coefficients on the leads and lags of $\tilde{D}_{st}$.

[40]For brevity, figure 4 presents only the coefficients for the treatment on the group of interest (low income non-fragile). The regression is weighted using the number of observations in the average used to construct each *cst* cell, and standard errors are clustered at the state level.

changes, but are not driven by changes in the direction of expected interest rate moves - only the uncertainty around that expectation.

5 Conclusion

The existence of wealth inequality alters many of the predictions of macroeconomic models, with fiscal policy a particularly strong example. I have shown that the source of the inequality also affects those predictions. In a model where wealth inequality is driven by endogenous asset return heterogeneity, fiscal expansions have smaller and more persistent effects on consumption than in a comparable two-agent model in which inequality comes from heterogeneous discount factors. Lower costs of information about asset returns, taxes on asset income, and interest rate dispersion all reduce the impact of fiscal expansions further in this model.

These results are due to a complementarity between saving and information processing. As poor households increase their saving after a fiscal expansion, they have greater incentives to pay attention to their choice of savings product. They make better savings choices, and the interest rate they face rises relative to the policy rate, causing them to save more and consume less. Cross-state variation in uncertainty about savings interest rates after the 2017 Tax Cuts and Jobs Act provides evidence for this response of attention to tax cuts.

References

Acharya, S. and Dogra, K. (2020). Understanding HANK: Insights From a PRANK. *Econometrica*, 88(3):1113–1158.

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

Adam, K. (2009). Monetary policy and aggregate volatility. *Journal of Monetary Economics*, 56(SUPPL.):S1–S18.

Afrouzi, H. (2020). Strategic Inattention, Inflation Dynamics, and the Non-Neutrality of Money.

Almgren, M., Gallegos Dago, J. E., Kramer, J., and Lima, R. (2019). Monetary Policy and Liquidity Constraints: Evidence from the Euro Area.

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

Auclert, A., Rognlie, M., and Straub, L. (2018). The Intertemporal Keynesian Cross. Technical report, National Bureau of Economic Research, Cambridge, MA.

Barrett, C. (2018). What would it take for you to switch your bank account?

Bayer, C., Born, B., and Luetticke, R. (2020). The Liquidity Channel of Fiscal Policy.

Benhabib, J., Bisin, A., and Luo, M. (2019). Wealth distribution and social mobility in the US: A quantitative approach. *American Economic Review*, 109(5):1623–1647.

Best, J. and Dogra, K. (2019). Capital Management and Inequality.

Bilbiie, F. O. (2008). Limited asset markets participation, monetary policy and (inverted) aggregate demand logic. *Journal of Economic Theory*, 140(1):162–196.

Bilbiie, F. O. (2019a). Monetary Policy and Heterogeneity: An Analytical Framework.

Bilbiie, F. O. (2019b). The New Keynesian cross. *Journal of Monetary Economics*.

Bilbiie, F. O., Monacelli, T., and Perotti, R. (2013). Public Debt and Redistribution with Borrowing Constraints. *The Economic Journal*, 123(566):F64–F98.

Blanchard, O. and Summers, L. H. (2019). Introduction. In Blanchard, O. and Summers, L. H., editors, *Evolution or Revolution? Rethinking Macroeconomic Policy after the Great Recession*. MIT Press.

Bracha, A. and Cooper, D. (2014). Asymmetric responses to income changes: The payroll tax increase versus tax refund in 2013. *Economics Letters*, 124(3):534–538.

Briggs, J., Cesarini, D., Lindqvist, E., and Östling, R. (2020). Windfall Gains and Stock Market Participation. *Journal of Financial Economics*.

Bunn, P., Le Roux, J., Reinold, K., and Surico, P. (2018). The consumption response to positive and negative income shocks. *Journal of Monetary Economics*, 96:1–15.

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

Cantore, C. and Freund, L. B. (2020). Workers, Capitalists, and the Government: Fiscal Policy and Income (Re)Distribution.

Christelis, D., Georgarakos, D., Jappelli, T., Pistaferri, L., and Van Rooij, M. (2019). Asymmetric consumption effects of transitory income shocks. *Economic Journal*, 129(622):2322–2341.

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

Debortoli, D. and Galí, J. (2018). Monetary policy with heterogeneous agents: Insights from TANK models. *Economics Working Papers 1686, Department of Economics and Business, Universitat Pompeu Fabra.*

Dupor, W. (2017). How the U.S. Debt-to-GDP Ratio Has Changed. *Federal Reserve Bank of St. Louis, On the Economy Blog.*

Eggertsson, G. B., Juelsrud, R. E., Summers, L. H., and Wold, E. G. (2019). Negative nominal interest rates and the bank lending channel.

Fagereng, A., Guiso, L., Malacrino, D., and Pistaferri, L. (2020). Heterogeneity and Persistence in Returns to Wealth. *Econometrica*, 88(1):115–170.

Galí, J., López-Salido, J. D., and Vallés, J. (2007). Understanding the effects of government spending on consumption. *Journal of the European Economic Association*, 5(1):227–270.

Giambattista, E. and Pennings, S. (2017). When is the government transfer multiplier large? *European Economic Review*, 100:525–543.

Gornemann, N., Kuester, K., and Nakajima, M. (2016). Doves for the Rich, Hawks for the Poor? Distributional Consequences of Monetary Policy. *International Finance Discussion Paper*, 2016(1167):1–40.

Hagedorn, M., Manovskii, I., and Mitman, K. (2019). The Fiscal Multiplier. *NBER Working Papers.*

Holm, M. B., Paul, P., and Tischbirek, A. (2020). The Transmission of Monetary Policy under the Microscope. *Federal Reserve Bank of San Francisco, Working Paper Series*, pages 01–87.

Holston, K., Laubach, T., and Williams, J. C. (2017). Measuring the natural rate of interest: International trends and determinants. *Journal of International Economics*, 108:S59–S75.

Joint Committee on Taxation (2017). Macroeconomic Analysis of the Conference Agreement for H.R. 1, the "Tax Cuts and Jobs Act" (JCX-69-17).

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

Kaplan, G., Moll, B., and Violante, G. L. (2018). Monetary policy according to HANK. *American Economic Review*, 108(3):697–743.

Kaplan, G. and Violante, G. L. (2018). Microeconomic heterogeneity and macroeconomic shocks. *Journal of Economic Perspectives*, 32(3):167–194.

Kumar, A. (2020). Did Tax Cuts and Jobs Act Create Jobs and Stimulate Growth?

Early Evidence Using State-Level Variation in Tax Changes. *Federal Reserve Bank of Dallas, Working Papers*, 2020(2001).

Lagarde, C. (2020). Opening remarks at the EUI's State of the Union event.

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

Liu, C. and Williams, N. (2019). State-level implications of federal tax policies. *Journal of Monetary Economics*, 105:74–90.

Lusardi, A., Michaud, P. C., and Mitchell, O. S. (2017). Optimal financial knowledge and wealth inequality. *Journal of Political Economy*, 125(2):431–477.

Macaulay, A. (2020). Cyclical Attention to Saving.

Maćkowiak, B. and Wiederholt, M. (2009). Optimal Sticky Prices under Rational Inattention. *American Economic Review*, 99(3):769–803.

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

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.

McKay, A., Nakamura, E., and Steinsson, J. (2016). The power of forward guidance revisited. *American Economic Review*, 106(10):3133–3158.

Occhino, F. (2019). The Macroeconomic Effects of the Tax Cuts and Jobs Act.

Powell, J. (2020). FOMC Press Conference, July 29 2020.

Rotemberg, J. J. (1982). Sticky Prices in the United States. *Journal of Political Economy*, 90(6):1187–1211.

Schmitt-Grohé, S. and Uribe, M. (2003). Closing small open economy models. *Journal of International Economics*, 61(1):163–185.

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

Sims, C. A. (2006). Rational inattention: Beyond the linear-quadratic case. *American Economic Review*, 96(2):158–163.

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

Werning, I. (2015). Incomplete Markets and Aggregate Demand. *National Bureau of Economic Research Working Paper Series.*

Zidar, O. (2019). Tax cuts for whom? Heterogeneous effects of income tax changes on growth and employment. *Journal of Political Economy*, 127(3):1437–1472.

A Model Equations

A.1 Taylor Rule and Firm Problem in the attention trap model

A perfectly competitive final goods producer combines varieties of intermediate goods $Y_t(i)$ to make the final good using a standard CES aggregator:

$$Y_t = \left(\int_0^1 Y_t(i)^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}} \tag{34}$$

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

$$Y_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\epsilon} Y_t \tag{35}$$

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

$$P_t = \left(\int_0^1 P_t(i)^{1-\epsilon} \right)^{\frac{1}{1-\epsilon}} \tag{36}$$

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

$$Y_t(i) = A_t N_t(i)^{1-\alpha} \tag{37}$$

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

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

$$\max_{P_t(i)} \mathbb{E}_0 \sum_{t=0}^{\infty} \Lambda_{0,t} \left[P_t(i) Y_t(i) - W_t N_t(i) - \frac{\psi}{2} \left(\frac{P_t(i)}{P_{t-1}(i)} - 1 \right)^2 P_t Y_t \right] \tag{38}$$

subject to

$$Y_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\epsilon} Y_t \tag{39}$$

$$Y_t(i) = A_t N_t(i)^{1-\alpha} \tag{40}$$

Firm ownership is equally distributed among households, who contract out firm operation to an agent with a stochastic discount factor of $\Lambda_{t,t+1} = \frac{\beta}{\Pi_{t+1}}\left(\frac{\bar{c}_t}{\bar{c}_{t+1}}\right)^{\sigma}$, where $\bar{c}$ is aggregate consumption[41]. All profits are distributed back to households equally.

The first order condition of an intermediate goods producer gives the Phillips curve when we substitute in that all intermediate firms are identical, so make identical choices:

$$Y_t(1-\epsilon)+\frac{\epsilon}{1-\alpha}w_t\left(\frac{Y_t}{A_t}\right)^{\frac{1}{1-\alpha}}-\psi(\Pi_t-1)\Pi_t Y_t+\psi\beta\mathbb{E}_t(\Pi_{t+1}-1)\Pi_{t+1}Y_{t+1}\left(\frac{\bar{c}_t}{\bar{c}_{t+1}}\right)^{\sigma}=0 \tag{41}$$

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

$$Y_t = A_t(\omega n_t^w + (1-\omega)n_t^p)^{1-\alpha} \tag{42}$$

Monetary policy is conducted according to a Taylor rule:

$$1 + i_t = \bar{i} + \left(\frac{\Pi_t}{\bar{\Pi}}\right)^{\delta_\Pi}\left(\frac{Y_t}{\bar{Y}}\right)^{\delta_Y} \tag{43}$$

A.2 Comparison Model Equations

The household side of the model consists of an Euler equation, intratemporal labour supply condition, and a budget constraint for each type of household, $h \in \{w, p\}$:

$$c_t^{h^{-\sigma}} = \beta^h(1+i_t)\mathbb{E}_t\left(\frac{c_{t+1}^{h^{-\sigma}}}{\Pi_{t+1}}\right) + (1+i_t)\lambda_t^h \tag{44}$$

$$c_t^{h^\sigma} n_t^{h^\varphi} = w_t \tag{45}$$

$$c_t^h + \frac{b_t^h}{1+i_t} = \frac{b_{t-1}^h}{\Pi_t} + w_t n_t^h + \mathcal{D}_t - T_t^h \tag{46}$$

The firm, policy and market clearing blocks of the model are as in section 2 and appendix A.1, except that there are no bad bond dealer transaction costs in the goods market clearing condition. The equations consist of a Phillips curve, production function, government

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

budget constraint, tax rule, and market clearing conditions for bonds and goods:

$$Y_t(1-\epsilon)+\frac{\epsilon}{1-\alpha}w_t\left(\frac{Y_t}{A_t}\right)^{\frac{1}{1-\alpha}}-\psi(\Pi_t-1)\Pi_t Y_t+\psi\beta\mathbb{E}_t(\Pi_{t+1}-1)\Pi_{t+1}Y_{t+1}\left(\frac{\bar{c}_t}{\bar{c}_{t+1}}\right)^{\sigma}=0 \quad (47)$$

$$Y_t = A_t(\omega n_t^w + (1-\omega)n_t^p)^{1-\alpha} \quad (48)$$

$$\frac{\bar{b}_{t-1}}{\Pi_t} = \frac{\bar{b}_t}{1+i_t} + (\omega T_t^w + (1-\omega)T_t^p) \quad (49)$$

$$1 + i_t = \bar{i} + \left(\frac{\Pi_t}{\bar{\Pi}}\right)^{\delta_\Pi}\left(\frac{Y_t}{\bar{Y}}\right)^{\delta_Y} \quad (50)$$

$$\bar{b}_t = \omega b_t^w + (1-\omega)b_t^p \quad (51)$$

$$T_t^w = T_t^p + z_t \quad (52)$$

$$Y_t = \omega c_t^w + (1-\omega)c_t^p + \frac{\psi}{2}(\Pi_t - 1)^2 Y_t \quad (53)$$

A.3 Calibration for section 3

The parameters common to the attention trap and heterogeneous β models are set at standard values, shown in the first block of the table below. Next is the novel parameter in the heterogeneous β model, the discount factor of poor households. The final portion of the table shows the parameter values for the information problem in the attention trap model. These are set such that λ^p, the shadow value of relaxing the borrowing constraint for the poor household, is equal in steady state across the two models. The information cost has been set sufficiently low that q^w is close to 1 in all periods of the simulations.

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

Parameter	Description	Value
β	Discount factor	0.99
σ	Coefficient of relative risk aversion	1
φ	Disutility of labour	1
ω	Proportion of households which are wealthy	0.65
ϵ	Elasticity of substitution between intermediate good varieties	9
A	Steady state technology	1
α	Degree of diminishing returns in intermediate firms' production	0.25
ψ	Price adjustment costs	372.8
$\mathcal{B}$	Steady state government debt	1
$\bar{\Pi}$	Gross inflation target	1
d_{Π}	Elasticity of the interest rate to inflation	1.5
$\bar{Y}$	Output target	0.858954
$\bar{i}$	Interest rate target	$1/\beta - 1$
d_Y	Elasticity of the interest rate to output	0.125
β^p	Poor household discount factor	0.985
μ	Information cost	0.0001
g	Fraction of good financial institutions	0.15
τ	Inefficient bond dealer costs	0.005882
τ^-	Lower inefficient bond dealer costs for section 3.3.3	0.002941

B Proofs

B.1 Derivation of equation 3

Suppose that there are N bond dealers, of which G are good. Each dealer is good with the same probability $g = \frac{G}{N}$. There are therefore S possible states of the world, where each state is a unique combination of matched dealers and interest rates, and each state is equally likely. S is given by:

$$S = \frac{N!}{G!(N-G)!} \tag{54}$$

Information processing is defined as the expected entropy reduction from prior to posterior beliefs, as is common in the rational inattention literature (Sims, 2003):

$$\mathcal{I} = H(h) - \mathbb{E}_s H(f|s) \tag{55}$$

Here H is the entropy function, a measure of the dispersion in a distribution:

$$H(\cdot) = -\sum_{k=1}^{S} \Pr(\text{state } k) \log \Pr(\text{state } k) \tag{56}$$

The distribution h is the prior belief, which is that each state is equally likely (i.e. no bond dealer is more likely to be good than any other). The distribution $f|s$ is the posterior, the distribution of beliefs implied by the signals households process, embodied in the decisions they took.

In the prior each state is equally likely, so:

$$H(h) = -\sum_{k=1}^{S} \frac{1}{S} \log \frac{1}{S} = \log S \tag{57}$$

Households choose signals that maximise their probability of successfully choosing a good bond dealer q for the minimum information cost. This means that if a signal instructs a household member to choose dealer d^*, their posterior belief will be that each state in which d^* is good is equally likely. There is no use in distinguishing between those states. The posterior therefore attaches probability $\frac{qN}{SG}$ to each of those states, where $\frac{SG}{N}$ is the number of states in which d^* is good. Similarly, they assign probability $\frac{(1-q)N}{N-SG}$ to each state of the world in which d^* is bad. The entropy of this posterior is therefore:

$$H(f|s \Rightarrow \text{choose } d^*) = -q \log\big(\frac{qN}{SG}\big) - (1-q)\log\big(\frac{(1-q)N}{N-SG}\big) \tag{58}$$

This is independent of the arbitrary d^*, so this expression also gives the expected entropy across all such signals. The total information processed to attain a success probability of q is therefore:

$$\mathcal{I} = \log S + q \log\big(\frac{qN}{SG}\big) + (1-q)\log\big(\frac{(1-q)N}{N-SG}\big) \tag{59}$$

Splitting apart the logs and cancelling terms we find:

$$\mathcal{I} = \log N - q\log G - (1-q)\log(N-G) + q\log q + (1-q)\log(1-q) \tag{60}$$

Now replace $G = gN$ and simplify to obtain equation 3:

$$\mathcal{I} = q\log\left(\frac{q}{g}\right) + (1-q)\log\left(\frac{1-q}{1-g}\right) \tag{61}$$

All terms involving N, the specific number of bond dealers, have cancelled out, leaving only terms in g, the proportion of bond dealers who are offering good interest rates. This implies that as $\mathcal{I}$ approaches $-\log(g)$, the success probability q approaches 1. While the information processing formula is not defined for $q = 1$, q gets arbitrarily close to 1 for this finite amount of information processing, which does not depend on the number of bond dealers in the market. $\mathcal{I}'(q) \to \infty$ as $q \to 1$.

B.2 The household first order conditions are sufficient for utility maximisation

Here I prove that the household first order conditions (equations 7 - 9) are sufficient for utility maximisation. I only consider a household away from the no-borrowing constraint, as at the constraint they fix attention at zero and their problem becomes identical to the poor household problem in existing TANK papers (e.g Bilbiie, 2008).

The utility maximisation problem of such a household can be written as an unconstrained maximisation by substituting out for consumption using the budget constraint:

$$\max_{n_t, b_t, \tilde{i}_t} U = \sum_t \beta^t \left(\frac{\left(\frac{b_{t-1}}{\Pi_t} + w_t n_t + \mathcal{D}_t - T_t - \frac{b_t}{1+\tilde{i}_t}\right)^{1-\sigma}}{1-\sigma} - \frac{n_t^{1+\varphi}}{1+\varphi} - \mu \mathcal{I}(\tilde{i}_t) \right) \tag{62}$$

The first order conditions are:

$$\frac{\partial U}{\partial b_t} = -\frac{1}{1+\tilde{i}_t}\left(\frac{b_{t-1}}{\Pi_t} + w_t n_t + \mathcal{D}_t - T_t - \frac{b_t}{1+\tilde{i}_t}\right)^{-\sigma} + \mathbb{E}_t \frac{\beta}{\Pi_{t+1}}\left(\frac{b_t}{\Pi_{t+1}} + w_{t+1} n_{t+1} + \mathcal{D}_{t+1} - T_{t+1} - \frac{b_{t+1}}{1+\tilde{i}_{t+1}}\right)^{-\sigma} = 0 \tag{63}$$

$$\frac{\partial U}{\partial n_t} = w_t\left(\frac{b_{t-1}}{\Pi_t} + w_t n_t + \mathcal{D}_t - T_t - \frac{b_t}{1+\tilde{i}_t}\right)^{-\sigma} - n_t^{\varphi} = 0 \tag{64}$$

$$\frac{\partial U}{\partial \tilde{i}_t} = \frac{b_t}{(1+\tilde{i}_t)^2}\left(\frac{b_{t-1}}{\Pi_t} + w_t n_t + \mathcal{D}_t - T_t - \frac{b_t}{1+\tilde{i}_t}\right)^{-\sigma} - \mu \mathcal{I}'(\tilde{i}_t) = 0 \tag{65}$$

The elements of the Hessian matrix are therefore (substituting consumption back in using the budget constraint where it simplifies the expressions):

$$\frac{\partial^2 U}{\partial b_t^2} = -\frac{\sigma}{(1+\tilde{i}_t)^2} c_t^{-\sigma-1} - \mathbb{E}_t \frac{\sigma\beta}{\Pi_{t+1}^2} c_{t+1}^{-\sigma-1} \tag{66}$$

$$\frac{\partial^2 U}{\partial n_t^2} = -w_t^2 \sigma c_t^{-\sigma-1} - \varphi n_t^{\varphi-1} \tag{67}$$

$$\frac{\partial^2 U}{\partial \tilde{i}_t^2} = -\frac{\sigma b_t^2}{(1+\tilde{i}_t)^4} c_t^{-\sigma-1} - \frac{2 b_t}{(1+\tilde{i}_t)^3} c_t^{-\sigma} - \mu \mathcal{I}''(\tilde{i}_t) \tag{68}$$

$$\frac{\partial^2 U}{\partial b_t \partial n_t} = \frac{\sigma w_t}{1+\tilde{i}_t} c_t^{-\sigma-1} \tag{69}$$

$$\frac{\partial^2 U}{\partial b_t \partial \tilde{i}_t} = \frac{1}{(1+\tilde{i}_t)^2} c_t^{-\sigma} + \frac{\sigma b_t}{(1+\tilde{i}_t)^3} c_t^{-\sigma-1} \tag{70}$$

$$\frac{\partial^2 U}{\partial n_t \partial \tilde{i}_t} = -\frac{\sigma b_t w_t}{(1+\tilde{i}_t)^2} c_t^{-\sigma-1} \tag{71}$$

The first order conditions are sufficient for utility maximisation if U is concave. This is the case if the Hessian matrix is negative semi-definite, i.e. if for any real-valued vector $[x \quad y \quad z]$:

$$[x \quad y \quad z] \begin{bmatrix} \frac{\partial^2 U}{\partial b_t^2} & \frac{\partial^2 U}{\partial b_t \partial n_t} & \frac{\partial^2 U}{\partial b_t \partial \tilde{i}_t} \\ \frac{\partial^2 U}{\partial b_t \partial n_t} & \frac{\partial^2 U}{\partial n_t^2} & \frac{\partial^2 U}{\partial n_t \partial \tilde{i}_t} \\ \frac{\partial^2 U}{\partial b_t \partial \tilde{i}_t} & \frac{\partial^2 U}{\partial n_t \partial \tilde{i}_t} & \frac{\partial^2 U}{\partial \tilde{i}_t^2} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \leq 0 \tag{72}$$

Multiplying this out and simplifying we obtain:

$$\begin{aligned} &-\sigma c_t^{-\sigma-1}\left(\frac{x}{1+\tilde{i}_t} - y w_t - \frac{z b_t}{(1+\tilde{i}_t)^2}\right)^2 \\ &\qquad - \mathbb{E}_t \frac{\sigma\beta x^2 c_{t+1}^{-\sigma-1}}{\Pi_{t+1}^2} + \frac{2xzc_t^{-\sigma}}{(1+\tilde{i}_t)^2} - \varphi y^2 n_t^{\varphi-1} - \frac{2z^2 b_t c_t^{-\sigma}}{(1+\tilde{i}_t)^3} - \mu z^2 \mathcal{I}''(\tilde{i}_t) \leq 0 \end{aligned} \tag{73}$$

This expression would clearly be true without the term in xz, which comes from $\frac{\partial^2 U}{\partial b_t \partial \tilde{i}_t}$. This is to be expected: it is the feedback between saving and attention (so interest rates) that is the cause for our concern that U may not be concave. I proceed by showing a condition under which this feedback is sufficiently weak that U remains concave, and so the first order conditions remain sufficient for utility maximisation.

Factorising the two terms containing $c_t^{-\sigma}$ gives:

$$\frac{2xzc_t^{-\sigma}}{(1+\tilde{i}_t)^2} - \frac{2z^2 b_t c_t^{-\sigma}}{(1+\tilde{i}_t)^3} = \frac{c_t^{-\sigma}}{(1+\tilde{i}_t)^3}\big(2xz(1+\tilde{i}_t) - z^2\big) = -\frac{c_t^{-\sigma}}{(1+\tilde{i}_t)^3}\big(z - x(1+\tilde{i}_t)\big)^2 + \frac{x^2 c_t^{-\sigma}}{1+\tilde{i}_t} \tag{74}$$

Putting this back into the full expression, the condition for concavity becomes:

$$\begin{aligned} &-\sigma c_t^{-\sigma-1}\left(\frac{x}{1+\tilde{i}_t} - y w_t - \frac{z b_t}{(1+\tilde{i}_t)^2}\right)^2 \\ &+ x^2\left(\frac{c_t^{-\sigma}}{1+\tilde{i}_t} - \mathbb{E}_t \frac{\sigma\beta c_{t+1}^{-\sigma-1}}{\Pi_{t+1}^2}\right) - \frac{c_t^{-\sigma}}{(1+\tilde{i}_t)^3}\big(z - x(1+\tilde{i}_t)\big)^2 - \varphi y^2 n_t^{\varphi-1} - \mu z^2 \mathcal{I}''(\tilde{i}_t) \leq 0 \end{aligned} \tag{75}$$

Assuming a non-negative information cost μ and an effective interest rate $\tilde{i}_t$ which is greater than -100% (so $1+\tilde{i}_t > 0$), a sufficient condition for the concavity of the utility function is therefore:

$$\frac{c_t^{-\sigma}}{1+\tilde{i}_t} - \mathbb{E}_t \frac{\sigma \beta c_{t+1}^{-\sigma-1}}{\Pi_{t+1}^2} \leq 0 \tag{76}$$

At the first order conditions, the Euler equation gives $c_t^{-\sigma}$ in terms of expected future consumption. I complete the proof by substituting this in to condition 76 and showing that for plausible calibrations, the condition holds. This shows that the first order conditions are a local maximum of the utility function. Since there is only one stationary point of the utility function and the function is continuous, this local maximum must be the global maximum.

Substituting the Euler equation in to expression 76 we obtain:

$$\mathbb{E}_t \frac{\beta c_{t+1}^{-\sigma-1}}{\Pi_{t+1}} \left(c_{t+1} - \frac{\sigma}{\Pi_{t+1}} \right) \leq 0 \tag{77}$$

Therefore if consumption is always less than σ divided by inflation, the first order conditions will be sufficient for utility maximisation. In all calibrations in this paper, $\sigma = 1$ and inflation remains close to its steady state at $\Pi = 1$. Steady state consumption of wealthy households is considerably less than 1 (it is 0.86 in the calibration used for the simulations in sections 3.2.2 and 3.3), and so for all the exercises I consider we remain well within this region, where the first order conditions are sufficient for utility maximisation.

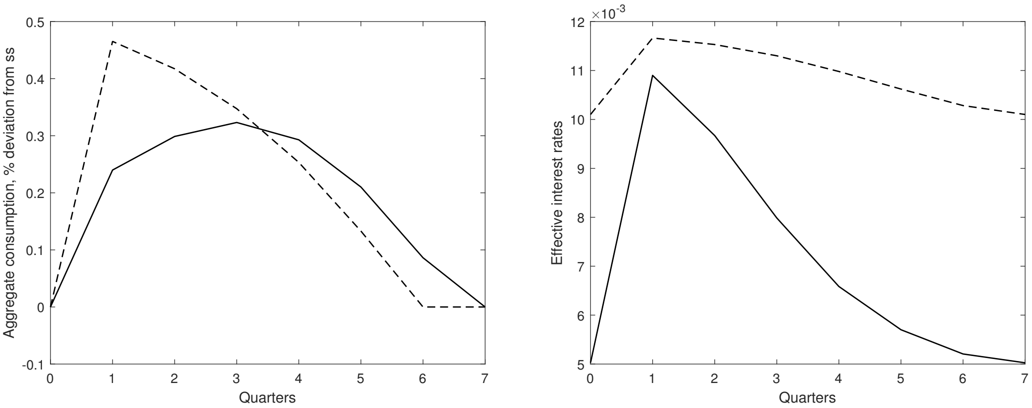
C Intratemporal Redistribution

Here I simulate a one-period redistribution shock z_t, which reduces T_t^p, funded by a contemporaneous rise in T_t^w. This is expansionary in both the attention trap and heterogeneous β models because poor households have a higher MPC than wealthy households.

As with debt-funded tax cuts, for small shocks poor households remain at their borrowing constraint and the two models are the same. If the shock is large enough that poor households temporarily save positive amounts, aggregate consumption responds less on impact, but is more persistent, in the attention trap model, for the same reasons discussed for debt-funded tax cuts in section 3.

Using the same calibration as in section 3, figure 5a shows the response of aggregate consumption in the two models after a redistribution that causes poor households to accumulate the same stock of assets in period 1 as the aggregate tax cut considered in section 3.2.2. Aggregate consumption rises by 48% less on impact in the attention trap model, but is more persistent, returning to steady state a quarter later than in the heterogeneous β model. Figure 5b shows why: effective interest rates facing poor households rise substantially relative to the policy rate in the attention trap model.

Figure 5: Effects of a redistributive tax shock



(a) Aggregate consumption in the attention trap model (solid) and the heterogeneous β model (dashed).

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

D Instability with respect to idiosyncratic shocks

Although there are two steady states of this model, the representative agent steady state is unstable. An arbitrarily small shock to the incomes of a subset of households is sufficient to push the economy permanently away from the representative agent steady state. This highlights a conceptual difference between this model and existing TANK models, where the identity of the poor households is pre-ordained by preferences or market access. Here, which households end up at the borrowing constraint, achieving a lower level of lifetime utility[42], is random. The households are all identical, but when some are hit with a small negative shock the attention trap leads to them becoming permanently poorer.

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

The first step of this argument consists of proposition 2.

Proposition 2 *Assume that the income of household p is strictly less than the income of household w in period t, and that both households expect their incomes to converge in the long run. If there are no further shocks, the income difference between the households*

[42]Constrained households have a lower consumption in steady state than wealthy households in this model. If this is not the case (as in e.g. Bilbiie, 2008) constrained households have a lower level of utility on average as long as asset supply is away from zero in at least one period. In that case constrained households have lower utility because they are unable to smooth consumption intertemporally, and their behaviour is different to those who can.

will be larger in period $t+1$ than in period t if:

$$\frac{1+\tilde{i}_t^p}{1+\tilde{i}_t^w} > \left(\frac{(\sigma-1)\varphi}{\sigma(1+\varphi)}\right)^{\frac{\sigma\varphi}{\sigma+\varphi}} w_t^{-\frac{\sigma(1+\varphi)}{\sigma+\varphi}} c_t^{w^{\sigma}} \tag{78}$$

Proof. Let $\bar{c}$ be the level of consumption both households expect to converge to in the long run. Iterating household p's Euler equation forward gives:

$$c_t^{p^{-\sigma}} = \bar{c}^{-\sigma}\mathbb{E}_t \prod_{s=t}^{\infty} \frac{\beta(1+\tilde{i}_s^p)}{\Pi_{s+1}} \tag{79}$$

Substituting the household p's period t budget constraint and the labour supply condition into the first order condition on attention we obtain:

$$c_t^{p^{-\sigma}}\left(\frac{b_{t-1}^p}{\Pi_t} + w_t^{\frac{1+\varphi}{\varphi}} c_t^{p^{-\frac{\sigma}{\varphi}}} - c_t^p + \mathcal{D}_t - T_t^p\right) = \mu(1+\tilde{i}_t^p)\mathcal{I}(\tilde{i}_t^p) \tag{80}$$

Consider the possibility that household p maintains the same attention, and so the same effective interest rates, as the household w. From the Euler equation, household p must therefore have the same consumption in period t as household w. The left hand side of equation 80 will therefore be smaller for p than w, as income is lower (either due to higher taxes or lower savings from the previous period) and consumption the same. For equation 80 to hold, household p cannot therefore face the same interest rates as household w.

If p faces lower interest rates than w in period t, because they pay less attention, then from the Euler equation they will consume more in period t. This is consistent with equation 80: the left hand side of that equation will be lower for p due to both lower income and higher consumption. Poorer households expecting to return to the representative agent steady state in the long run therefore pay less attention to choosing between different bond dealers and pay a higher price per bond.

I now proceed to show the conditions under which this implies that the gap between the incomes of the households will grow without further shocks. Let the savings of household p coming in to period t be $b_{t-1}^p = b_{t-1}^w - \epsilon_b$, with $\epsilon_b > 0$. The only reason incomes will differ between households in period $t+1$ is because p saves less in period t and so has less asset income in the next period. From the Euler equation, the minimum that household p could consume in period t is given by:

$$c_t^p = c_t^w \left(\frac{1+\tilde{i}_t^w}{1+\tilde{i}_t^p}\right)^{\frac{1}{\sigma}} \tag{81}$$

This is the consumption we would see if the poor household effective interest rate equals that of the wealthy households from period $t+1$ onwards. This provides a maximum for poor household saving, so if the gap in saving between households grows for this level of c_t^p it must grow for all possible c_t^p.

Using the budget constraints of the two households and the labour supply conditions we find that the difference between the savings of the two households in period t with this minimum c_t^p is:

$$b_t^w - b_t^p = \left(\frac{b_{t-1}}{\Pi_t} + \mathcal{D}_t\right)(\tilde{i}_t^w - \tilde{i}_t^p) + w_t^{\frac{1+\varphi}{\varphi}} c_t^{w^{-\frac{\sigma}{\varphi}}}(1+\tilde{i}_t^w)\left(1 - \left(\frac{1+\tilde{i}_t^p}{1+\tilde{i}_t^w}\right)^{\frac{1+\varphi}{\varphi}}\right)$$
$$- c_t^w(1+\tilde{i}_t^w)\left(1 - \left(\frac{1+\tilde{i}_t^p}{1+\tilde{i}_t^w}\right)^{\frac{\sigma-1}{\sigma}}\right) + \frac{\epsilon_b}{\Pi_t}\left(1 + \tilde{i}_t^p + (1-\omega)(\tilde{i}_t^w - \tilde{i}_t^p)\left(1 - \frac{\Pi_t}{1+i_t}\right)\right) \quad (82)$$

Where I have substituted out for taxes using the government budget constraint, noting that in period $t-1$ government debt is $\bar{b}_{t-1} = \omega b_{t-1}^w + (1-\omega)b_{t-1}^p = b_{t-1}^w - (1-\omega)\epsilon_b$.

The first term is always positive, as we have already shown that $\tilde{i}_t^w > \tilde{i}_t^p$.

The second term must also be positive, as $\tilde{i}_t^p < \tilde{i}_t^w$ and $\varphi > 0$. The sign of the third term depends on the magnitude of σ, the inverse elasticity of intertemporal substitution. If $\sigma = 1$ (log utility), this term becomes zero. If $\sigma \in (0,1)$ then $\left(\frac{1+\tilde{i}_t^p}{1+\tilde{i}_t^w}\right)^{\frac{\sigma-1}{\sigma}} > 1$, so the whole term is positive. If $\sigma > 1$ then $\left(\frac{1+\tilde{i}_t^p}{1+\tilde{i}_t^w}\right)^{\frac{\sigma-1}{\sigma}} < 1$ and the term is negative.

A sufficient condition for the sum of terms 2 and 3 to be positive is:

$$\frac{1+\tilde{i}_t^p}{1+\tilde{i}_t^w} > \left(\frac{(\sigma-1)\varphi}{\sigma(1+\varphi)}\right)^{\frac{\sigma\varphi}{\sigma+\varphi}} w_t^{-\frac{\sigma(1+\varphi)}{\sigma+\varphi}} c_t^{w^{\sigma}} \quad (83)$$

To obtain this condition, first note that the sum of the terms is positive if and only if:

$$F := w_t^{\frac{1+\varphi}{\varphi}} c_t^{w^{-\frac{\sigma+\varphi}{\varphi}}}\left(1 - \left(\frac{1+\tilde{i}_t^p}{1+\tilde{i}_t^w}\right)^{\frac{1+\varphi}{\varphi}}\right) - \left(1 - \left(\frac{1+\tilde{i}_t^p}{1+\tilde{i}_t^w}\right)^{\frac{\sigma-1}{\sigma}}\right) > 0 \quad (84)$$

If $\tilde{i}_t^p = \tilde{i}_t^w$, $F = 0$ and the two middle terms in equation 82 sum to zero. F is maximised with respect to $\frac{1+\tilde{i}_t^p}{1+\tilde{i}_t^w}$ at:

$$\left(\frac{1+\tilde{i}_t^p}{1+\tilde{i}_t^w}\right)^* = \left(\frac{(\sigma-1)\varphi}{\sigma(1+\varphi)}\right)^{\frac{\sigma\varphi}{\sigma+\varphi}} w_t^{-\frac{\sigma(1+\varphi)}{\sigma+\varphi}} c_t^{w^{\sigma}} \quad (85)$$

If condition 83 holds, we therefore have that $\frac{1+\tilde{i}_t^p}{1+\tilde{i}_t^w}$ is above the value that maximises F. Since we know $\frac{1+\tilde{i}_t^p}{1+\tilde{i}_t^w} < 1$, this also ensures that the unique maximum of F occurs at $\left(\frac{1+\tilde{i}_t^p}{1+\tilde{i}_t^w}\right)^* < 1$, and since $F = 0$ when $\tilde{i}_t^p = \tilde{i}_t^w$ and F is continuous, we know that this maximum has $F > 0$. With $\frac{1+\tilde{i}_t^p}{1+\tilde{i}_t^w}$ between the level that maximises F and 1, we therefore have that $F \in (0, F^*)$. The second and third terms in equation 82 sum to a positive number. This means that $b_t^w - b_t^p > 0$, so the household that entered period t with more income (w) saves more than the other household.

This translates into a larger income differential in period $t+1$ than in period t if:

$$\frac{b_t^w - b_t^p}{\Pi_{t+1}} > \frac{b_{t-1}^w - b_{t-1}^p}{\Pi_t} = \frac{\epsilon_b}{\Pi_t} \tag{86}$$

A sufficient condition for income differentials to grow is therefore that the real interest rate earned by poor households is positive:

$$\frac{b_t^w - b_t^p}{\Pi_{t+1}} > \frac{\epsilon_b}{\Pi_t} \quad \text{if} \quad \frac{1+\tilde{i}_t^p}{\Pi_{t+1}} > 1 \tag{87}$$

This is because the right hand side of equation 82 is strictly greater than the final term in isolation, which in turn is strictly greater than $\frac{\epsilon_b}{\Pi_t}(1+\tilde{i}_t^p)$ (as long as $1+i_t > \Pi_t$, which will be assured by the Taylor Rule as long as Π_t doesn't deviate too far from steady state[43]).

We know that in the representative agent steady state the real interest rate equals $\beta^{-1} > 1$, so for small shocks this condition will hold. This is also an extremely small subset of the conditions under which income differentials will grow, as it misses the effects of the other terms in equation 82. ■

Condition 78 will hold for a wide range of sensible calibrations. If $\sigma \leq 1$ it will hold, as the right hand side will be (weakly) negative while the left hand side is always strictly positive. All of the numerical results in the paper are for log utility ($\sigma = 1$), so the condition holds.

In general condition 78 is more likely to be satisfied when consumption is low, wages are high, and σ and φ are low, but these are not very restrictive bounds. The non-shaded area in figure 6 shows the range of w_t and c_t^w for which condition 78 holds for all possible values of $\tilde{i}_t^w$ and $\tilde{i}_t^p$, if the policy rate is at its steady state level, for the extremely conservative calibration $\sigma = 3$, $\varphi = 5$. The diamond shows the steady state w, c^w in the representative agent steady state, which is easily within the region in which condition 78 is satisfied. For that combination of w, c^w to be outside of the white region with these parameters, we would require $\frac{1+\tilde{i}_t^p}{1+\tilde{i}_t^w} < 0.75$, implying a more than 2500 basis point difference in the returns faced by the two households, which is implausibly large.

[43] To see this, notice that since $q^w < 1$, $\bar{i} > \bar{\tilde{i}}^w = \beta^{-1} - 1 > 0$. Therefore $1 + i_t > \Pi_t$ if $\beta^{-1} > \Pi_t + (1 - \frac{\Pi_t}{\bar{\Pi}}^{\delta_\Pi} \frac{Y_t}{\bar{Y}}^{\delta_Y})$. Close to steady state the right hand side of this expression will be close to Π_t, which will be close to $\bar{\Pi} = 1$, and so the condition will be satisfied. Another approach is to note that $1+i_t > \Pi_t$ if and only if lump sum taxes are positive in period t (from the government budget constraint assuming asset supply is the same in periods $t-1$ and t). In steady state with $\bar{i} > \bar{\tilde{i}}^w = \beta^{-1} - 1$, $\bar{T} > 1 - \beta > 0$. Close to steady state the condition will therefore be satisfied.

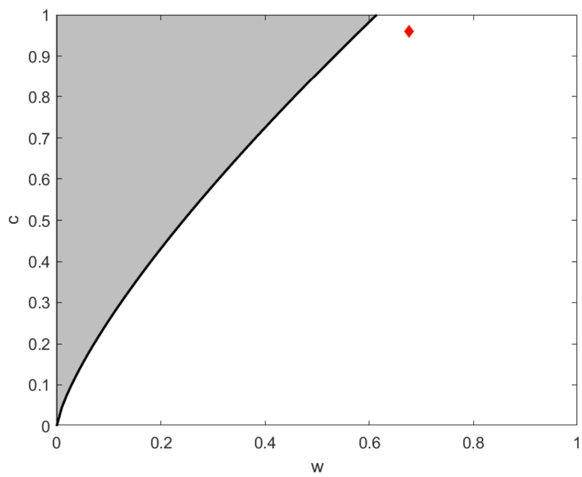


Figure 6: Values of w_t, c_t^w such that condition 78 holds for any possible interest rate differential, for $\sigma = 3$, $\varphi = 5$ (non-shaded region). Steady state values as diamond.

Given this, we have that the representative agent steady state is unstable with respect to small idiosyncratic shocks that reduce the income of some, but not all, households. If the economy is to return to the representative agent steady state after such a shock, the gap between the incomes of the households must at some point begin to shrink. If households expect this, proposition 2 tells us that the income differential can never shrink, and so there is no equilibrium path back to the representative agent steady state that is consistent with rational expectations.

Intuitively, after a small idiosyncratic shock to incomes at the representative agent steady state, the household left poorer will save less, and so will choose to pay less attention to the choice between savings products, than the wealthier household. The poorer household therefore faces a lower effective interest rate, which encourages them to consume more and save even less. Conversely, the wealthy household increases savings and attention, and decreases consumption.

In the next period the poor household therefore has less asset income, because they have saved less at a lower interest rate. They therefore pay even less attention in the period after the idiosyncratic shock. This process continues, with the savings and attention levels of the two households diverging even though the income shock has disappeared. In all periods after the shock, there are no differences between households except their wealth, and that is sufficient to cause further divergence in wealth. The poor household's consumption starts to fall after a while as their asset income declines. Similarly, the wealthy household increases consumption. The new steady state therefore has the wealthy household consuming more, and working less, than the poor household, because of their extra asset income. This feedback between attention and wealth is the 'attention trap'.

E Empirical Exercises

E.1 Demographics

Here I present some summary statistics for demographic variables across responsive and unresponsive states. There are no significant differences race or gender before or after the tax cut, and the differences in age and education are quantitatively small, suggesting that changes in uncertainty between the groups before and after the tax cut are not caused by changes in the composition of the states. The individual fixed effects in the regressions account for the differences in age and education (as respondents only state these in their first month in the survey, so they do not change within individuals).

Table 2: Demographic variables in the two groups of states. $k = 1$ indicates states whose GDP responds significantly to corporate tax cuts in the analysis of Liu and Williams (2019). The means post- tax cut are calculated using households entering the survey after the tax cut, so households in the survey both before and after the cut are counted in the pre- tax cut column.

Variable	Mean pre- tax cut		Difference	Mean post- tax cut		Difference
	$k = 0$	$k = 1$		$k = 0$	$k = 1$	
Age	49.84	50.56	-0.72^{**}	49.72	50.93	-1.20^{**}
			(-2.15)			(-2.20)
Male	0.52	0.52	0.00	0.50	0.48	0.01
			(0.17)			(0.85)
College degree	0.55	0.55	-0.00	0.56	0.54	0.02^{**}
			(-0.77)			(2.74)
White	0.84	0.84	-0.01	0.83	0.83	-0.00
			(-0.92)			(-0.10)

t statistics in parentheses

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

E.2 Examining the mechanism

I have shown that uncertainty about savings interest rates falls after the 2018 tax cut in states likely to be responsive to that cut. To inspect the mechanism, I now examine how the tax cut affected two other variables. I run regression 32 again, replacing the dependent variable with the subjective probability of a rise in savings rates (from which the uncertainty in savings was computed) and log income[44]. The results are in table 3.

The first column shows that the tax cut did not have a significant effect on the subjective probability of a rise in interest rates. The fall in uncertainty associated with

[44]Income is reported in bins in the survey. I take the midpoint of each bin to be the income of everyone in that bin before taking logs, with income in the top-coded bin ($200,000+) set at $250,000. I also remove income from X_{ist} when income is the dependent variable.

Table 3: Regression on interest rate forecasts and income.

	(1) Probability of rate rise	(2) Income
Low income non-fragile	-0.00975	0.165***
	(-0.95)	(3.96)
High income non-fragile	0.00293	-0.443***
	(0.28)	(-9.05)
Low income fragile	0.00110	0.194***
	(0.11)	(4.42)
High income fragile	-0.00700	-0.419***
	(-0.59)	(-7.56)
Individual FE	Yes	Yes
State FE	Yes	Yes
Time FE	Yes	Yes
Clustered SE	No	No
Observations	99748	99748

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

the tax cut found above is not driven by individuals all getting more (or less) optimistic about a rate rise. On average the views of individuals about whether rates will rise or fall remain unchanged, but in responsive states low income non-fragile households move further away from a probability of 50%: they become more certain in their beliefs.

The second column shows that the tax cut is associated with a rise in income for the group of interest. These two results are consistent with the model, in which tax cuts cause poorer households to pay more attention to their saving decisions, and so become less uncertain about savings markets, through a rise in their income[45].

[45] Interestingly, the tax cut is associated with a fall in the income of high income individuals in responsive states. As the mechanism in the model is driven by attention decisions of poor households, I leave exploration of this result to future research.