Sticky Expectations and Consumption Dynamics

February 28, 2019

Christopher D. Carroll[1] JHU
Edmund Crawley[2] JHU
Jiri Slacalek[3] ECB
Kiichi Tokuoka[4] MoF Japan
Matthew N. White[5] UDel

Abstract

To match evidence about aggregate consumption dynamics, representative agent macroeconomic models require a mechanism to generate 'excess smoothness' in consumption. But a large literature has found no evidence of corresponding smoothness in micro data; consequently, existing microfounded models constructed to match the micro data fail to match the macro smoothness facts. We show that the micro and macro evidence can be reconciled by a microfounded model in which consumers have accurate knowledge of their personal circumstances but 'sticky expectations' about the macroeconomy. The persistence of aggregate consumption growth in our model reflects consumers' imperfect attention to aggregate shocks. Our proposed degree of (macro) inattention has negligible utility costs because aggregate shocks constitute only a tiny proportion of the uncertainty that consumers face.

Keywords Consumption, Expectations, Habits, Inattention

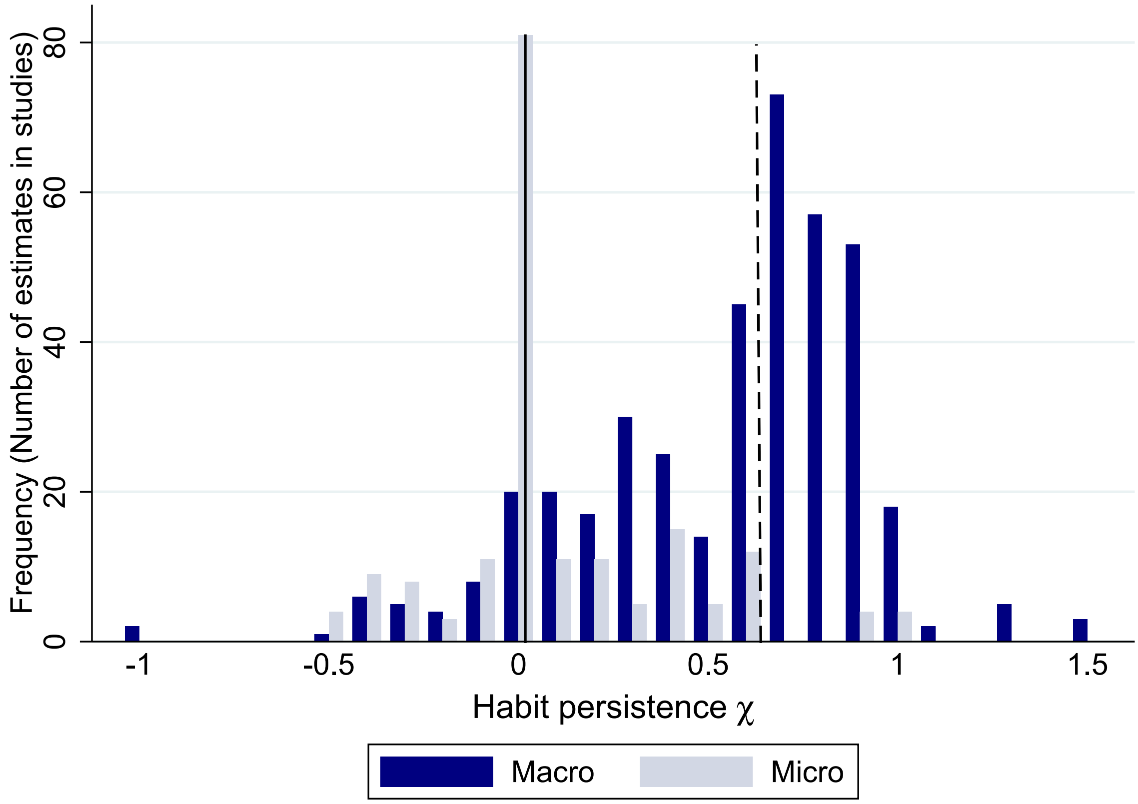
JEL codes D83, D84, E21, E32

Repo: https://github.com/llorracc/cAndCwithStickyE
PDF: http://econ.jhu.edu/people/ccarroll/papers/cAndCwithStickyE.pdf
Appendix: http://econ.jhu.edu/people/ccarroll/papers/cAndCwithStickyE-App.pdf
Web: http://econ.jhu.edu/people/ccarroll/papers/cAndCwithStickyE/
Slides: Versions to View or Print

[1]Carroll: Department of Economics, Johns Hopkins University, http://econ.jhu.edu/people/ccarroll/, ccarroll@jhu.edu [2]Crawley: Department of Economics, Johns Hopkins University, ecrawle2@jhu.edu [3]Slacalek: DG Research, European Central Bank, http://www.slacalek.com/, jiri.slacalek@ecb.europa.eu [4]Tokuoka: Ministry of Finance, Japan, kiichi.tokuoka@mof.go.jp [5]White: Department of Economics, University of Delaware, mnwecon@udel.edu

The computational results in this paper were constructed using tools in the Econ-ARK/HARK toolkit. The toolkit

Figure 1 Distribution of Estimates of Habit Persistence in Macro and Micro Studies



Notes: Reproduced from Havranek, Rusnak, and Sokolova (2017), Figure 2. The figure shows the distribution of estimates of habit persistence in studies based on macro and micro data. Solid and dashed lines show the median estimates in micro (0.0) and macro (0.6) studies, respectively.

1 Introduction

Starting with Campbell and Deaton (1989), the macroeconomics, finance, and international economics literatures have concluded that aggregate consumption exhibits 'excess smoothness' compared to the benchmark Hall (1978) random walk model of consumption.[1] For a standard measure of excess smoothness χ (defined more precisely below), Figure 1 shows that studies using aggregate data estimate that $\chi = 0.6$ on average.[2] A careful reading of the literature suggests that the coefficient is higher, perhaps 0.75, in papers where the data are better measured.

[1] In finance, early references are Abel (1990) and Constantinides (1990); in international economics, Gruber (2004).

[2] Figure 1 is reproduced from a recent comprehensive meta-analysis of 597 published estimates by Havranek, Rusnak, and Sokolova (2017).

can be cited by its digital object identifier, 10.5281/zenodo.1001068, in the references as Carroll, White, and Econ-ARK (2017). Thanks to Robert King, Dirk Krueger, Bartosz Maćkowiak, Giorgio Primiceri, Kathrin Schlafmann, Lenno Uusküla, Gianluca Violante, Mirko Wiederholt and seminar participants in the NBER Summer Institute, the Copenhagen Conference on Heterogeneity, the McMaster University, the University of Michigan, and the University of Delaware for constructive and insightful comments which substantially improved this paper. The views presented in this paper are those of the authors, and should not be attributed to the European Central Bank or to the Japanese Ministry of Finance.

In contrast, parallel work using household-level data rejects the existence of any meaningful degree of excess smoothness. The modal estimate of the micro literature is χ of 0; the mean estimate is about 0.1.[3]

We add a simple (and tractable) information friction to an existing benchmark 'microfounded' macro model, and show that the modified model can reconcile the micro and macro empirical facts. As in the standard full-information rational expectations setup, consumers in our framework perfectly ('frictionlessly') perceive their own personal circumstances (employment status, wage rate, wealth, etc). However, information about *macroeconomic* quantities (e.g., aggregate productivity growth) arrives only occasionally (as in the Calvo model of firms' price updating), so that households' macroeconomic expectations are "sticky," as in Mankiw and Reis (2002) and Carroll (2003). We calculate that our proposed degree of (macro) inattention has negligible utility costs because aggregate shocks are small compared to idiosyncratic shocks.

Aggregate consumption sluggishness *à la* Campbell and Deaton (1989) arises in our model as follows. A household whose beliefs about the aggregate economy are out of date will behave in the ways that would have been macroeconomically appropriate (for the consumer's currently observed level of wealth, etc) at the time of their last perception of macroeconomic circumstances. The lag in perception generates a lag in the response of aggregate spending to aggregate developments; the amount of sluggishness will depend on the frequency with which consumers update. When our model's updating frequency is calibrated to match estimates of the degree of inattention for other aggregate variables (e.g., inflation) using direct expectations data from surveys of households, the model's implications for the persistence in aggregate consumption growth match the estimates of the 'excess smoothness' in the macro literature. We demonstrate this using a Campbell–Mankiw regression augmented with lagged consumption growth.

Despite generating appropriate aggregate smoothness, when estimated on simulated *individual* data (corresponding to microeconomic evidence), regressions in the spirit of Dynan (2000) (the seminal paper in the micro 'excess smoothness' literature) reproduce her finding that at the level of individual households, consumption growth has little predictability at quarterly frequency—Dynan (2000)'s regressions typically get $\bar{R}^2$'s of about 0.01, and her largest reported value is 0.02, figures in the ballpark of those generated in our model.[4]

Because our model is formulated as a deviation from a maximizing model, we can calculate explicit utility costs of that deviation, which are small because the comparatively small size of the aggregate shocks means that neglecting them temporarily causes only small and temporary errors in the level of consumption. Consistent with a theme in the literature all the way back to Akerlof and Yellen (1985), we find that the utility penalty from these small errors is low, so that our consumers would be willing to pay very little for even perpetually perfect information about macroeconomic conditions.

[3]For examples of such studies on aggregate data, see results and references in Fuhrer (2000) or Christiano, Eichenbaum, and Evans (2005). For micro data, Dynan (2000) is the best-known study; many others are reported in Havranek, Rusnak, and Sokolova (2017).

[4]Below we explain why the Havranek, Rusnak, and Sokolova (2017) findings, and our model, are not in substantial conflict with the 'excess sensitivity' literature—indeed, our model can explain some puzzles in that literature.

There are many ways besides ours in which information can be imperfect. But the review of the literature in the next section of the paper shows that the alternative imperfect information frameworks that have been proposed to explain excess smoothness are inconsistent with first-order facts from the micro or the macro literatures (sometimes both).

After the literature review, we begin explaining our ideas with a 'toy model' (section 3) in which the key mechanisms can be derived analytically, thanks to extreme simplifying assumptions like quadratic utility and constant factor prices. We next (section 4) present the full version of our model, which abides by the more realistic assumptions (CRRA utility, aggregate as well as individual shocks, etc) that have become conventional respectively in the micro and macro literatures. After calibrating the model (section 4.7), we describe the stylized facts from both the micro and macro literatures that need to be explained by a good microfounded macroeconomic model of consumption, and show that our model robustly reproduces those facts (section 5). We then (section 6) calculate how much a fully informed consumer would be willing to pay at birth to enjoy instantaneous and perfect knowledge of aggregate developments (not much, it turns out).

2 Background and Literature Review

2.1 Imperfect Information

Our approach is related to extensive work on other forms of information frictions. These frictions include 'noisy information' (cf Pischke (1995)); costly information processing, as in models with rational inattention (cf Sims (2003)); and models of bounded rationality (cf Gabaix (2014)).

In rational inattention models, agents have a limited ability to pay attention and allocate it optimally. Early work by Reis (2006) showed explicitly how rational inattention could lead to excess consumption smoothness. Maćkowiak and Wiederholt (2009) built on that work, and more recently Maćkowiak and Wiederholt (2015) study a DSGE model with inattentive consumers and firms using a simple New Keynesian framework in which they replace all sources of slow adjustment (habit formation, Calvo pricing, and wage setting frictions) with rational inattention.[5] Their setup with rational inattention can match the sluggish responses observed in aggregate data, in response both to monetary policy shocks and to technology shocks. A new paper by Luo, Nie, Wang, and Young (2017) studies implications of rational inattention for the dynamics and cross-sectional dispersion of consumption and wealth in a general equilibrium model with CARA utility.

A challenge to the rational inattention approach has been the complexity of solving models that aim to work out the full implications of rational inattention in contexts where the models that match the microeconomic evidence are already formidably mathematically and computationally complex (see below for why this complexity is necessary to match first-order micro consumption facts). The consumption literature on rational

[5] In Maćkowiak and Wiederholt (2015) also households' consumption–saving decisions are subject to rational inattention. In addition, the model features stochastic real interest rates.

inattention has therefore had to adopt simplifying assumptions about the utility function like quadratic (Sims (2003), section 6; Luo (2008)) or CARA (Luo, Nie, Wang, and Young (2017); Reis (2006)), or a highly stylized setup of idiosyncratic and aggregate income shocks.[6]

As a compromise, Gabaix (2014) has recently proposed a framework that is much simpler than the full rational inattention framework of Sims (2003), but aims to capture much of its essence. This approach is relatively new, and while it does promise to be more tractable than the full-bore Simsian framework, even the simplified Gabaix approach would be difficult to embed in a model with a standard treatment of transitory and persistent income shocks, precautionary motives, liquidity constraints, and other complexities entailed in modern models of microeconomic consumption decisions.[7] .

It would be similarly challenging to determine how to apply the approaches of Woodford (2002) or Morris and Shin (2006) to our question.[8]

Finally, even for a perfectly attentive consumer, information itself can be imperfect. The seminal work contemplating this possibility was by Muth (1960), whose most direct descendant in the consumption literature is Pischke (1995) (building on Lucas (1973); see also Ludvigson and Michaelides (2001)). The idea is that (perfectly attentive) consumers face a signal extraction problem in determining whether a shock to income is transitory or permanent. When a permanent shock occurs, the immediate adjustment to the shock is only partial, since agents' best guess is that the shock is partly transitory and partly permanent. With the right calibration, such a model could in principle explain any amount of excess smoothness. But we argue in section 7 that when a model of this kind is calibrated to the actual empirical data, it generates far less smoothness than exhibited in the data.

2.2 Microfoundations

No review of the empirical literature on smoothness is needed; Havranek, Rusnak, and Sokolova (2017) have done an admirable job.

As for matching "first-order" micro facts, a large empirical literature over the last several decades has documented the importance of modeling precautionary saving behavior under uncertainty. For example, in micro data there is incontrovertible evidence—most recently from millions of datapoints from the Norwegian population registry examined by Fagereng, Holm, and Natvik (2017)—that the consumption function is not linear with

[6]Sims (2006) considers a 2-period consumption–saving model with log utility. Otherwise, to our knowledge, the only paper that employs the CRRA utility to solve a consumption–saving problem under rational inattention is Tutino (2013). Her contribution is mainly methodological, as her setup is quite stylized (e.g., an i.i.d. income process). It would be interesting to extend her work to a more realistic setup (with permanent/persistent income shocks) and study quantitative implications of rational inattention in a model with both idiosyncratic and aggregate income components.

[7]Gabaix (2014) proposes a framework in which consumers perceive a simplified version of the world because there is a cost to paying attention. The existence of a fixed cost of paying attention means that beliefs are not updated continuously but episodically, and the framework generates dynamics that, when aggregated, resemble partial adjustment dynamics. It is beyond the scope of this paper (and would be an interesting project in itself) to determine how this framework would apply in a context like ours, where there are four distinct kinds of shocks (aggregate and idiosyncratic, transitory and permanent), each with very different rewards to attention.

[8]Arguably, our Calvo-style updating is not too different from what one might get in a suitably adapted version of the Morris and Shin (2006) model.

respect to wealth.[9] It is concave, as the general theory says it should be (Carroll and Kimball (1996)), and this concavity matters greatly for matching the main micro facts. In addition, there is also nothing that looks either like the Reis model's prediction that there will be extended periods in which consumption does not change at all, or its prediction that there will be occasional periods in which consumption moves a lot (at dates of adjustment) and then remains anchored at that newer level for another extended period (a similar result holds in the rational-inattention setup of Tutino (2013)). This critique applies generically to models that incorporate a convex cost of adjustment—whether to the consumer's stock of information (Reis (2006)) or to the level of consumption as in Chetty and Szeidl (2016). All such models imply counterfactually 'jerky' behavior of spending at the microeconomic level.[10]

To better match the micro data, we use the now-conventional microeconomic formulation in which utility takes the Constant Relative Risk Aversion form and uncertainty is calibrated to match micro estimates. Our assumption that consumers can perfectly observe the idiosyncratic components of their income allows us to use essentially the same solution methods as in the large recent literature exploring models of this kind. Implementing the state of the art in the micro literature adds a great deal of complexity and precludes a closed form solution for consumption like the one used by Reis. The payoff is that the model is quantitatively plausible enough that, for example, it might actually be usable by policymakers who wanted to assess the likely aggregate dynamics entailed by specific alternative fiscal policy options.

Finally, there is an interesting and growing literature that uses expectations data from surveys in an attempt to directly measure sluggishness in expectations dynamics.[11] For example, Coibion and Gorodnichenko (2015) find that the implied degree of information rigidity in inflation expectations is high, with an average duration of six to seven months between information updates. Fuhrer (2017) and Fuhrer (2018) find that even for professional forecasters, forecast revisions are explainable using lagged information, which would not be the case under perfect information processing. These empirical results are consonant with the spirit of our exercise.

3 A Quadratic Utility 'Toy Model'

Here we briefly introduce concepts and notation, and motivate the key result using a simple framework, the classic Hall (1978) random walk model, with time separable quadratic utility and geometric discounting by factor β. Overall wealth $\mathbf{o}$ (the sum of

[9] More empirical evidence that households that are in some way 'constrained' (e.g., have low liquid assets, low income or low credit scores) have large marginal propensities to consume, especially in newer papers, includes: Johnson, Parker, and Souleles (2006), Agarwal, Liu, and Souleles (2007), Blundell, Pistaferri, and Preston (2008), Kaplan, Violante, and Weidner (2014), Jappelli and Pistaferri (2014), Parker (2017) and Aydın (2018).

[10] This pattern *does* match consumers' purchases of durable goods like automobiles; but the 'excess smoothness' facts hold as strongly for aggregate nondurables as for durable goods. The fixed-adjustment-cost framework matches many other economic decisions well—for instance, individual investors adjust their portfolios sporadically even though the prices of many assets experience large fluctuations at high frequency—and Alvarez, Guiso, and Lippi (2012) find "a robust pattern consistent with the assumption that a component of adjustment costs is information gathering" (p. 2273).

[11] We omit a fuller survey of this interesting literature because expectations data will not be our focus here.

human and nonhuman wealth) evolves according to the dynamic budget constraint

$$\mathbf{o}_{t+1} = (\mathbf{o}_t - \mathbf{c}_t)\mathsf{R} + \zeta_{t+1}, \tag{1}$$

where $\mathsf{R} = (1 + \mathsf{r})$ is the interest factor, ζ_{t+1} is a shock to (total) wealth, and $\mathbf{c}$ is the level of consumption.

With no informational frictions, the usual derivations lead to the standard Euler equation:

$$\mathrm{u}'(\mathbf{c}_t) = \mathsf{R}\beta\,\mathbb{E}_t\left[\mathrm{u}'(\mathbf{c}_{t+1})\right],$$

where $\mathbb{E}_t$ denotes an assumption of instantaneous perfect frictionless updating of all information. Quadratic u and $\mathsf{R}\beta = 1$ imply Hall's random walk proposition:

$$\Delta\mathbf{c}_{t+1} = \varepsilon_{t+1}.$$

Consumers spend

$$\mathbf{c}_t = (\mathsf{r}/\mathsf{R})\mathbf{o}_t,$$

because this is exactly the amount that maintains expected wealth unchanged:

$$\mathbb{E}_t[\mathbf{o}_{t+1}] = (\mathbf{o}_t - \mathbf{c}_t)\mathsf{R} = \mathbf{o}_t.$$

3.1 Sticky Expectations

Now suppose consumers update their information about $\mathbf{o}_t$, and therefore their behavior, only occasionally. A consumer who updates in period t obtains precisely the same information that a consumer in a frictionless model would receive, forms the same expectations, and makes the same choices. Nonupdaters, however, behave as though their former expectations had actually come true (since by definition they have learned nothing to disconfirm their prior beliefs). For example, consider a consumer who updates in periods t and $t+n$ but not between. Designating $\widetilde{\mathbf{o}}$ as the consumer's *perception* of wealth:

$$\widetilde{\mathbf{o}}_{t+j} \equiv \mathbb{E}_t[\mathbf{o}_{t+j}] = \mathbf{o}_t \qquad \text{for } 1 \le j < n,$$

the consumer spends according to perceived wealth so that

$$\mathbf{c}_{t+j} = (\mathsf{r}/\mathsf{R})\widetilde{\mathbf{o}}_{t+j} = (\mathsf{r}/\mathsf{R})\mathbf{o}_t = \mathbf{c}_t \qquad \text{for } 1 \le j < n.$$

The dynamics of actual (as distinct from perceived) wealth are given by (1),

$$\mathbf{o}_{t+n} = \mathbf{o}_t + \underbrace{\sum_{s=1}^{n} \mathsf{R}^{n-s}\zeta_{t+s}}_{\equiv \Delta^n \mathbf{o}_{t+n}},$$

so for a consumer who updates in periods t and $t+n$ but not between, the change in consumption is

$$\mathbf{c}_{t+n} - \mathbf{c}_t = (\mathsf{r}/\mathsf{R})\Delta^n \mathbf{o}_{t+n},$$

where $\Delta^n \mathbf{o}_{t+n}$ is white noise because it is a weighted sum of the white noise errors ζ. Thus, consumption follows a random walk across updating periods; consumers who were only observed during their updating periods would never be seen to deviate from the predictions of Hall (1978).

3.2 Aggregation

The economy is populated by consumers indexed by i, distributed uniformly along the unit interval. Aggregate (or equivalently, per capita) consumption is:

$$\mathbf{C}_t = \int_0^1 \mathbf{c}_{t,i}\, \mathrm{d}i.$$

Whether the consumer at location i updates in period t is determined by the realization of the binary random variable $\pi_{t,i}$, which takes the value 1 if consumer i updates in period t and 0 otherwise. Each period's updaters are chosen randomly such that a constant proportion Π update in each period:

$$\begin{aligned} \mathbb{E}[\pi_{t+1,i}] &= \Pi \qquad \forall\, t \text{ and } i, \\ \int_0^1 \pi_{t,i}\, \mathrm{d}i &= \Pi \qquad \forall\, t. \end{aligned}$$

Aggregate consumption is the population-weighted average of per-capita consumption of updaters $\mathbf{C}^{\pi}$ and nonupdaters $\mathbf{C}^{\not\pi}$:

$$\mathbf{C}_{t+1} = \Pi \mathbf{C}^{\pi}_{t+1} + (1-\Pi)\underbrace{\mathbf{C}^{\not\pi}_{t+1}}_{=\mathbf{C}_t}, \tag{2}$$

where per-capita consumption $\mathbf{C}^{\not\pi}_{t+1} = \mathbf{C}_t$ because the nonupdaters at time $t+1$ are a random subset of the population at time t. The first difference of (2) yields:

$$\Delta \mathbf{C}_{t+1} = (1-\Pi)\Delta \mathbf{C}_t + \underbrace{\Pi \Delta \mathbf{C}^{\pi}_{t+1}}_{\equiv \varepsilon_{t+1}},$$

and online Appendix G.1 shows that ε_{t+1} is approximately mean zero.[12] Thus, in the quadratic utility framework the serial correlation of aggregate per-capita consumption changes is an approximate measure of the proportion of nonupdaters.

This is the mechanism behind the exercises presented in section 5. While the details of the informational friction are different in the more realistic model we present in section 4, the same logic and quantitative result hold: the serial correlation of consumption growth approximately equals the proportion of nonupdaters.

Note further that the model does *not* introduce any explicit reason that consumption growth should be related to the predictable component of income growth *a la* Campbell and Mankiw (1989). In a regression of consumption growth on the predictable component of income growth (and nothing else), the coefficient on income growth would entirely derive from whatever correlation predictable income growth might have with lagged

[12] Intuitively, this term mainly reflects the behavior of updating consumers, and is therefore unpredictable with respect to sufficiently delayed information.

consumption growth. This is the pattern we will find below, in both our theoretical and empirical work.

4 Realistic Model

One of the lessons of the consumption literature after Hall (1978) is that his simplifying assumptions (quadratic utility, perfect capital markets, $\mathsf{R}\beta = 1$) are far from innocuous; more plausible assumptions can lead to very different conclusions. In particular, a host of persuasive theoretical and empirical considerations has led to the now-standard assumption of constant relative risk aversion utility, $\mathrm{u}(\mathbf{c}) = \mathbf{c}^{1-\rho}/(1-\rho)$. But when utility is not quadratic, solution of the model requires specification of the exact stochastic structure of the income and transition processes.

Below, we present a model that will be used to simulate the economy under frictionless and sticky expectations. We specify a small open economy (or partial equilibrium) model with a rich and empirically realistic calibration of idiosyncratic and aggregate risk but exogenous interest rates and wages. In the online appendix, we present two alternative closed economy (general equilibrium) models, along with simulation results analogous to those of section 5, replicating our findings in other settings.[13]

In our model, a continuum of agents care about expected lifetime utility derived from CRRA preferences over a unitary consumption good; they geometrically discount future utility flows by discount factor β. Agents inelastically supply one unit of labor, and their only decision in each period t is how to divide their market resources $\mathbf{m}$ between consumption $\mathbf{c}$ and saving in a single asset $\mathbf{a}$. We assume agents are Blanchard (1985) "perpetual youth" consumers: They have a constant probability of death D between periods, and upon death they are are immediately replaced, while their assets are distributed among surviving households in proportion to the recipient's wealth.

4.1 Output, Income, and Productivity

Output is produced by a Cobb–Douglas technology using capital $\mathbf{K}_t$ and (effective) labor $\mathbf{L}_t$; capital depreciates at rate δ immediately after producing output, leaving portion $(1-\delta)$ intact, and as usual the effectiveness of labor depends on the level of aggregate labor productivity. We consider a small open economy with perfect international capital mobility, so that the returns to capital and labor r_t and W_t are exogenously determined (at constant values r and W); this permits a partial equilibrium analysis using only the solution to the individual households' problem.[14]

[13] In online Appendix B, we extend the SOE model to a heterogeneous agents dynamic stochastic general equilibrium (HA-DSGE) model that endogenizes factor returns at the cost of considerably more computation, which gives results substantially the same as the SOE model. Online Appendix C presents a model that abstracts from idiosyncratic income risk (essentially, setting $\sigma_\psi^2 = \sigma_\theta^2 = 0$), and which produces results similar to those of our 'realistic' models. The simplification enables general equilibrium analysis at a small fraction of the computational cost. However, it is neither a representative agent model—the distribution of beliefs must be tracked—nor a respectable heterogeneous agents model, which may reduce its appeal to both audiences.

[14] See the online appendix for model variants with endogenous factor prices.

We represent both aggregate and idiosyncratic productivity levels as having both transitory and permanent components. Large literatures have found that this representation is difficult to improve upon much in either context, and the simplicity of this description yields considerable benefits both in the tractability of the model, and in making its mechanics as easy to understand as possible.

In more detail, aggregate permanent labor productivity P_t grows by factor Φ_t, subject to mean one iid aggregate permanent shocks Ψ_t, so the aggregate productivity state evolves according to a finite Markov chain:

$$P_{t+1} = \Phi_{t+1} P_t \Psi_{t+1}, \text{ where } \qquad \text{Prob}[\Phi_{t+1} = \Phi_k | \Phi_t = \Phi_j] = \Xi_{j,k}, \tag{3}$$

where j and k index the states. The productivity growth factor Φ_t follows a bounded random walk, as in (for example) Edge, Laubach, and Williams (2007), which is part of a literature whose aim is to capture in a simple statistical way the fact that underlying rates of productivity growth seem to vary substantially over time (e.g., fast in the 1950s, slow in the 1970s and 1980s, moderate in the 1990s, and so on; see also Jorgenson, Ho, and Stiroh (2008)).[15] We introduce these slow-moving productivity growth rates not just for realism, but also because we need to perform simulated exercises analogous to those of Campbell and Mankiw (1989) on empirical data, in which consumption growth is regressed on the component of income growth that was predictable using data lagged several quarters. We therefore need a model in which there *is* some predictability in income growth several quarters in the future.

The transitory component of productivity in any period is represented by a mean-one variable Θ_t, so the overall level of aggregate productivity in a given period is $P_t \Theta_t$.

Similarly, each household has an idiosyncratic labor productivity level $p_{t,i}$, which (conditional on survival) evolves according to:

$$p_{t+1,i} = p_{t,i} \psi_{t+1,i}, \tag{4}$$

and like their aggregate counterparts, idiosyncratic permanent productivity shocks are mean one iid ($\mathbb{E}_t[\psi_{t+n,i}] = \mathbb{E}_t[\Psi_{t+n}] = 1 \ \forall \ n > 0$). Total labor productivity for the individual is determined by the interaction of transitory idiosyncratic (θ), transitory aggregate (Θ), permanent idiosyncratic (p), and permanent aggregate (P) factors. When the household supplies one unit of labor, effective labor is:

$$\boldsymbol{\ell}_{t,i} = \underbrace{\theta_{t,i} \Theta_t}_{\equiv \boldsymbol{\theta}_{t,i}} \underbrace{p_{t,i} P_t}_{\equiv \boldsymbol{p}_{t,i}}. \tag{5}$$

Here, θ can be thought of as reflecting, for example, individual unemployment spells, while Θ captures, e.g., disruptions in output due to bad weather. Just like permanent shocks, transitory shocks are mean one and iid, $\mathbb{E}_t[\theta_{t+n,i}] = \mathbb{E}_t[\Theta_{t+n}] = 1 \ \forall \ n > 0$. The idiosyncratic transitory shock has a minimum possible value of 0 (corresponding to an unemployment spell) which occurs with a small finite probability $\wp$. This has the effect of imposing a 'natural borrowing constraint' (cf. Zeldes (1989b)) at zero.

[15] We capture the process by discretizing the range of productivity growth rates within our bounds, and calibrate the Markov transition probability matrix Ξ so that the statistical properties of productivity growth rates exhibited by our process match the corresponding properties measured in U.S. data since the 1950s.

4.2 Perceptions and Behavior

For understanding the decisions of an individual consumer in a frictionless (i.e. perfect information) world the aggregate and idiosyncratic transitory shocks can be combined into a single overall transitory shock indicated by the boldface $\boldsymbol{\theta}$, and the aggregate and idiosyncratic levels of permanent income can be combined as $\boldsymbol{p}$ (likewise, the combined permanent shock is boldface $\boldsymbol{\psi}_{t,i} \equiv \psi_{t,i}\Psi_t$).

All households (frictionless and sticky-expectations alike) in our models always correctly observe the *level* of all household-specific variables—they are able to read their bank statement and paycheck. As will be shown below, frictionless consumers' optimal behavior depends on the *ratios* of those household-specific variables to permanent productivity $\boldsymbol{p}$. That is, for some state variable $\mathbf{x}$ (like market wealth), the optimal choice for the frictionless consumer would depend on $x \equiv \mathbf{x}/\boldsymbol{p}$, where our definition of nonboldface x reflects our notational convention that when a level variable has been normalized by the corresponding measure of productivity, it loses its boldness. The same applies for aggregate variables, e.g. $X \equiv \mathbf{X}/P$.

One reason we assume that both frictionless and sticky-expectations consumers can perceive the idiosyncratic components of their income (the p and θ) is that this is the assumption made by almost all of the 'modern' literature, and therefore makes our paper's results easily comparable with that literature.

But the assumption can be defended on its own terms; it is consistent with evidence from a number of sources.

First, there are at least some shocks whose transitory nature is impossible to misperceive; the best example is lottery winnings in Norway, see again Fagereng, Holm, and Natvik (2017). The consumption responses to those shocks resemble the responses measured in the previous literature to shocks that economists presumed that consumers knew to be transitory. If consumers respond to such shocks in ways similar to their responses to unambiguously transitory shocks like lottery winnings, that would seem to support the proposition that consumers correctly perceive as transitory those other shocks that economists have presumed consumers identified as transitory.

Second, one reason to believe that perception of the idiosyncratic permanent shocks is not difficult comes from Low, Meghir, and Pistaferri (2010), who show that a large proportion of permanent shocks to income occur at the times of job transitions (mostly movements from one job to another). It would be hard to believe that consumers switching jobs were not acutely aware of the difference between the incomes yielded by those two jobs.

Earlier work by Pistaferri (2001) developed a method for decomposing income shocks into permanent and transitory components. He finds that data from a survey in which consumers are explicitly asked about their income expectations provides a powerful tool to estimate the magnitude of permanent versus transitory shocks; relatedly, Guvenen and Smith (2014) find that consumption choices provide important information about subsequent income movements.

More direct and more recent evidence comes from Karahan, Mihaljevich, and Pilossoph (2017). Using data from the New York Fed's Survey of Consumer Expectations

(SCE), they find that on average, the difference between four-month-ahead realizations of household income and four-month-ahead expectations is near zero and the average error is only 0.5 percent. Karahan, Mihaljevich, and Pilossoph (2017) explicitly interpret their evidence from the survey as suggesting that consumers have accurate perceptions of the permanent and transitory components of their income.

A final bit of evidence comes from metadata associated with the *Survey of Consumer Finances*, which asks a question designed to elicit consumers' perceptions of their permanent ("usual") income. A well-known fact in among survey methodologists is that the speed and ease with which consumers answer a question is an indicator of the extent to which they have a clear understanding of the question and are confident in their answer. The SCF question designed to elicit consumers perceptions of their permanent income is an example of such a question: Consumers answer quickly and easily and do not seem to exhibit any confusion about what they are being asked (Kennickell (1995)).

In contrast, we are aware of no corresponding evidence that consumers are well informed about aggregate income (especially at high frequencies). This is why we have assumed that the inattention that drives our model applies only to perceptions of the (tiny) contribution that aggregate productivity state variables $\{P_t, \Phi_t\}$ make to consumers' overall income.

We denote consumer i's *perceptions* about the aggregate state $\{\widetilde{P}_{t,i}, \widetilde{\Phi}_{t,i}\}$. Our key behavioral assumption is twofold:

1. Households always act *as if* their perception of the aggregate state $\{\widetilde{P}_{t,i}, \widetilde{\Phi}_{t,i}\}$ were the true aggregate state $\{P_t, \Phi_t\}$.

2. As in the 'toy model', households form their perception of the aggregate state according to the *expectation* of today's state that corresponds to the information they had the last time they observed the aggregate state.

Given the assumption that the productivity growth factor Φ_t follows a random walk, the second part of the behavioral assumption says that an agent who last observed the true aggregate state n periods ago perceives:

$$\{\widetilde{P}_{t,i}, \widetilde{\Phi}_{t,i}\} = \mathbb{E}_{t-n}\left[\{P_t, \Phi_t\} \middle| \{P_{t-n}, \Phi_{t-n}\}\right] = \left\{\Phi_{t-n}^{n} P_{t-n}, \Phi_{t-n}\right\}. \tag{6}$$

That is, our assumed random walk in productivity growth means that the household believes that the aggregate productivity factor has remained at Φ_{t-n} for the past n periods, and remains there today.[16] For households who observed the true aggregate state this period, $n = 0$ and thus (6) says that $\{\widetilde{P}_{t,i}, \widetilde{\Phi}_{t,i}\} = \{P_t, \Phi_t\}$.

Given their perception of the aggregate level of productivity, the household perceives that their overall permanent productivity level to be $\widetilde{\boldsymbol{p}}_{t,i} = p_{t,i}\widetilde{P}_{t,i}$.

The behavior of a 'sticky expectations' consumer thus differs from that of a frictionless consumer only to the extent that the 'sticky expectations' consumer's perception of aggregate productivity is out of date.

[16] Because of the boundedness of the random walk in productivity growth, this does not hold exactly when the productivity growth state is in the bottommost or topmost Markov bin; we neglect the small error entailed by this fact.

When a household's perception of productivity $\widetilde{\boldsymbol{p}}$ differs from actual productivity, we denote the perceived ratio as, e.g., $\widetilde{x} \equiv \mathbf{x}/\widetilde{\boldsymbol{p}} = \mathbf{x}/(p\widetilde{P})$ where the last equality reflects our assumption that the household perceives the idiosyncratic component of their productivity p without error.

4.3 Transition Dynamics

Infinitely-lived households with a productivity process like (4) would generate a nonergodic distribution of idiosyncratic productivity—as individuals accumulated ever more shocks to their permanent productivities, those productivities would spread out indefinitely across the population with time. To avoid this inconvenience, we make the Blanchard (1985) assumption: Each consumer faces a constant probability of mortality of D. We track death events using a binary indicator:

$$\mathsf{d}_{t+1,i} = \begin{cases} 0 & \text{if consumer at location } i \text{ survives from time } t \text{ to } t+1 \\ 1 & \text{if consumer at location } i \text{ dies between } t \text{ and } t+1. \end{cases}$$

We refer to this henceforth as a 'replacement' event, since the consumer who dies is replaced by an unrelated newborn who happens to inhabit the same location on the number line. The *ex ante* probability of death is identical for each consumer, so that the aggregate mass of consumers who are replaced is time invariant at $\mathsf{D} = \int_0^1 \mathsf{d}_{t,i}\, di$ (again using the notation that capital letters represent aggregate values).

Under the assumption that 'newborns' have the population-average productivity level of 1, the population mean of the idiosyncratic component of permanent income is always $\int_0^1 p_{t,i} di = 1$.[17] Our earlier equation (4) is thus adjusted to:[18]

$$p_{t+1,i} = \begin{cases} p_{t,i}\psi_{t+1,i} & \text{if } \mathsf{d}_{t+1,i} = 0 \\ 1 & \text{if } \mathsf{d}_{t+1,i} = 1. \end{cases}$$

Along with its productivity level, the household's primary state variable when the consumption decision is made is the level of market resources $\mathbf{m}_{t,i}$, which captures both current period labor income $\mathbf{y}_{t,i}$ (the wage rate times the household's effective labor supply) and the resources that come from the agent's capital stock $\mathbf{k}_{t,i}$ (the value of the capital itself plus the capital income it yields):

$$\mathbf{m}_{t,i} = \underbrace{\mathsf{W}_t \boldsymbol{\ell}_{t,i}}_{\equiv \mathbf{y}_{t,i}} + \underbrace{\mathcal{R}_t}_{1-\delta+\mathsf{r}_t} \mathbf{k}_{t,i}. \tag{7}$$

The transition process for $\mathbf{m}$ is broken up, for convenience of analysis, into three steps. 'Assets' at the end of the period are market resources minus consumption:

$$\mathbf{a}_{t,i} = \mathbf{m}_{t,i} - \mathbf{c}_{t,i}. \tag{8}$$

[17] The dynamics and steady state of the population variance of the idiosyncratic component of permanent income p are derived in online Appendix G.2. The variance will exist so long as $(1 - \mathsf{D})\,\mathbb{E}[\psi^2] < 1$, which imposes a loose restriction on the magnitude of the idiosyncratic permanent shocks.

[18] The constant population size permits the analytical convenience of replacing the dying with newborns, but it is important to understand that there is no relationship between successive persons at the same location on the number line; this is not a dynastic model.

Next period's capital is determined from this period's assets via:

$$\mathbf{k}_{t+1,i} \quad = \quad \mathsf{d}_{t+1,i} \cdot 0 + (1 - \mathsf{d}_{t+1,i})\mathbf{a}_{t,i} \big/ (1 - \mathsf{D}), \tag{9}$$

where the first term represents 'newborns' having zero assets, and the second term's division of $\mathbf{a}$ by the survival probability $(1 - \mathsf{D})$ reflects returns to survivors from the Blanchardian insurance scheme.

4.4 Aggregation

The foregoing assumptions permit straightforward aggregation of individual-level variables. Aggregate capital is the population integral of (9):

$$\mathbf{K}_t = \int_0^1 \mathbf{k}_{t,i}\,\mathrm{d}i = \int_0^1 \big((1 - \mathsf{d}_{t,i})\mathbf{a}_{t-1,i} \big/ (1 - \mathsf{D})\big)\,\mathrm{d}i = \int_0^1 \mathbf{a}_{t-1,i}\,\mathrm{d}i = \mathbf{A}_{t-1}. \tag{10}$$

The third equality holds because $(1 - \mathsf{D})^{-1} \int_0^1 (1 - \mathsf{d}_{t,i})\,\mathrm{d}i = 1$ since $\mathsf{d}_{t,i}$ is independent of $\mathbf{a}_{t-1,i}$. Because $\int_0^1 \theta_{t,i} = \int_0^1 p_{t,i} = 1$, aggregate labor supply is

$$\mathbf{L}_t \quad = \quad \int_0^1 \boldsymbol{\ell}_{t,i}\,\mathrm{d}i = \Theta_t P_t. \tag{11}$$

Aggregate market resources can be written as per-capita resources of the survivors times their population mass $(1 - \mathsf{D})$, plus per-capita resources of the newborns times their population mass D:[19]

$$\begin{aligned}
\mathbf{M}_t \quad &= \quad \big(\overbrace{\mathbf{A}_{t-1}\mathcal{R}_t/(1 - \mathsf{D}) + \Theta_t P_t \mathsf{W}}^{\text{per-capita } \mathbf{m} \text{ for survivors}}\big)(1 - \mathsf{D}) + \overbrace{\Theta_t P_t \mathsf{W}_t}^{\text{per-capita } \mathbf{m} \text{ for newborns}} \quad \mathsf{D} \\
&= \quad \mathbf{A}_{t-1}\mathcal{R}_t + \Theta_t P_t \mathsf{W}_t \\
&= \quad \mathbf{K}_t\mathcal{R}_t + \mathbf{L}_t \mathsf{W}_t.
\end{aligned} \tag{12}$$

The productivity-normalized version of (12) says that

$$M_t = A_{t-1}\mathcal{R}_t \big/ (\Psi_t \Phi_t) + \Theta_t \mathsf{W}_t. \tag{13}$$

We will sometimes refer to the factor $P_t \big/ \widetilde{P}_{t,i}$ as the household's 'productivity misperception,' the scaling factor between actual and perceived market resources.

4.5 Model Solution

Because of the assumption of a small open economy, the frictionless consumer's state variables are simply $(\mathbf{m}_{t,i}, p_{t,i}, P_t, \Phi_t)$. As we assume that the sticky expectations consumer behaves according to the decision rules that are optimal for the frictionless consumer but using perceived rather than true values of the state variables, we need only to solve for the frictionless solution.

[19]This identity can also be derived directly as the population integral of (7).

The household's problem in levels can be written in Bellman form as:[20]

$$\mathbf{v}(\mathbf{m}_{t,i}, p_{t,i}, P_t, \Phi_t) \;=\; \max_{\mathbf{c}_{t,i}} \left\{ \mathrm{u}(\mathbf{c}_{t,i}) + \beta\,\mathbb{E}_t\left[(1-\mathsf{d}_{t+1,i})\mathbf{v}(\mathbf{m}_{t+1,i}, p_{t+1,i}, P_{t+1}, \Phi_{t+1})\right]\right\}.$$

Our assumption that the aggregate and idiosyncratic productivity levels both reflect a combination of purely transitory and purely permanent components now permits us to make a transformation that considerably simplifies analysis and solution of the model: When the utility function is in the CRRA class, the problem can be simplified by dividing by $\boldsymbol{p}_{t,i}^{1-\rho} = (p_{t,i}P_t)^{1-\rho}$ while converting to normalized variables as above (e.g., $m_{t,i} = \mathbf{m}_{t,i}/\boldsymbol{p}_{t,i}$).[21] This yields the normalized form of the problem, which has only $m_{t,i}$ and Φ_t as state variables:

$$\begin{aligned}
\mathrm{v}(m_{t,i}, \Phi_t) \;&=\; \max_{c_{t,i}} \;\left\{\mathrm{u}(c_{t,i}) + (1-\mathsf{D})\beta\,\mathbb{E}_t\left[(\Phi_{t+1}\boldsymbol{\psi}_{t+1,i})^{1-\rho}\mathrm{v}(m_{t+1,i}, \Phi_{t+1})\right]\right\} \\
&\text{s.t.} \\
a_{t,i} \;&=\; m_{t,i} - c_{t,i}, \\
k_{t+1,i} \;&=\; a_{t,i}\big/((1-\mathsf{D})\Phi_{t+1}\boldsymbol{\psi}_{t+1,i}), \\
m_{t+1,i} \;&=\; \mathcal{R}k_{t+1,i} + \mathsf{W}\boldsymbol{\theta}_{t+1,i}.
\end{aligned} \tag{14}$$

Defining $\mathsf{R} = \mathcal{R}/(1-\mathsf{D})$, the main requirement for this problem to have a useful solution is an impatience condition:[22]

$$\mathsf{R}\beta\,\mathbb{E}[\boldsymbol{\psi}^{-\rho}] \;<\; 1.$$

Designating the converged normalized consumption function that solves (14) as $\mathrm{c}(m, \Phi)$, the level of consumption for the frictionless consumer can be obtained[23] from

$$\mathbf{c}_{t,i} \;=\; \boldsymbol{p}_{t,i}\mathrm{c}(m_{t,i}, \Phi_t).$$

Because the model is homothetic in $\boldsymbol{p}_{t,i} = p_{t,i}P_t$, this can be equivalently written with the un-normalized consumption function $\mathbf{c}$ as:

$$\mathbf{c}_{t,i} \;=\; \mathbf{c}(\mathbf{m}_{t,i}, p_{t,i}, P_t, \Phi_t).$$

4.6 Frictionless vs Sticky Expectations

Following the same notation as in the motivating section 3, we define an indicator variable for whether household i updates their perception to the true aggregate state in

[20] Subject to definitions (3), (4), (5), (7), (8) and (9).

[21] This is well understood in the literature; for a full exposition, see, e.g., Carroll (2016).

[22] Other parametric restrictions are also necessary, but for typical parameterizations are not likely to be binding; see Carroll (2016) for details. Note that the relevant interest factor is the within-period productivity of capital $\mathcal{R}$ is the gross return adjusted for both of the influences, δ and $(1-\mathsf{D})$, that intervene between the amount of assets with which the consumer ends period t and the amount of productive capital owned by a survivor when capital income is received in period $t+1$.

[23] Online Appendix D.1 provides details for numerically solving the model.

period t:[24]

$$\pi_{t,i} = \begin{cases} 1 & \text{if consumer } i \text{ updates in period } t \\ 0 & \text{if consumer } i \text{ does not update in period } t. \end{cases}$$

The Bernoulli random variable $\pi_{t,i}$ is iid for each household each period, with a probability Π of returning 1. Consistent with (6), household beliefs about the aggregate state evolve according to:

$$\{\widetilde{P}_{t,i}, \widetilde{\Phi}_{t,i}\} = \begin{cases} \{P_t, \Phi_t\} & \text{if } \pi_{t,i} = 1 \\ \{\widetilde{\Phi}_{t-1,i}\widetilde{P}_{t-1,i}, \widetilde{\Phi}_{t-1,i}\} & \text{if } \pi_{t,i} = 0. \end{cases} \tag{15}$$

Under the assumption that consumers treat their belief about the aggregate state as if it were the truth, the relevant inputs for the normalized consumption function $\mathrm{c}(m, \Phi)$ are the household's *perceived* normalized market resources $\widetilde{m}_{t,i} = \mathbf{m}_{t,i}/\widetilde{\boldsymbol{p}}_{t,i} = \big(P_t/\widetilde{P}_{t,i}\big)m_{t,i}$ and perceived aggregate productivity growth $\widetilde{\Phi}_{t,i}$. The household chooses the level of consumption by:

$$\mathbf{c}_{t,i} = \widetilde{\boldsymbol{p}}_{t,i}\mathrm{c}(\widetilde{m}_{t,i}, \widetilde{\Phi}_{t,i}) = \mathbf{c}(\mathbf{m}_{t,i}, p_{t,i}, \widetilde{P}_{t,i}, \widetilde{\Phi}_{t,i}).$$

The behavior of the 'sticky expectations' consumer converges to that of the frictionless consumer as Π approaches 1.

Because households in our model never misperceive the *level* of their own market resources ($\widetilde{\mathbf{m}}_{t,i} = \mathbf{m}_{t,i}$), they can never choose consumption that would violate the budget constraint. Households observe both their level of income $\mathbf{y}_{t,i}$ and its idiosyncratic components $\theta_{t,i}$ and $p_{t,i}$. If they wanted to do so, households could therefore calculate the aggregate component $\Theta_t \times P_t$, which would correspond with the reports of a statistical agency; but they do *not* observe Θ_t and P_t separately (because, in our model as in reality, statistical agencies do not report these objects).

Our assumption is simply that households with sticky expectations neither perceive nor attempt to extract an estimate of the decomposition of that aggregate state into transitory and permanent components. Consumers' misperceptions of aggregate permanent income *do* cause them to make systematic errors—but below we present calculations showing that for the value of Π that we estimate, those errors have small utility costs.

The utility costs would be a bit smaller if consumers were to perform a certainty-equivalent signal extraction and behaved as though the signal-extracted estimate of the aggregate state is "truth" (that is, they ignore the fact that their estimate has an error term), but section 7 analyzes the alternative model in which households perform such a signal extraction and shows that the dynamics of aggregate consumption under this assumption do not match the dynamics that are observed in the aggregate data.

[24] For simplicity, newborns begin life with correct beliefs about the aggregate state. This assumption about newborns' beliefs is numerically inconsequential because the quarterly replacement rate is so low; see section 4.7 for details.

4.6.1 Alternative Beliefs About the Aggregate Income Process

A model in which households understand that their macroeconomic beliefs are out-of-date due to inattention and prudently change their behavior to account for the extent of their uncertainty at any given moment would be far more computationally costly to solve (adding several additional state variables). This reflects the fact that the mathematically correct treatment of widening aggregate uncertainty is formidably difficult. But if the benefits to consumers of keeping track of the consequences of their growing ignorance were large, we might feel that we had no choice but to go down that path.

Consumers' motivation to take account of the widening of their uncertainty as the distance from the last update increases springs from the convexity of marginal utility with respect to larger shocks: Compared to experiencing four shocks of a given size, experiencing one shock that is four times is large is strictly worse. The magnitude of the benefit to consumers from accounting for correctly for their expanding aggregate uncertainty is related to the *degree* to which the one big shock is worse than the four smaller shocks.

To gauge that magnitude, we conducted an experiment. In online Appendix F, we present a specification in which sticky expectations households optimize under the belief that aggregate shocks only arrive in one in four quarters, but with four times the variance of the quarterly shocks, matching approximately how they will actually *perceive* the arrival of macroeconomic information; the consumption function and main results are virtually identical under these alternate beliefs, which makes us comfortable in not attempting the challenging task of computing the optimal behavior that takes into account the widening uncertainty about the aggregate state as the time since the last update increases.

4.7 Calibration

The full set of parameters is presented in Table 1. We offer complete discussion of our calibration in online Appendix A, but a few aspects warrant comment here.

In the SOE model, we set a much lower value of β (0.97) than would be expected given our calibrated return factor ($\mathcal{R} = 1.015$), resulting in agents with wealth holdings around the median observed in the data. This reflects the recent literature finding that for purposes of capturing aggregate consumption dynamics it may be more important to match the behvior of the typical consumer rather than the behavior of the typical holder of a dollar of wealth.[25] Readers who prefer a calibration matching *mean* observed wealth can consult the online appendix for a closed economy general equilibrium model, in which we show that the main results still hold.

We calibrated the process for trend aggregate productivity growth Φ to match measured U.S. productivity data. A Markov process with eleven states ranging between -3.0 percent and $+3.0$ percent (annual), and in which the state changes on average every two quarters, allowed us to fit both the high frequency autocorrelation evidence cited above and the low-frequency component of productivity growth obtained, e.g., by

[25] See, for example, Olafsson and Pagel (2018).

Staiger, Stock, and Watson (2001), Figure 1.9 and Fernald, Hall, Stock, and Watson (2017), Figure 10.

In our calibration, the variance of the idiosyncratic permanent innovations at the quarterly frequency is about 100 times the variance of the aggregate permanent innovations (4×0.00004 divided by 0.012). This is a point worth emphasizing: Idiosyncratic uncertainty is *approximately two orders of magnitude larger* than aggregate uncertainty. While reasonable people could differ a bit from our calibration of either the aggregate or idiosyncratic risk, no plausible calibration of either magnitude will change the fundamental point that the aggregate component of risk is tiny compared to the idiosyncratic component. This is why assuming that people do not pay close attention to the macroeconomic environment is plausible: It makes a negligible contribution to the total uncertainty they face.

4.7.1 Small Aggregate Shocks and Consumption Concavity

A reader who is persuaded of the general importance of precautionary motives and other causes of nonlinearity in the microeconomic consumption function might feel uneasy about our assumption that consumers act in what seems to be a 'certainty equivalent' way with respect to aggregate shocks. The prior paragraph explains why the consequences of this assumption are negligible: Misperception of the level of aggregate productivity is so small that the consumption function is *approximately* linear over the span between the level of consumption that would be correct with full knowledge, and the level of consumption that the consumer actually chooses. The global concavity of the consumption function (and the curvature of marginal utility), which are important for many other purposes, are of little consequence for errors small enough not to interact meaningfully with that nonlinearity. The importance of this insight has recently been emphasized by Boppart, Krusell, and Mitman (2018), who show that assuming that behavior is linear with respect to aggregate shocks has huge benefits for computation of the solution to heterogeneous agent economies, at little cost to microeconomic realism.

We calibrate the probability of updating at $\Pi = 0.25$ per quarter, for several reasons. First, this is the parameter value assumed for the speed of expectations updating by Mankiw and Reis (2002) in their analysis of the consequences of sticky expectations for inflation. They argue that an average frequency of updating of once a year is intuitively plausible. Second, Carroll (2003) estimates an empirical process for the adjustment process for household inflation expectations in which the point estimate of the corresponding parameter is 0.27 for inflation expectations and 0.32 for unemployment expectations; the similarity of these figures suggests 0.25 is a reasonable benchmark, and provides some insulation against the charge that the model is *ad hoc*: It is calibrated in a way that corresponds to estimates of the stickiness of expectations in a fundamentally different context. Finally, empirical results presented below will also suggest a speed of updating for U.S. consumption dynamics of about 0.25 per quarter.

5 Results

The calibrated model can now be used to evaluate the effects of sticky expectations on consumption dynamics. We begin this section with an empirical benchmark on U.S. data that will guide our investigation of the implications of the model. We then demonstrate that simulated data from the sticky expectations models quantitatively and qualitatively reproduces the key patterns of aggregate and idiosyncratic consumption data.

5.1 U.S. Empirical Benchmark

The random walk model provides the framework around which both micro and macro consumption literatures have been organized. Reinterpreted to incorporate CRRA utility and permit time-varying interest rates, the random walk proposition has frequently been formulated as a claim that $\mu = 0$ in regressions of the form:

$$\Delta \log \mathbf{C}_{t+1} \quad = \quad \varsigma + \nu\, \mathbb{E}_t[\mathbf{r}_{t+1}] + \mu X_t + \epsilon_{t+1}, \tag{16}$$

where X_t is any variable whose value was known to consumers when the period-t consumption decision was made, and ϵ_{t+1} is white noise.

For macroeconomic models (including the HA-DSGE setup in online Appendix B), our simulation analysis[26] shows that the relationship between the normalized asset stock A_t and the expected interest rate $\mathbb{E}_t[\mathbf{r}_{t+1}]$ is nearly linear, so (16) can be reformulated with no loss of statistical power as

$$\Delta \log \mathbf{C}_{t+1} \quad = \quad \varsigma + \alpha A_t + \mu X_t + \epsilon_{t+1}.$$

This reformulation is convenient because the literatures on precautionary saving and liquidity constraints since at least Zeldes (1989a,b) have argued that the effects of capital market imperfections can be captured by incorporating a lagged measure of resources like A_t in consumption growth regressions.

Campbell and Mankiw (1989) famously proposed a modification of this model in which a proportion η of income goes to rule-of-thumb consumers who spend $C = Y$ in every period. They argued that η can be estimated by incorporating the predictable component of income growth as an additional regressor. Finally, Dynan (2000) and Sommer (2007) show that in standard habit formation models, the size of the habit formation parameter can be captured by including lagged consumption growth as a regressor. These considerations lead to a benchmark specification of the form:

$$\Delta \log \mathbf{C}_{t+1} \quad = \quad \varsigma + \chi \Delta \log \mathbf{C}_t + \eta\, \mathbb{E}_t[\Delta \log \mathbf{Y}_{t+1}] + \alpha A_t + \epsilon_{t+1}. \tag{17}$$

There is an extensive existing literature on aggregate consumption dynamics, but Sommer (2007) is the only paper we are aware of that estimates an equation of precisely this form in aggregate data. He interprets the serial correlation of consumption growth as reflecting habit formation.[27] However, Sommer's choice of instruments, estimation

[26] Readers can confirm these results using the toolkit for solving the model available at the [Econ-ARK/REMARK](https://github.com/econ-ark/REMARK/tree/master/REMARKs/cAndCwithStickyE) resource; the authors can provide particular specifications to produce all claimed results.

[27] Weber (2002) makes a similar point using a different methodology.

methodology, and tests do not correspond precisely to our purposes here, so we have produced our own estimates using U.S. data.

In Table 2 we conduct a simple empirical exercise along the lines of Sommer's work, modified to correspond to the testable implications of our model for aggregate U.S. data.

First, while the existing empirical literature has tended to focus on spending on nondurables and services, there are reasons to be skeptical about the measurement of quarterly dynamics (or lack of such dynamics) in large portions of the services component of measured spending. Hence, we report results both for the traditional measure of nondurables and services spending, and for the more restricted category of nondurables spending alone. Fortunately, as the table shows, our results are robust to the measure of spending.[28]

Second, Sommer (2007) emphasizes the importance of taking account of the effects of measurement error and transitory shocks on high frequency consumption data. In principle, measurement error in the level of consumption could lead to a severe downward bias in the estimated serial correlation of measured consumption growth as distinct from 'true' consumption growth. The simplest solution to this problem is the classic response to measurement error in any explanatory variable: Instrumental variables estimation. This point is illustrated in the fact that instrumenting drastically increases the estimated serial correlation of consumption growth.

Finally, we needed to balance the desire for the empirical exercise to match the theory with the need for sufficiently powerful instruments. This would not be a problem if, in empirical work, we could use once-lagged instruments as is possible for the theoretical model. However, empirical consumption data are subject to time aggregation bias (Working (1960), Campbell and Mankiw (1989)), which can be remedied by lagging the time-aggregated instruments an extra period. To increase the predictive power of the lagged instruments, we augmented with two variables traditionally known to have predictive power: The Federal Funds rate and the expectations component of the University of Michigan's Index of Consumer Sentiment (cf. Carroll, Fuhrer, and Wilcox (1994)).[29]

Table 2 demonstrates three main points. First, when lagged consumption growth is excluded from the regression equation, the classic Campbell and Mankiw (1989) result holds: Consumption growth is strongly related to predictable income growth. Second, when predictable income growth is excluded but lagged consumption growth is included, the serial correlation of consumption growth is estimated to be in the range of 0.7–0.8, consistent with the Havranek, Rusnak, and Sokolova (2017) survey of the 'habits' literature and very far from the benchmark random walk coefficient of zero. Finally, in the 'horse race' regression that pits predictable income growth against lagged consumption growth, lagged consumption growth retains its statistical significance and

[28] Indeed, similar results hold even when the measure of spending is the broader measure of total personal consumption expenditures, or for an even stricter version of nondurables spending.

[29] An extensive literature has found a broad range of other variables with predictive power for spending growth; our experience is that results similar to those in the table can be obtained with any collection of instruments with a statistically robust predictive capacity for consumption growth.

large point estimate, while the predictable income growth term becomes statistically insignificant (and economically small).[30]

5.2 Simulated Small Open Economy Empirical Estimation

We now present in Table 3 the results that an econometrician would obtain from estimating an equation like (17) using aggregate data generated by our calibrated model. In short, the table shows that aggregate consumption growth in an economy populated by such consumers exhibits a high degree of serial correlation, quantitatively similar to that in empirical data. This occurs even though simulated households with sticky expectations exhibit only modest predictability of idiosyncratic consumption growth, as discussed below in section 5.3.

To generate these results, we simulate the small open economy model for 200 quarters, tracking aggregate dynamics to generate a dataset whose size is similar to the 57 years of NIPA data used for Table 2. Because there is some variation in coefficient estimates depending on the random number generator's seed, we repeat the simulation exercise 100 times. Table 3 reports average point estimates and standard errors across those 100 samples.

Given the relatively long time frame of each sample, and that the idiosyncratic shocks to income are washed away by the law of large numbers, it is feasible to use instrumental variables techniques to obtain the coefficient on the expected growth term. This is the appropriate procedure for comparison with empirical results in any case, since instrumental variables estimation is the standard way of estimating the benchmark Campbell–Mankiw model. As instruments, we use lags of consumption growth, income growth, the wealth–permanent income ratio, and income growth over a two-year span.[31]

Finally, for comparison to empirical results, we take into account Sommer (2007)'s argument (based on Wilcox (1992)) that transitory components of aggregate spending[32] (hurricanes, etc) and high-frequency measurement problems introduce transitory components in measured NIPA consumption expenditure data. Sommer finds that measurement error produces a severe downward bias in the empirical estimate of the serial correlation in consumption growth, relative to the 'true' serial correlation coefficient. To make the simulated data comparable to the measurement-error-distorted empirical data, we multiply our model's simulated aggregate spending data by a white noise error ξ_t:

$$\mathbf{C}_t^* \quad = \quad \mathbf{C}_t \times \xi_t.$$

The standard deviation of ξ_t is set to the value that would cause the observed difference between the OLS and IV estimates of χ in the univariate regression in Table 2 ($\chi^{OLS} = 0.468$ and $\chi^{IV} = 0.830$): $\text{std}(\log(\xi)) = 0.375 \times \text{std}(\Delta \log \mathbf{C}_t)$.

[30] None of these points is a peculiarity of the U.S. data. Carroll, Sommer, and Slacalek (2011) performed similar exercises for all eleven countries for which they could obtain the required data, and robustly obtained similar results across almost all of those countries.

[31] Instruments $\mathbf{Z}_t = \{\Delta \log \mathbf{C}_{t-2}, \Delta \log \mathbf{C}_{t-3}, \Delta \log \mathbf{Y}_{t-2}, \Delta \log \mathbf{Y}_{t-3}, A_{t-2}, A_{t-3}, \Delta_8 \log \mathbf{C}_{t-2}, \Delta_8 \log \mathbf{Y}_{t-2}\}$, where $\Delta_8 \log x_{t-2} \equiv \log x_{t-2} - \log x_{t-10}$.

[32] It is worth pointing out here that in Friedman (1957)'s original statement of the Permanent Income Hypothesis, transitory shocks to expenditures were given equal billing with transitory shocks to income. The subsequent literature deemphasized expenditure shocks, perhaps inappropriately.

The top panel of Table 3 estimates (17) on simulated data for the frictionless economy. The second and third rows indicate that consumption growth is moderately predictable by (instrumented versions of) both its own lag and expected income growth, of comparable magnitude to the empirical benchmark. However, the 'horse race' regression in the bottom row reveals that neither variable is significantly predictive of consumption growth when both are present as regressors—contrary to the robust empirical results from the U.S. and other countries (cf Carroll, Sommer, and Slacalek (2011)). The problem is that for both consumption growth and income growth, most of the predictive power of the instruments stems from the serial correlation of productivity growth Φ_t in the model, so the instrumented versions of the variables are highly correlated with each other. Thus neither has distinct statistical power when they are both included.

In the sticky expectations specification (lower panel), the second-stage $\bar{R}^2$'s are all much higher than in the frictionless model, and more in keeping with the corresponding statistics in NIPA data. This is because high frequency aggregate consumption growth is being driven by the predictable sticky expectations dynamics. The first two rows show that when we introduce measurement error as described above, the OLS estimate is biased downward significantly. As suggested by the analysis of our 'toy model' above, the IV estimate of χ in the second row is close to the $(1-\Pi) = 0.75$ figure that measures the proportion of consumers who do not adjust their expectations in any given period; thus the intuition derived from the toy model survives all the subsequent complications and elaborations. The third row reflects what would have been found by Campbell and Mankiw had they estimated their model on data produced by the simulated 'sticky expectations' economy: The coefficient on predictable component of perceived income growth term is large and highly statistically significant.

The last row of the table presents the 'horse race' between the Campbell–Mankiw model and the sticky expectations model, and shows that the dynamics of consumption are dominated by the serial correlation in the predictable component of consumption growth stemming from the stickiness of expectations. This can be seen not only from the magnitude of the coefficients, but also by comparison of the second-stage $\bar{R}^2$'s, which indicate that the contribution of predictable income growth to the predictability of consumption growth is negligible, increasing the $\bar{R}^2$ from 0.260 to 0.261.

5.3 Simulated Micro Empirical Estimation

Havranek, Rusnak, and Sokolova (2017)'s meta-analysis of the micro literature is consistent with Dynan (2000)'s early finding that there is little evidence of serial correlation in household-level consumption growth. Such a lack of serial correlation is a direct implication of the canonical Hall (1978) certainty-equivalent model with quadratic utility. But in principle, even without habits, a more modern model like ours with precautionary saving motives predicts that there *will* be *some* positive serial correlation in consumption growth. To see why, think of the behavior of a household whose wealth, leading up to date t, was near its target value. In period t, this household experiences a large negative transitory shock to income, pushing buffer stock wealth far below its target. The model says the household will cut back sharply on consumption to rebuild its buffer stock, and during that period of rebuilding the expected growth rate of consumption will be persistently above its long-term rate (but decline toward that rate). That is, in a univariate analysis, consumption growth will exhibit serial correlation.

But as the foregoing discussion suggests, the model says there is a much more direct indicator than lagged consumption growth for current consumption growth: The lagged value of a, the buffer stock of assets.

The same fundamental point holds for a model in which there is an explicit liquidity constraint (our model has no such constraint, but the precautionary motive induces something that looks like a 'soft' liquidity constraint). Zeldes (1989a) pointed out long ago that the Euler equation on which the random walk proposition is based fails to hold for consumers who are liquidity constrained; if consumers with low levels of wealth (relative to their permanent income) are more likely to be constrained, then low wealth consumers will experience systematically faster consumption growth than otherwise-similar high-wealth consumers. Zeldes found empirical evidence of such a pattern, as has a large subsequent literature.

What is less clear is whether models in this class imply that any residual serial correlation will remain once the lagged level of assets has been controlled for. In numerical models like ours, such quantitative questions can be answered only by numerically solving and simulating the model, which is what we do here.

The model predicts that the relationship between $\mathbb{E}_t[\Delta \log c_{t+1,i}]$ and $a_{t,i}$ will be nonlinear and downward sloping, but theory does not imply any specific functional form. We experimented with a number of ways of capturing the role of $a_{t,i}$ but will spare the reader the unedifying discussion of those experiments because they all reached conclusions similar to those of a particularly simple case, inspired by the original analysis of Zeldes (1989a): We simply include a dummy variable that indicates whether last period's $a_{t,i}$ is low. Specifically, we define $\bar{a}_{t,i}$ as 0 if household i's level of a in period t is in the bottom 1 percent of the distribution, and $\bar{a}_{t,i} = 1$ otherwise. (We could have chosen, say, 10 or 20 percent with qualitatively similar, though less quantitatively impressive, results).

So, in data simulated from our SOE model, we estimate regressions of the form:[33]

$$\Delta \log \mathbf{c}_{t+1,i} \quad = \quad \varsigma + \chi \Delta \log \mathbf{c}_{t,i} + \eta \mathbb{E}_{t,i}[\Delta \log \mathbf{y}_{t+1,i}] + \alpha \bar{a}_{t,i} + \epsilon_{t+1,i}.$$

Results for the frictionless model are presented the upper panel of Table 4. For our purposes, the most important conclusion is that the predictable component of idiosyncratic consumption growth is very modest. In the version of the model that corresponds to the thought experiment above, in which consumption growth should have *some* positive serial correlation, the magnitude of that correlation is only 0.019.[34]

The second row of the table presents the results of a Campbell and Mankiw (1989)-

[33] Details of the simulation procedure are in online Appendix D.2.

[34] We do not report standard errors because in exercises like this the only source of such errors should be the finiteness of the sample of simulated agents; we have taken care to always simulate enough agents that the regression standard errors are essentially zero.

type exercise regressing $\Delta \log \mathbf{c}_{t+1,i} = \eta \, \mathbb{E}_{t,i}[\Delta \log \mathbf{y}_{t+1,i}]$. From our definitions above,

$$
\begin{aligned}
\mathbb{E}_{t,i}[\Delta \log \mathbf{y}_{t+1,i}] &= \mathbb{E}_{t,i}[\log \mathbf{p}_{t,i}\Phi_{t+1}\psi_{t+1,i}\Psi_{t+1}\theta_{t+1,i}\Theta_{t+1}] - \log \mathbf{p}_{t,i}\theta_{t,i}\Theta_t, \\
&= \log \mathbf{p}_{t,i}\Phi_t - \log \mathbf{p}_{t,i}\theta_{t,i}\Theta_t, \\
&= \log \Phi_t - \log \theta_{t,i}\Theta_t.
\end{aligned}
$$

Predictable income growth thus has two components: One deriving from the consumer's beliefs about the underlying aggregate productivity growth rate, and one deriving from the expectation that transitory shocks will revert to their mean value of $\mathbb{E}[\theta\Theta] = 1$. But as noted earlier, our idiosyncratic shocks are vastly larger than aggregate ones, so virtually all of the variation in predicted income growth comes from the $-\log \theta_{t,i}\Theta_t$ term. This explains why the η coefficient, while positive, is close to zero: The model says that the quarterly MPC out of a known-to-be-transitory shock is small, so knowledge that the shock will reverse itself quickly yields only modest predictability.

The third row confirms the proposition articulated above: For people with very low levels of wealth, the model implies rapid consumption growth as they dig themselves out of their hole.

The final row presents the results when all three terms are present. Interestingly, the coefficient on lagged consumption growth actually increases, to about 0.06, when we control for the other two terms. But this is still easily in the range of estimates from 0.0 to 0.1 that Havranek, Rusnak, and Sokolova (2017) indicate characterizes the micro literature.

The crucial point to note from the frictionless model is the very small values of the $\bar{R}^2$'s. Even the version of the model including all three explanatory variables can explain only about 2 percent of the variation in consumption growth—around the maximum degree $\bar{R}^2$ found in the above-cited work of Dynan (2000).

The table's lower panel contains results from estimating the same regressions on the sticky expectations version of the model. These results are virtually indistinguishable from those obtained for the frictionless expectations model. As before, aside from the precautionary component captured by α, idiosyncratic consumption growth is largely unpredictable.

5.4 Excess Sensitivity of Consumption

5.4.1 Relation to the Literature

Our results here might seem to be at variance with the 'excess sensitivity' literature, with prominent contributions for example by Souleles (1999), Johnson, Parker, and Souleles (2006), and Parker, Souleles, Johnson, and McClelland (2013). That literature finds a number of natural experiments in which microeconomic consumers' spending growth is related to changes in their income that, in principle, they could have known about in advance (see also work by Kueng (2012), who finds similar results).

Browning and Collado (2001), in an early summary of the literature, argue that the best way to reconcile the varying microeconomic findings is to suppose that consumers are not always fully aware of the predictable components of their incomes, an explanation that has recently been echoed by Parker (2017).

When we assumed that consumers generally know the idiosyncratic components of their income, we were thinking of the kinds of shocks that are normal everyday occurrences and about which information flows automatically to consumers through regular channels like receipt of their paycheck or taking a new job. Rare events that are outside of ordinary experience, like a once-every-ten-years stimulus check, seem more like our macro than micro shocks. The channels by which consumers might be imagined to learn about these things in advance—news stories, in particular—are the same kinds of sources through which consumers presumably learn about macroeconomic news to which we have assumed they are inattentive.[35]

The plausibility of the inattention explanation of the 'excess sensitivity' results is bolstered by the fact that the natural experiments that this literature relies upon tend to be rare (stimulus checks designed to ward off recessions come once every ten years or so), or quirky (payments from the Alaska Permanent Fund (Hsieh (2003), Kueng (2018)). Furthermore, while many of the individual studies are statistically convincing with respect to their particular experiment, the conclusions across studies are sometimes difficult to reconcile (see Hsieh (2003) or Coulibaly and Li (2006) for counterexamples to the general tendency of the literature's findings); Kueng (2018), for example, finds a higher MPC for high-income than for low-income consumers, in contrast with much of the rest of the literature).

5.4.2 Excess Sensitivity of Consumption to a Fiscal Stimulus

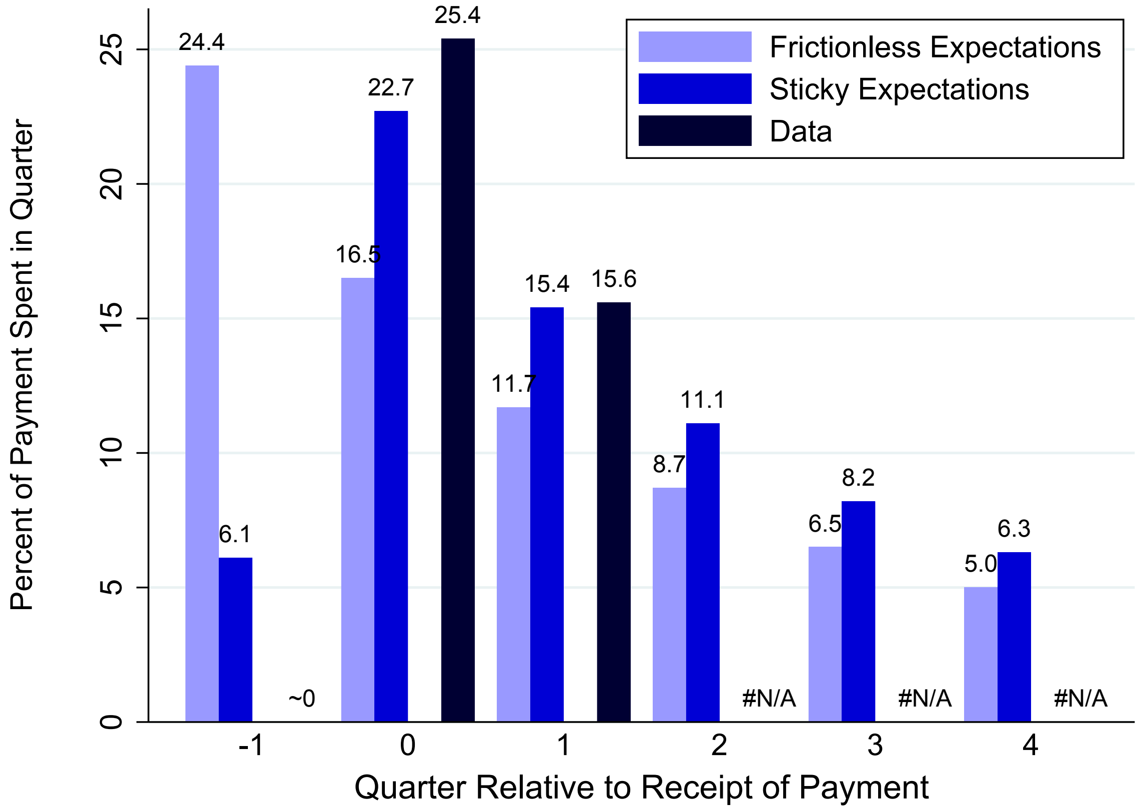
We will now consider the implications of our model for what we take to be the best-established work, by Parker and various collaborators, on the consumption response to fiscal stimulus checks. We focus on this work in part because it has found roughly comparable results across a number of different experiments and in part because it addresses a question that is clearly of first order importance for macroeconomics and in particular fiscal policy. Specifically, we perform a model experiment designed to correspond to the 2008 U.S. federal economic stimulus in which stimulus checks are announced before they are received, and we assume that the announcement of this program is treated in the same way other macro news is treated. We will show that a version of our model is consistent with little reaction of spending upon announcement (Broda and Parker (2014), Parker (2017)) and also with the result that 12–30 percent of the payments was spent on nondurables in the three months in which the payment arrived (Parker, Souleles, Johnson, and McClelland (2013)).

For this experiment, we employ a variant of our model that allows for ex-ante heterogeneity in households' discount factors, following Carroll, Slacalek, Tokuoka, and White (2017).[36] By allowing for heterogeneity in the discount factor, we are able to calibrate the

[35] A prominent exception to this is Olafsson and Pagel (2018), who use daily data and find excess sensitivity to regular payday receipts, a clear micro phenomenon. While this finding is important, our paper is focused on longer (quarterly) frequencies.

[36] We also add unemployment insurance for this experiment.

Figure 2 Effects of Fiscal Stimulus Payments on Consumption, Models vs. Data



Notes: The figure shows how consumption reacts to a fiscal stimulus payment in data and in models with frictionless and sticky expectations. The evidence from data is based on Parker, Souleles, Johnson, and McClelland (2013), Table 5 and Broda and Parker (2014) (the lack of reaction of consumption in quarter −1, "∼ 0" before the payment is received). To our knowledge, literature does not estimate the reaction in quarters 2 through 4, "#N/A".

model to the distribution of wealth (and in particular the large fraction of the population with low levels of liquid wealth). In keeping with related work by Kaplan, Violante, and Weidner (2014), Kaplan, Moll, and Violante (2018), and others who emphasize the role of liquid assets, we calibrate the distribution of discount factors to match the empirical distribution of liquid wealth; Carroll, Slacalek, Tokuoka, and White (2017) show that when their model is calibrated in that way, it generates an annual MPC of around 0.5.[37]

Our exact experiment is as follows. An announcement is made in quarter $t-1$ that stimulus checks will arrive in consumers' bank accounts in period t.[38] In line with our sticky expectation parameter, we assume 25 percent of households learn about the payment when it is announced, while the other three quarters of households are unaware until the payment arrives in period t. Furthermore, we assume the households who know about the upcoming payment are able to borrow against it in period $t-1$.

The experiment sharply differentiates the models with frictionless and sticky expectations both upon announcement of the payments and when households receive the payments (Figure 2). Upon announcement, consumption in the frictionless model substantially increases (households spend 24.4 percent of the payment), but under sticky expectations only one quarter of households update their beliefs when the announcement is made and consumption only rises by 6.1 percent of the stimulus payment. This small effect is in line with Broda and Parker (2014), who estimate no economically or statistically significant change in spending when the household learns that it will receive a payment. Instead, once the stimulus payment is received, sticky expectations households substantially increase their spending—by 22.7 percent of the payment, right in the middle of the 12–30 percent range estimated in Parker, Souleles, Johnson, and McClelland (2013)—as three quarters of them then learn about the payment by seeing it arrive in their bank account. In contrast, in the frictionless setup the reaction of spending upon the receipt of the payment is more muted (16.5 percent).[39] In the following two quarters, consumption in the sticky expectations model is higher by 15.4 and 11.1 percent of the payment amount respectively. This also fits with the empirical evidence suggesting around 40 percent of the stimulus payment is spent in the first three quarters (Parker, Souleles, Johnson, and McClelland (2013)).

The reader's intuition might have been that because our model exhibits little predictability in micro consumption growth when the consumer is experiencing ordinary income shocks (the R^2 of the predictive regression was only a few percent), and because it generates sluggishness in consumption with respect to aggregate shocks, the model would not be able to match the ample micro evidence showing high average MPCs, or

[37] This variant of the model produces similar results to our baseline model with respect to aggregate smoothness. An alternative approach to calibrating the distribution of β would be to target the distribution of MPCs by liquid wealth quantile, as reported for example by Fagereng, Holm, and Natvik (2017) or Crawley and Kuchler (2018). We also did this, but the results are too similar to the liquid wealth calibration to justify reporting. We get similar (albeit lower) consumption responses when we calibrate the distribution of β to match the distribution of net wealth.

[38] This approximately fits the 2008 stimulus timetable. The announcement was made in February and the payments arrived between May and July. We also ran the experiment with two and three quarters advance notice and find the response on receipt of the payment remains in the right empirical range (19.9 and 16.7 percent respectively).

[39] The identification method of Parker, Souleles, Johnson, and McClelland (2013) retrieves the difference between households who have received the payment and those who have not. In the sticky expectations model this is 14 percent of the payment, while it is zero in the frictionless model.

the evidence from Parker and his coauthors showing that there is little "anticipatory" spending in advance of stimulus payments but a strong response to such payments once they have arrived. This section shows that, in fact, the model is capable of matching the broad sweep of those micro facts, while continuing to match the aggregate excess smoothness facts. The key is simple: In the version of our model calibrated to match high micro MPC's, people react robustly to shocks they know about, but they mostly don't know about the macro shocks until they see the money appear in their bank accounts.

6 The Utility Costs of Sticky Expectations

To this point, we have taken Π to be exogenous (though reasonably calibrated). Now, we ask what choices consumers would make if they could choose how much attention to pay in a framework where attention has costs. Specifically, we imagine that newborns make a once-and-for-all choice of their idiosyncratic value of Π, yielding an intuitive approximating formula for the optimal updating frequency.[40] We then conduct a numerical exercise to compute the cost of stickiness for our calibrated models. The utility penalty of having Π equal to our calibrated value of 0.25, rather than updating every period ($\Pi = 1$), are on the order of one two-thousandth of lifetime consumption, so that even small informational costs would justify updating aggregate information only occasionally. Benefits of updating would be even smaller if the update yielded imperfect information about the true state of the macroeconomy; see below.

In the first period of life, we assume that the consumer is employed and experiences no transitory shocks, so that market resources are nonstochastically equal to W_t; value can therefore be written as $\mathrm{v}(\mathsf{W}_t, \cdot)$. There is no analytical expression for v; but, fixing all parameters aside from the variance of the permanent aggregate shock, theoretical considerations suggest (and numerical experiments confirm) that the consequences of permanent uncertainty for value can be well approximated by:

$$\mathrm{v}(\mathsf{W}_t, \cdot) \approx \grave{\mathrm{v}}(\mathsf{W}_t, \cdot) - \kappa \sigma_\Psi^2,$$

where $\grave{\mathrm{v}}(\mathsf{W}_t, \cdot)$ is the value that would be generated by a model with no aggregate permanent shocks and κ is a constant of approximation that captures the cost of aggregate permanent uncertainty (effectively, it is the coefficient on a first order Taylor expansion of the model around the point $\sigma_\Psi^2 = 0$).

Suppose now (again confirmed numerically—see Figure 3) that the effect of sticky expectations is approximately to reduce value by an amount proportional to the inverse of the updating probability:

$$\widetilde{\mathrm{v}}(\mathsf{W}_t, \cdot) \approx \grave{\mathrm{v}}(\mathsf{W}_t, \cdot) - (\kappa/\Pi)\sigma_\Psi^2. \tag{18}$$

This assumption has appropriate scaling properties in three senses:

- If $\sigma_\Psi^2 = 0$ so that there are no permanent shocks, then the cost of stickiness is zero (given our assumption that initial perceptions are correct).

[40] For a more thorough theoretical examination of the tradeoffs in a related model, see Reis (2006).

- If the probability of updating is $\Pi = 1$ so that perceptions are always accurate, value is the same as in the frictionless model.

- If expectations never adjust, then $\Pi = 0$ and the utility cost of stickiness is infinite, which is appropriate because consumers would be making choices based on expectations that would eventually be arbitrarily far from the truth.

Now imagine that newborns make a once-and-for-all choice of the value of Π; a higher Π (faster updating) is assumed to have a linear cost ι in units of normalized value.[41] The newborn's objective is therefore to choose the Π that solves:

$$\max_{\Pi} \ \grave{\mathrm{v}}(\mathsf{W}_t, \cdot) - (\kappa/\Pi)\sigma_{\Psi}^2 - \iota\Pi.$$

The first order condition is:

$$\begin{aligned} 0 &= \Pi^{-2}\kappa\sigma_{\Psi}^2 - \iota, \\ \Pi^2 &= (\kappa\sigma_{\Psi}^2)/\iota, \end{aligned}$$

which leads to the conclusion that the consumer will pick the Π satisfying:

$$\Pi = (\kappa/\iota)^{0.5}\sigma_{\Psi}.$$

Thus, the speed of updating should be related directly to the utility cost of permanent uncertainty (κ), inversely to the cost of information (cheaper information induces faster updating), and linearly to the standard deviation of permanent aggregate shocks.

Our calibrated models can be used to numerically calculate the welfare loss from our specification of sticky expectations as an agent's willingness to pay at birth in order to avoid having $\Pi = 0.25$ for his entire lifetime.[42] Specifically, we calculate the percentage loss of permanent income that would make a newborn indifferent between living in the world with $\Pi = 0.25$, or living in a frictionless world after paying the cost of abolishing the friction.[43]

Using notation from the theoretical exercise above, define a newborn's average lifetime (normalized) value at birth under frictionless and sticky expectations as respectively:

$$\overline{\mathrm{v}}_0 \equiv \mathbb{E}\left[\mathrm{v}(\mathsf{W}_t, \cdot)\right], \qquad \overline{\widetilde{\mathrm{v}}}_0 \equiv \mathbb{E}\left[\widetilde{\mathrm{v}}(\mathsf{W}_t, \cdot)\right],$$

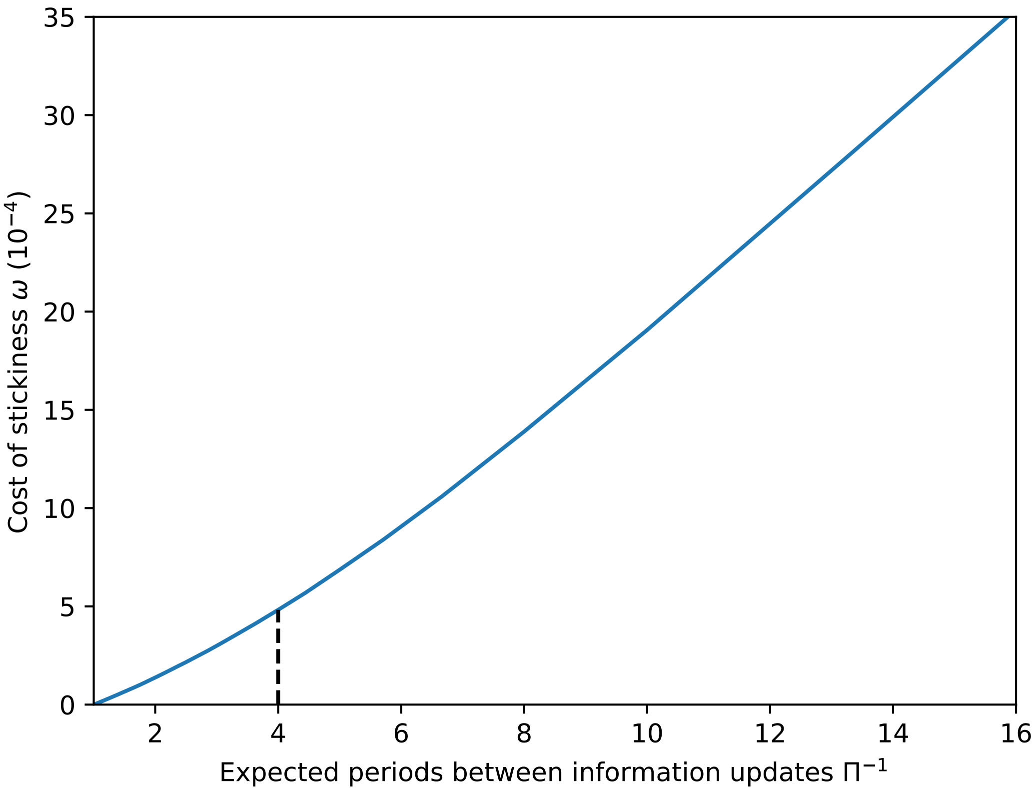
where the expectation is taken over the distribution of state variables other than $m_{t,i}$ that an agent might be born into. We compute these quantities by averaging the discounted sum of consumption utilities experienced by households over their simulated lifetimes. A newborn's willingness to pay (as a fraction of permanent income) to avoid having sticky

[41] Think of this as utility costs of paying attention to macroeconomic news instead of, say, sports news or other more pleasurable news items; since we are examining a model that has been normalized by productivity, this could alternatively be loosely interpreted as a time cost of gathering information.

[42] Additional numeric details of the cost of stickiness calculation can be found in online Appendix D.3.

[43] The measure of the cost of stickiness we report is almost surely larger than the willingness to pay of an agent at a *random period of life*. As newborns do not yet have a buffer stock and are on the steeper portion of their consumption function than they will generally find themselves, consumption errors from sticky expectations are larger in the first few periods of life, and so the utility costs of stickiness are somewhat front-loaded.

Figure 3 Costs of Stickiness ω and Probability of Aggregate Information Updating Π



Notes: The figure shows how the utility costs of updating ω depend on the probability of updating of aggregate information Π in the SOE model.

expectations can then be calculated as:

$$\omega \;=\; 1-\left(\frac{\overline{\overline{\mathrm{v}}}_0}{\overline{\mathrm{v}}_0}\right)^{\frac{1}{1-\rho}}. \tag{19}$$

A newborn in our model is willing to give up about 0.05 percent of his permanent income to remain frictionless. These values are comparable to the findings of Maćkowiak and Wiederholt (2015), who construct a model in which, as in Reis (2006), agents optimally choose how much attention to pay to economic shocks by weighing off costs and benefits. They find (p. 1519) that the cost of suboptimal tracking of aggregate shocks is 0.06 percent of steady state consumption.[44]

Now that we have explained how to compute the cost of stickiness numerically, we can test our supposition in equation (18) that the cost of stickiness might have a roughly inverse linear relationship to Π. Figure 3 plots numerically computed willingness-to-pay ω for various values of Π^{-1}; the relationship is close to linear, as we speculated.

Our preferred interpretation is not that households deliberately choose Π optimally due to a cost of updating, but instead that Π is exogenous and represents the speed with which macroeconomic news arrives "for free" from the news media. This could explain why the parameter 0.25 seems to work about equally well for inflation, unemployment expectations, and consumption – all of them are informed by the same flow of free information. An objection to this interpretation is that a household who has not updated for several years would face a substantially larger loss from continuing to be oblivious and would eventually feel the need to deliberately look up some aggregate facts. At the cost of a large computational and theoretical investment, we could modify the model to allow consumers to behave in this way, but it seems clear that the *ex ante* benefit would be extremely small, because the likelihood of being sufficiently out of date to make costly mistakes is negligible. Intuitively, we can calculate that at any given moment, only 3 percent of households will have information that is more than 3 years out of date ($(1-\Pi)^{12} \approx 0.03$), of households will be in this position. Furthermore, simple calculations show that if we change the simulations so that households always exogenously update after three years, this barely changes aggregate dynamics (the estimate of χ slightly increases from 0.660 to 0.667 in the small open economy model).

7 Muth–Lucas–Pischke and Reis (2006)

Now that our calibrations and results have been presented, we are in position to make some quantitative comparisons of our model to two principal alternatives to habit formation (or our model) for explaining excess smoothness in consumption growth, by Pischke and by Reis.

[44] The theme goes back at least to Cochrane (1989).

7.1 Muth–Lucas–Pischke

The longest-standing rival to habit formation as an explanation of consumption sluggishness is what we will call the Muth–Lucas–Pischke (henceforth, MLP) framework. The idea is not that agents are inattentive, but instead that they have imperfect information on which they perform an optimal signal extraction problem.

Muth (1960)'s agents could observe only the level of their income, but not the split between its permanent and transitory components. He derived the optimal (mean-squared-error-minimizing) method for estimating the level of permanent income from the observed signal about the level of actual income. Lucas (1973) applied the same mathematical toolkit to solve a model in which firms are assumed to be unable to distinguish idiosyncratic from aggregate shocks. Pischke (1995) combines the ideas of Muth and Lucas and applies the result to micro consumption data: His consumers have no ability at all to perceive whether income shocks that hit them are aggregate or idiosyncratic, transitory or permanent. They see only their income, and perform signal extraction on it.

Pischke calibrates his model with micro data in which he calculates that transitory shocks vastly outweigh permanent shocks.[45] So, when a shock arrives, consumers always interpret it as being almost entirely transitory and change their consumption by little. However, macroeconometricians have long known that *aggregate* income shocks are close to permanent. When an aggregate permanent shock comes along, Pischkian consumers spend very little of it, confounding the aggregate permanent shock's effect on their income with the mainly transitory idiosyncratic shocks that account for most of the total variation in their income. This misperception causes sluggishness in *aggregate* consumption dynamics in response to aggregate shocks.

In its assumption that consumers fail to perceive aggregate shocks immediately and fully, Pischke's model resembles ours. However, few papers in the subsequent literature have followed Pischke in making the assumption that households have no idea, when an idiosyncratic income shock occurs, whether it is transitory or permanent. Especially in the last decade or so, the literature instead has almost always assumed that consumers can perfectly perceive the transitory and permanent components of their income; see our defense of this assumption above.

Granting our choice to assume that consumers correctly perceive the events that are idiosyncratic to them (job changes, lottery winnings, etc), there is still a potential role for application of the MLP framework: Instead of assuming sticky expectations, we could instead have assumed that consumers perform a signal extraction exercise on *only* the aggregate component of their income, because they cannot perceive the transitory/permanent split for the (tiny) part of their income change that reflects aggregate macroeconomic developments.

In principle, such confusion could generate excess smoothness; for a detailed description of the mechanism, see online Appendix G.4. But, defining the signal-to-noise ratio

[45] Pischke's estimates constructed from the *Survey of Income and Program Participation* are rather different from the magnitudes of transitory and permanent shocks estimated in the extensive literature—mostly subsequent to Pischke's paper—cited in our calibration section above.

$\varphi = \sigma_{\Psi}^2/\sigma_{\Theta}^2$, Muth's derivations imply that the optimal updating coefficient is:

$$\Pi = \varphi\sqrt{1+\varphi^2/4} - (1/2)\varphi^2 \tag{20}$$

Plugging our calibrations of σ_{Ψ}^2 and σ_{Θ}^2 from section 4.7 into (20), the model yields a predicted value of $(1-\Pi) \approx 0.17$—very far below the approximately 0.6 estimate from Havranek, Rusnak, and Sokolova (2017) and even farther below our estimate of roughly 0.7–0.8 for U.S. data. This reflects the well-known fact that aggregate income is hard to distinguish from a random walk; if it were perceived to be a perfect random walk with no transitory component at all, the serial correlation in its growth would be zero. So, in practice, allowing signal extraction with respect to the aggregate data is not a path to explaining excess smoothness.

7.2 Reis (2006)

Leaving aside our earlier criticisms of its fidelity to microeconomic evidence, the model of Reis (2006) has a further disadvantage relative to any of the other three stories (habits, MLP, or our model) with respect to aggregate dynamics. In Reis's model consumers update their information on a regular schedule—under a plausible calibration of the model, once a year. One implication of the model is that the change in consumption at the next reset is unpredictable; this implies that aggregate consumption growth would be unpredictable at any horizon beyond, say, a year.[46] But, business cycle analysts felt compelled to incorporate sluggishness into macroeconomic models in large part to explain the fact that consumption growth is forecastable over extended periods—empirical impulse response functions indicate that a macroeconomically substantial component of the adjustment to shocks takes place well beyond the one year horizon. A calibration of the Reis model in which consumers update once a year therefore fails to solve a large part of the original problem (of medium-term predictability).

8 Conclusion

Using a traditional utility function that does not incorporate habits, the literature on the microfoundations of consumption behavior has made great strides over the past couple of decades in constructing models that are faithful to many of the microeconomic facts about consumption, income dynamics, and the distribution of wealth. But over roughly the same interval, habit formation has gone from an exotic hypothesis to a standard assumption in the representative agent macroeconomics literature, because habits allow representative agent models to match the measured smoothness in aggregate consumption growth that is of practical importance in quantitative macroeconomic dynamics. This conflict, thrown into sharp focus by the recent meta-analysis of both literatures by Havranek, Rusnak, and Sokolova (2017), is arguably the most important puzzle in the microfoundations of macroeconomic consumption dynamics.

[46] In contrast, our model exhibits significant predictability beyond one year. The value of χ in the 'horse-race' regression for the SOE economy is 0.66 when the right hand side is lagged by one quarter (see Table 3). Adding an extra one and two years' lag to the right hand side sees χ decline approximately as an AR(1), to 0.20 and 0.06 respectively.

We show that this conflict can be resolved by applying insights from the literature on 'inattention' that has developed robustly since the early contributions of Sims (2003), Woodford (2002), Mankiw and Reis (2002), and others. In the presence of such inattention, aggregation of the behavior of microeconomic consumers without habits generates aggregate consumption dynamics that match the 'excess smoothness' facts that have induced the representative agent literature to embrace habits.

The sticky expectations assumption is actually more attractive for modeling consumption than for other areas where it has been more widely applied, because in the consumption context there is a well-defined utility-based metric for calculating the cost of sticky expectations. This is in contrast with, say, models in which households' inflation expectations are sticky; the welfare cost of misperceiving the inflation rate in those models is typically harder to quantify. The cost to consumers of our proposed degree of macroeconomic inattention is quite modest, for reasons that will be familiar to anyone who has worked with both micro and macro data: Idiosyncratic variation is vastly greater than aggregate variation. This means that the small imperfections in macroeconomic perceptions proposed here have very modest utility consequences. So long as consumers respond appropriately to their idiosyncratic shocks (which we assume they do), the failure to keep completely up-to-date with aggregate developments simply does not matter much.

While a number of previous papers have proffered the idea that inattention (or imperfect information) might generate excess smoothness, the modeling question is a quantitative one ('how much excess smoothness can a sensible model explain?'). We argue that the imperfect information models and mechanisms proposed in the prior literature are quantitatively unable simultaneously to match the micro and macro quantitative facts, while our model matches the main stylized facts from both literatures.

In future work, it would be interesting to enrich the model so that it has plausible implications for how the degree of attention might vary over time or across people, and to connect the model to the available expectations data—for example, measures of consumer sentiment, or measures of uncertainty constructed from news sources, cf Baker, Bloom, and Davis (2016). Such work might be particularly useful in any attempt to understand how behavioral dynamics change between normal times in which news coverage of macroeconomic dynamics is not front-page material versus crisis times, when it *is*.

Table 1 Calibration

	Macroeconomic Parameters	
γ	0.36	Capital's Share of Income
δ	$1 - 0.94^{1/4}$	Depreciation Rate
σ^2_{Θ}	0.00001	Variance Aggregate Transitory Shocks
σ^2_{Ψ}	0.00004	Variance Aggregate Permanent Shocks
	Steady State of Perfect Foresight DSGE Model	
	$(\sigma_{\Psi} = \sigma_{\Theta} = \sigma_{\psi} = \sigma_{\theta} = \wp = \mathsf{D} = 0,\ \Phi_t = 1)$	
K/K^{γ}	12.0	SS Capital to Output Ratio
K	48.55	SS Capital to Labor Productivity Ratio ($= 12^{1/(1-\gamma)}$)
W	2.59	SS Wage Rate ($= (1-\gamma)K^{\gamma}$)
r	0.03	SS Interest Rate ($= \gamma K^{\gamma-1}$)
$\mathcal{R}$	1.015	SS Between-Period Return Factor ($= 1 - \delta + \mathsf{r}$)
	Preference Parameters	
ρ	2.	Coefficient of Relative Risk Aversion
β	0.970	Discount Factor (SOE Model)
Π	0.25	Probability of Updating Expectations (if Sticky)
	Idiosyncratic Shock Parameters	
σ^2_{θ}	0.120	Variance Idiosyncratic Tran Shocks ($=4\times$ Annual)
σ^2_{ψ}	0.003	Variance Idiosyncratic Perm Shocks ($=\frac{1}{4}\times$ Annual)
$\wp$	0.050	Probability of Unemployment Spell
D	0.005	Probability of Mortality

Note: As discussed in online Appendix A, we calibrate to the steady state values from a perfect foresight DGSE model.

Table 2 Aggregate Consumption Dynamics in US Data

$$\Delta \log \mathbf{C}_{t+1} = \varsigma + \chi \Delta \log \mathbf{C}_t + \eta \, \mathbb{E}_t[\Delta \log \mathbf{Y}_{t+1}] + \alpha A_t + \epsilon_{t+1}$$

Measure of Consumption Independent Variables			OLS or IV	2nd Stage $\bar{R}^2$	KP p-val Hansen J p val
Nondurables and Services					
$\Delta \log \mathbf{C}_t$	$\Delta \log \mathbf{Y}_{t+1}$	A_t			
0.468***			OLS	0.216	
(0.076)					
0.830***			IV	0.278	0.222
(0.098)					0.439
	0.587***		IV	0.203	0.263
	(0.110)				0.319
		−0.17e−4	IV	−0.005	0.081
		(5.71e−4)			0.181
0.618***	0.305*	−4.96e−4*	IV	0.304	0.415
(0.159)	(0.161)	(2.94e−4)			0.825
Memo: For instruments $\mathbf{Z}_t, \Delta \log \mathbf{C}_t = \mathbf{Z}_t \zeta,\ \bar{R}^2 = 0.358$					
Nondurables					
$\Delta \log \mathbf{C}_t$	$\Delta \log \mathbf{Y}_{t+1}$	A_t			
0.200***			OLS	0.036	
(0.058)					
0.762***			IV	0.083	0.504
(0.284)					0.727
	0.849**		IV	0.061	0.398
	(0.357)				0.731
		9.09e−4	IV	0.008	0.118
		(9.05e−4)			0.446
0.620**	0.313	−3.25e−4	IV	0.077	0.523
(0.292)	(0.286)	(8.32e−4)			0.821
Memo: For instruments $\mathbf{Z}_t, \Delta \log \mathbf{C}_t = \mathbf{Z}_t \zeta,\ \bar{R}^2 = 0.080$					

Notes: Robust standard errors are in parentheses. Instruments $\mathbf{Z}_t = \{\Delta \log \mathbf{C}_{t-2}, \Delta \log \mathbf{C}_{t-3}, \Delta \log \mathbf{Y}_{t-2}, \Delta \log \mathbf{Y}_{t-3}, A_{t-2}, A_{t-3}, \Delta_8 \log \mathbf{C}_{t-2}, \Delta_8 \log \mathbf{Y}_{t-2}$, lags 2 and 3 of differenced Fed funds rate, lags 2 and 3 of the Michigan Index of Consumer Sentiment Expectations$\}$. The penultimate column reports the $\bar{R}^2$ from a regression of the dependent variable on the RHS variables (instrumented, when indicated); the final column reports two tests of instrument validity: The p-value from the Kleibergen–Paap Wald rk F statistic of first-stage instrument validity (top), and the p-value from the Hansen–Sargan overidentification test (bottom). $\{^*, ^{**}, ^{***}\}$ = Statistical significance at $\{10, 5, 1\}$ percent.
Data sources are NIPA and US Financial Accounts, 1960Q1–2016Q4. Income ($\mathbf{Y}$) is measured as as wages, salaries and transfers, net of social insurance. Wealth–income ratio (A_t) is measured as the ratio of net worth to income.

Table 3 Aggregate Consumption Dynamics in SOE Model

$$\Delta \log \mathbf{C}_{t+1} = \varsigma + \chi \Delta \log \mathbf{C}_t + \eta \mathbb{E}_t[\Delta \log \mathbf{Y}_{t+1}] + \alpha A_t + \epsilon_{t+1}$$

Expectations : Dep Var Independent Variables			OLS or IV	2nd Stage $\bar{R}^2$	KP p-val Hansen J p-val
Frictionless : $\Delta \log \mathbf{C}^*_{t+1}$ (with measurement error $\mathbf{C}^*_t = \mathbf{C}_t \times \xi_t$);					
$\Delta \log \mathbf{C}^*_t$	$\Delta \log \mathbf{Y}_{t+1}$	A_t			
0.295$^{\bullet\bullet\bullet}$			OLS	0.087	
(0.066)					
0.660$^{\bullet\bullet}$			IV	0.040	0.237
(0.309)					0.600
	0.457$^{\bullet\bullet}$		IV	0.035	0.059
	(0.209)				0.421
		−6.92e–4	IV	0.026	0.000
		(5.87e–4)			0.365
0.420	0.258	0.45e–4	IV	0.041	0.516
(0.428)	(0.365)	(9.51e–4)			0.529
Memo: For instruments $\mathbf{Z}_t$, $\Delta \log \mathbf{C}^*_t = \mathbf{Z}_t \zeta$, $\bar{R}^2 = 0.039$; var(log(ξ_t)) = 5.99e–6					
Sticky : $\Delta \log \mathbf{C}^*_{t+1}$ (with measurement error $\mathbf{C}^*_t = \mathbf{C}_t \times \xi_t$);					
$\Delta \log \mathbf{C}^*_t$	$\Delta \log \mathbf{Y}_{t+1}$	A_t			
0.508$^{\bullet\bullet\bullet}$			OLS	0.263	
(0.058)					
0.802$^{\bullet\bullet\bullet}$			IV	0.260	0.000
(0.104)					0.554
	0.859$^{\bullet\bullet\bullet}$		IV	0.198	0.060
	(0.182)				0.233
		−8.26e–4$^{\bullet\bullet}$	IV	0.066	0.000
		(3.99e–4)			0.002
0.660$^{\bullet\bullet\bullet}$	0.192	0.60e–4	IV	0.261	0.359
(0.187)	(0.277)	(5.03e–4)			0.546
Memo: For instruments $\mathbf{Z}_t$, $\Delta \log \mathbf{C}^*_t = \mathbf{Z}_t \zeta$, $\bar{R}^2 = 0.260$; var(log(ξ_t)) = 5.99e–6					

Notes: Reported statistics are the average values for 100 samples of 200 simulated quarters each. Bullets indicate that the average sample coefficient divided by average sample standard error is outside of the inner 90%, 95%, and 99% of the standard normal distribution. Instruments $\mathbf{Z}_t = \{\Delta \log \mathbf{C}_{t-2}, \Delta \log \mathbf{C}_{t-3}, \Delta \log \mathbf{Y}_{t-2}, \Delta \log \mathbf{Y}_{t-3}, A_{t-2}, A_{t-3}, \Delta_8 \log \mathbf{C}_{t-2}, \Delta_8 \log \mathbf{Y}_{t-2}\}$.

Table 4 Micro Consumption Regression on Simulated Data

$$\Delta \log \mathbf{c}_{t+1,i} \quad = \quad \varsigma + \chi \Delta \log \mathbf{c}_{t,i} + \eta \mathbb{E}_{t,i}[\Delta \log \mathbf{y}_{t+1,i}] + \alpha \bar{a}_{t,i} + \epsilon_{t+1,i}.$$

Model of Expectations	χ	η	α	$\bar{R}^2$
Frictionless				
	0.019			0.000
	(–)			
		0.011		0.004
		(–)		
			−0.190	0.010
			(–)	
	0.061	0.016	−0.183	0.017
	(–)	(–)	(–)	
Sticky				
	0.012			0.000
	(–)			
		0.011		0.004
		(–)		
			−0.191	0.010
			(–)	
	0.051	0.015	−0.185	0.016
	(–)	(–)	(–)	

Notes: $\mathbb{E}_{t,i}$ is the expectation from the perspective of person i in period t; $\bar{a}$ is a dummy variable indicating that agent i is in the top 99 percent of the normalized a distribution. Simulated sample size is large enough such that standard errors are effectively zero. Sample is restricted to households with positive income in period t. The notation "(—)" indicates that standard errors are close to zero, given the very large simulated sample size.

References

Abel, Andrew B. (1990): "Asset Prices under Habit Formation and Catching Up with the Joneses," *American Economic Review*, 80(2), 38–42.

Agarwal, Sumit, Chunlin Liu, and Nicholas S. Souleles (2007): "The Reaction of Consumer Spending and Debt to Tax Rebates–Evidence from Consumer Credit Data," *Journal of Political Economy*, 115(6), 986–1019.

Akerlof, George A., and Janet L. Yellen (1985): "A Near-rational Model of the Business Cycle, with Wage and Price Intertia," *The Quarterly Journal of Economics*, 100(5), 823–38.

Alvarez, Fernando, Luigi Guiso, and Francesco Lippi (2012): "Durable Consumption and Asset Management with Transaction and Observation Costs," *American Economic Review*, 102(5), 2272–2300.

Aydin, Deniz (2018): "Consumption Response to Credit Expansions: Evidence from Experimental Assignment of 45,307 Credit Lines," mimeo, Washington University, St. Louis.

Baker, Scott R, Nicholas Bloom, and Steven J Davis (2016): "Measuring economic policy uncertainty," *The Quarterly Journal of Economics*, 131(4), 1593–1636.

Blanchard, Olivier J. (1985): "Debt, Deficits, and Finite Horizons," *Journal of Political Economy*, 93(2), 223–247.

Blundell, Richard, Luigi Pistaferri, and Ian Preston (2008): "Consumption Inequality and Partial Insurance," *American Economic Review*, 98(5), 1887–1921.

Boldrin, Michele, Lawrence J. Christiano, and Jonas D. Fisher (2001): "Habit Persistence, Asset Returns and the Business Cycle," *American Economic Review*, 91(1), 149–66.

Boppart, Timo, Per Krusell, and Kurt Mitman (2018): "Exploiting MIT Shocks in Heterogeneous-Agent Economies: The Impulse Response as a Numerical Derivative," *Journal of Economic Dynamics and Control*, 89(C), 68–92.

Broda, Christian, and Jonathan A. Parker (2014): "The Economic Stimulus Payments of 2008 and the Aggregate Demand for Consumption," *Journal of Monetary Economics*, 68(S), 20–36.

Browning, Martin, and M. Dolores Collado (2001): "The Response of Expenditures to Anticipated Income Changes: Panel Data Estimates," *American Economic Review*, 91(3), 681–692.

Campbell, John, and Angus Deaton (1989): "Why is Consumption So Smooth?," *The Review of Economic Studies*, 56(3), 357–373, http://www.jstor.org/stable/2297552.

Campbell, John Y., and N. Gregory Mankiw (1989): "Consumption, Income, and Interest Rates: Reinterpreting the Time-Series Evidence," in *NBER Macroeconomics Annual, 1989*, ed. by Olivier J. Blanchard, and Stanley Fischer, pp. 185–216. MIT Press, Cambridge, MA, http://www.nber.org/papers/w2924.pdf.

Carroll, Christopher D. (2003): "Macroeconomic Expectations of Households and Professional Forecasters," *Quarterly Journal of Economics*, 118(1), 269–298, http://econ.jhu.edu/people/ccarroll/epidemiologyQJE.pdf.

——— (2016): "Theoretical Foundations of Buffer Stock Saving," manuscript, Department of Economics, Johns Hopkins University, Available at http://econ.jhu.edu/people/ccarroll/papers/BufferStockTheory.

Carroll, Christopher D., Jeffrey C. Fuhrer, and David W. Wilcox (1994): "Does Consumer Sentiment Forecast Household Spending? If So, Why?," *American Economic Review*, 84(5), 1397–1408, http://econ.jhu.edu/people/ccarroll/SentAERCarrollFuhrerWilcox.pdf.

Carroll, Christopher D., and Miles S. Kimball (1996): "On the Concavity of the Consumption Function," *Econometrica*, 64(4), 981–992, http://econ.jhu.edu/people/ccarroll/concavity.pdf.

Carroll, Christopher D., and Andrew A. Samwick (1997): "The Nature of Precautionary Wealth," *Journal of Monetary Economics*, 40(1), 41–71.

Carroll, Christopher D, Jiri Slacalek, and Kiichi Tokuoka (2015): "Buffer-Stock Saving in a Krusell–Smith World," *Economics Letters*, 132, 97–100, At http://econ.jhu.edu/people/ccarroll/papers/cstKS/; extended version available as ECB Working Paper number 1633, https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1633.pdf.

Carroll, Christopher D., Jiri Slacalek, Kiichi Tokuoka, and Matthew N. White (2017): "The Distribution of Wealth and the Marginal Propensity to Consume," *Quantitative Economics*, 8, 977–1020, At http://econ.jhu.edu/people/ccarroll/papers/cstwMPC.

Carroll, Christopher D., Martin Sommer, and Jiri Slacalek (2011): "International Evidence on Sticky Consumption Growth," *Review of Economics and Statistics*, 93(4), 1135–1145, http://econ.jhu.edu/people/ccarroll/papers/cssIntlStickyC/.

Carroll, Christopher D, Matthew N White, and Team Econ-ARK (2017): "econ-ark/HARK: 0.8.0," Available at via doi:10.5281/zenodo.1001068 or at https://doi.org/10.5281/zenodo.1001068.

Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan (2005): "A Critique of Structural VARs Using Real Business Cycle Theory," working paper 631, Federal Reserve Bank of Minneapolis.

Chetty, Raj, and Adam Szeidl (2016): "Consumption Commitments and Habit Formation," *Econometrica*, 84, 855–890.

Christiano, Laurence J., Martin Eichenbaum, and Charles L. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1), 1–45.

Cochrane, John H (1989): "The Sensitivity of Tests of the Intertemporal Allocation of Consumption to Near-Rational Alternatives," *American Economic Review*, 79(3), 319–337.

Coibion, Olivier, and Yuriy Gorodnichenko (2015): "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts," *American Economic Review*, 105(8), 2644–2678.

Constantinides, George M. (1990): "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy*, 98(3), 519–543.

Coulibaly, Brahima, and Geng Li (2006): "Do Homeowners Increase Consumption after the Last Mortgage Payment? An Alternative Test of the Permanent Income Hypothesis," *The Review of Economics and Statistics*, 88(1), 10–19.

Crawley, Edmund, and Andreas Kuchler (2018): "Consumption Heterogeneity: Micro Drivers and Macro Implications," working paper 129, Danmarks Nationalbank.

Dynan, Karen E. (2000): "Habit Formation in Consumer Preferences: Evidence from Panel Data," *American Economic Review*, 90(3), http://www.jstor.org/stable/117335.

Edge, Rochelle M, Thomas Laubach, and John C Williams (2007): "Learning and shifts in long-run productivity growth," *Journal of Monetary Economics*, 54(8), 2421–2438.

Fagereng, Andreas, Martin B. Holm, and Gisle J. Natvik (2017): "MPC Heterogeneity and Household Balance Sheets," discussion paper, Statistics Norway.

Fernald, John G., Robert Hall, James Stock, and Mark Watson (2017): "The Disappointing Recovery of Output after 2009," *Brookings Papers on Economic Activity*, Spring.

Friedman, Milton A. (1957): *A Theory of the Consumption Function*. Princeton University Press.

Fuhrer, Jeffrey C. (2000): "Habit Formation in Consumption and its Implications for Monetary Policy Models," *American Economic Review*, 90(3), 367–390, http://www.jstor.org/stable/117334.

——— (2017): "Expectations as a Source of Macroeconomic Persistence: Evidence from Survey Expectations in a Dynamic Macro Model," *Journal of Monetary Economics*, 86, 22–55.

——— (2018): "Intrinsic Expectations Persistence: Evidence from Professional and Household Survey Expectations," Working Papers 18-9, Federal Reserve Bank of Boston.

Gabaix, Xavier (2014): "A Sparsity-Based Model of Bounded Rationality," *The Quarterly Journal of Economics*, 129(4), 1661–1710.

Gruber, Joseph W. (2004): "A Present Value Test of Habits and the Current Account," *Journal of Monetary Economics*, 51(7), 1495–1507.

Guvenen, Fatih, and Anthony A. Smith (2014): "Inferring Labor Income Risk and Partial Insurance From Economic Choices," *Econometrica*, 82(6), 2085–2129.

Hall, Robert E. (1978): "Stochastic Implications of the Life-Cycle/Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 96, 971–87, Available at http://www.stanford.edu/~rehall/Stochastic-JPE-Dec-1978.pdf.

Havranek, Tomas, Marek Rusnak, and Anna Sokolova (2017): "Habit Formation in Consumption: A Meta-Analysis," *European Economic Review*, 95, 142–167.

Hsieh, Chang-Tai (2003): "Do consumers react to anticipated income changes? Evidence from the Alaska permanent fund," *American Economic Review*, 93(1), 397–405.

Jappelli, Tullio, and Luigi Pistaferri (2014): "Fiscal Policy and MPC Heterogeneity," *AEJ: Marcoeconomics*, 6(4), 107–36.

Jermann, Urban J. (1998): "Asset Pricing in Production Economies," *Journal of Monetary Economics*, 42(2), 257–75.

Johnson, David S., Jonathan A. Parker, and Nicholas S. Souleles (2006): "Household Expenditure and the Income Tax Rebates of 2001," *American Economic Review*, 96(5), 1589–1610.

Jorgenson, Dale W., Mun S. Ho, and Kevin J. Stiroh (2008): "A Retrospective Look at the U.S. Productivity Growth Resurgence," *Journal of Economic Perspectives*, 22(1), 3–24.

Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante (2018): "Monetary Policy According to HANK," *American Economic Review*, 108(3), 697–743, http://www.aeaweb.org/articles?id=10.1257/aer.20160042.

Kaplan, Greg, Gianluca Violante, and Justin Weidner (2014): "The Wealthy Hand-to-Mouth," *Brookings Papers on Economic Activity*, Spring, 77–138.

Karahan, Fatih, Sean Mihaljevich, and Laura Pilossoph (2017): "Understanding Permanent and Temporary Income Shocks," URL link retrieved on 03/02/2018 here.

Kennickell, Arthur (1995): "Saving and Permanent Income: Evidence from the 1992 SCF," mimeo, Board of Governors of the Federal Reserve System.

Krusell, Per, and Anthony A. Smith (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106(5), 867–896.

Kueng, Lorenz (2012): "Tax News: Identifying the Household Consumption Response to Tax Expectations Using Municipal Bond Prices," working paper, Northwestern University.

——— (2018): "Excess Sensitivity of High-Income Consumers," *The Quarterly Journal of Economics*, 133(4), 1693–1751.

Low, Hamish, Costas Meghir, and Luigi Pistaferri (2010): "Wage Risk and Employment Risk over the Life Cycle," *The American Economic Review*, 100(4), 1432–1467.

Lucas, Robert E. (1973): "Some International Evidence on Output-Inflation Tradeoffs," *American Economic Review*, 63, 326–334.

Ludvigson, Sydney, and Alexander Michaelides (2001): "Does Buffer Stock Saving Explain the Smoothness and Excess Sensitivity of Consumption?," *American Economic Review*, 91(3), 631–647.

Luo, Yulei (2008): "Consumption Dynamics under Information Processing Constraints," *Review of Economic Dynamics*, 11(2), 366–385.

Luo, Yulei, Jun Nie, Gaowang Wang, and Eric R. Young (2017): "Rational Inattention and the Dynamics of Consumption and Wealth in General Equilibrium," *Journal of Economic Theory*, (172), 55–87.

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

——— (2015): "Business Cycle Dynamics under Rational Inattention," *The Review of Economic Studies*, 82(4), 1502–1532.

Mankiw, N. Gregory, and Ricardo Reis (2002): "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*, 117(4), 1295–1328.

Morris, Stephen, and Hyun Song Shin (2006): "Inertia of Forward-Looking Expectations," *The American Economic Review*, 96(2), 152–157.

Muth, John F. (1960): "Optimal Properties of Exponentially Weighted Forecasts," *Journal of the American Statistical Association*, 55(290), 299–306.

Nielsen, Helena Skyt, and Annette Vissing-Jorgensen (2006): "The Impact of Labor Income Risk on Educational Choices: Estimates and Implied Risk Aversion," *Manuscript*.

Olafsson, Arna, and Michaela Pagel (2018): "The Liquid Hand-to-Mouth: Evidence from Personal Finance Management Software," *The Review of Financial Studies*, 31(11), 4398–4446.

Parker, Jonathan A. (2017): "Why Don't Households Smooth Consumption? Evidence from a $25 Million Experiment," *American Economic Journal: Macroeconomics*, 4(9), 153–183.

Parker, Jonathan A, Nicholas S Souleles, David S Johnson, and Robert McClelland (2013): "Consumer Spending and the Economic Stimulus Payments of 2008," *The American Economic Review*, 103(6), 2530–2553.

Pischke, Jörn-Steffen (1995): "Individual Income, Incomplete Information, and Aggregate Consumption," *Econometrica*, 63(4), 805–40.

Pistaferri, Luigi (2001): "Superior Information, Income Shocks, And The Permanent Income Hypothesis," *The Review of Economics and Statistics*, 83(3), 465–476.

Reis, Ricardo (2006): "Inattentive Consumers," *Journal of Monetary Economics*, 53(8), 1761–1800.

Sims, Christopher (2003): "Implications of Rational Inattention," *Journal of Monetary Economics*, 50(3), 665–690, available at http://ideas.repec.org/a/eee/moneco/v50y2003i3p665-690.html.

Sims, Christopher A. (2006): "Rational Inattention: Beyond the Linear–Quadratic Case," *American Economic Review*, 96(2), 158–163.

Sommer, Martin (2007): "Habit Formation and Aggregate Consumption Dynamics," *Advances in Macroeconomics*, 7(1), Article 21.

Souleles, Nicholas S. (1999): "The Response of Household Consumption to Income Tax Refunds," *American Economic Review*, 89(4), 947–958.

Staiger, Douglas, James H. Stock, and Mark W. Watson (2001): "Prices Wages and the US NAIRU in the 1990s," in *The Roaring Nineties: Can Full Employment Be Sustained?*, ed. by Alan B. Krueger, and Robert Solow. The Russell Sage Foundation and Century Press, New York.

Storesletten, Kjetil, Chris I. Telmer, and Amir Yaron (2004): "Consumption and Risk Sharing Over the Life Cycle," *Journal of Monetary Economics*, 51(3), 609–633.

Tutino, Antonella (2013): "Rationally Inattentive Consumption Choices," *Review of Economic Dynamics*, 16(3), 421–439.

Weber, Christian E. (2002): "Intertemporal Non-Separability and 'Rule-Of-Thumb' Consumption," *Journal of Monetary Economics*, 49, 293–308.

Wilcox, David W. (1992): "The Construction of U.S. Consumption Data: Some Facts and Their Implications for Empirical Work," *American Economic Review*, 82(4), 922–941.

Woodford, Michael (2002): "Imperfect Common Knowledge and the Effects of Monetary Policy," in *Knowledge, Information and Expectations in Modern Macroeconomics*, ed. by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford. Princeton University Press, Princeton.

Working, Holbrook (1960): "Note on the Correlation of First Differences of Averages in a Random Chain," *Econometrica*, 28(4), 916–918.

Zeldes, Stephen P. (1989a): "Consumption and Liquidity Constraints: An Empirical Investigation," *Journal of Political Economy*, 97, 305–46, Available at http://www.jstor.org/stable/1831315.

——— (1989b): "Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence," *Quarterly Journal of Economics*, 104(2), 275–298.

A Calibration

This appendix presents more complete details and justification for the calibrated parameters in Table 1. We begin by calibrating market-level and preference parameters by standard methods, then specify additional parameters to characterize the idiosyncratic income shock distribution.

A.1 Macroeconomic Calibration

We assume a coefficient of relative risk aversion of 2. The quarterly depreciation rate δ is calibrated by assuming annual depreciation of 6 percent, i.e., $(1-\delta)^4 = 0.94$. Capital's share in aggregate output takes its usual value of $\gamma = 0.36$.

We set the variances of the quarterly transitory and permanent shocks at the approximate values respectively:

$$\begin{aligned} \sigma_\Theta^2 &= 0.00001, \\ \sigma_\Psi^2 &= 0.00004, \end{aligned}$$

which allow the model to match high degree of persistence in aggregate labor income.[47] These values are consistent with papers such as Jermann (1998), Boldrin, Christiano, and Fisher (2001), and Chari, Kehoe, and McGrattan (2005), considered standard in the RBC literature. These authors model the state of technology as either a highly persistent AR(1) process or a random walk; but the underlying calibrations come from the autocorrelation properties of measured aggregate dynamics, which are matched about as well by our specification of the income process.

To finish the calibration, we consider a simple perfect foresight model (PF-DSGE), with all aggregate and idiosyncratic shocks turned off. We set the perfect foresight steady state aggregate capital-to-output ratio to 12 on a quarterly basis (corresponding to the usual ratio of 3 for capital divided by annual income). Along with the calibrated values of γ and δ, this choice implies values for the other steady-state characteristics of the PF-DSGE model:

$$\begin{aligned} K &= 12^{1/(1-\gamma)}, \\ \mathsf{W} &= (1-\gamma)K^\gamma, \\ \mathcal{R} &= (1-\delta) + \gamma K^{\gamma-1}. \end{aligned}$$

In the SOE model, we fix the interest factor $\mathcal{R}$ and wage rate W to these PF-DSGE steady state values.

A perfect foresight representative agent would achieve this steady state if his discount factor satisfied $\mathcal{R}\beta = 1$. For the SOE model, however, we choose a much lower value of β (0.97), resulting in agents with wealth holdings around the median observed in the

[47] We measure labor income using U.S. NIPA data as wages and salaries plus transfers minus personal contributions for social insurance.

data;[48] the value of β satisfying $\mathcal{R}\beta = 1$ is used in the closed economy models presented in the online appendix, allowing those models to fit the *mean* observed wealth.

A.2 Calibration of Idiosyncratic Shocks

The annual-rate idiosyncratic transitory and permanent shocks are assumed to be:

$$\begin{aligned}\sigma_\theta^2 &= 0.03,\\ \sigma_\psi^2 &= 0.012.\end{aligned}$$

Our calibration for the sizes of the idiosyncratic shocks are conservative relative to the literature;[49] using data from the *Panel Study of Income Dynamics*, for example, Carroll and Samwick (1997) estimate $\sigma_\psi^2 = 0.0217$ and $\sigma_\theta^2 = 0.0440$; Storesletten, Telmer, and Yaron (2004) estimate $\sigma_\psi^2 \approx 0.017$, with varying estimates of the transitory component. But recent work by Low, Meghir, and Pistaferri (2010) suggests that controlling for participation decisions reduces estimates of the permanent variance somewhat; and using very well-measured Danish administrative data, Nielsen and Vissing-Jorgensen (2006) estimate $\sigma_\psi^2 \approx 0.005$ and $\sigma_\theta^2 \approx 0.015$, which presumably constitute lower bounds for plausible values for the truth in the U.S. (given the comparative generosity of the Danish welfare state).

We assume that the probability of unemployment is 5 percent per quarter. This approximates the historical mean unemployment rate in the U.S., but model unemployment differs from real unemployment in (at least) two important ways. First, the model does not incorporate unemployment insurance, so labor income of the unemployed is zero. Second, model unemployment shocks last only one quarter, so their duration is shorter than the typical U.S. unemployment spell (about 6 months). The idea of the calibration is that a single quarter of unemployment with zero benefits is roughly as bad as two quarters of unemployment with an unemployment insurance payment of half of permanent labor income (a reasonable approximation to the typical situation facing unemployed workers). The model could be modified to permit a more realistic treatment of unemployment spells; this is a promising topic for future research, but would involve a considerable increase in model complexity because realism would require adding the individual's employment situation as a state variable.

The probability of mortality is set at $\mathsf{D} = 0.005$, which implies an expected working life of 50 years; results are not sensitive to plausible alternative values of this parameter, so long as the life length is short enough to permit a stationary distribution of idiosyncratic permanent income.

[48]The exact value of the median is depends in part on whether housing equity should be viewed as part of the precautionary buffer stock, the age range of the households being matched, the measure of permanent income, and many other extraneous issues.

[49]See Table 1 in the ECB working paper version of Carroll, Slacalek, and Tokuoka (2015) for a comprehensive overview of estimates of variances of idiosyncratic income shocks; Carroll, Christopher D., Jiri Slacalek, and Kiichi Tokuoka (2014): "Buffer-Stock Saving in a Krusell–Smith World," working paper 1633, European Central Bank, `https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1633.pdf`.

B Heterogeneous Agents Dynamic Stochastic General Equilibrium (HA-DSGE) Model

Our HA-DSGE model relaxes the simplifying assumption in the SOE model of a frictionless global capital market. In this closed economy, factor prices W_t and r_t are determined in the usual way from the aggregate production function and aggregate state variables, including the stochastic aggregate shocks, putting the model in the (small, but rapidly growing) class of heterogeneous agent DSGE models.

For the HA-DSGE model, we set the discount factor to $\beta = \mathcal{R}^{-1} = 0.986$, roughly matching the target capital-to-output ratio.[50] Households in the HA-DSGE model thus hold significantly more wealth than their counterparts in the baseline SOE model, who were calibrated to approximate the *median* observed wealth-to-income ratio. This reflects our goal of presenting results that span the full range of calibrations in the micro and macro literatures; the micro literature has often focused on trying to explain the wealth holdings of the median household, which are much smaller than average wealth holdings. Experimentation has indicated that our results are not sensitive to such choices.

B.1 Model and Solution

We make the standard assumption that markets are competitive, and so factor prices are the marginal product of (effective) labor and capital respectively. Denoting capital's share as γ, so that $\mathbf{Y}_t = \mathbf{K}_t^{\gamma}\mathbf{L}_t^{1-\gamma}$, this yields the usual wage and interest rates:

$$\begin{aligned} \mathsf{W}_t = \frac{\partial \mathbf{Y}_t}{\partial \mathbf{L}_t} &= (1-\gamma)(\mathbf{K}_t/\mathbf{L}_t)^{-\gamma}, \\ \mathsf{r}_t = \frac{\partial \mathbf{Y}_t}{\partial \mathbf{K}_t} &= \gamma(\underbrace{\mathbf{K}_t/\mathbf{L}_t}_{=K_t/\Theta_t})^{\gamma-1}. \end{aligned} \tag{21}$$

Net of depreciation, the return factor on capital is $\mathcal{R}_t = 1 - \delta + \mathsf{r}_t$.

An agent's relevant state variables at the time of the consumption decision include the levels of household and aggregate market resources $(\mathbf{m}_{t,i}, \mathbf{M}_t)$, as well as household and aggregate labor productivity $(p_{t,i}, P_t)$ and the aggregate growth rate Φ_t. We assume that agents correctly understand the operation of the economy, including the production and shock processes, and have *beliefs* about aggregate saving—how aggregate market resources $\mathbf{M}_t$ become aggregate assets $\mathbf{A}_t$ (equivalently, next period's aggregate capital $\mathbf{K}_{t+1}$). Following Krusell and Smith (1998) and Carroll, Slacalek, Tokuoka, and White (2017), we assume that households believe that the aggregate saving rule is linear in logs, conditional on the current aggregate growth rate:

$$\mathbb{E}[\mathbf{A}_t] = \aleph(\mathbf{M}_t, \Phi_t = \Phi_j) \equiv \exp\big(\kappa_{j,0} + \kappa_{j,1}\log(\mathbf{M}_t)\big). \tag{22}$$

[50] The addition of aggregate and idiosyncratic risk implies that the capital-to-output ratio will be higher in the HA-DSGE model than the PF-DSGE calibration. Moreover, the Markov growth process will move aggregate capital holdings away from the calibrated target.

The growth-rate-conditional parameters $\kappa_{j,0}$ and $\kappa_{j,1}$ are exogenous to the individual's (partial equilibrium) optimization problem, but are endogenous to the general equilibrium of the economy. Taking the aggregate saving rule $\aleph$ as given, the household's problem can be written in Bellman form as:[51]

$$\mathbf{v}(\mathbf{m}_{t,i}, \mathbf{M}_t, p_{t,i}, P_t, \Phi_t) \;=\; \max_{\mathbf{c}_{t,i}} \left\{ \mathrm{u}(\mathbf{c}_{t,i}) + \beta\, \mathbb{E}\left[(1-\mathsf{d}_{t,i})\mathbf{v}(\mathbf{m}_{t+1,i}, \mathbf{M}_{t+1}, p_{t+1,i}, P_{t+1}, \Phi_{t+1})\right]\right\}. \tag{23}$$

As in the SOE model, the household's problem can be normalized by the combined productivity level $\boldsymbol{p}_{t,i}$, reducing the state space by two continuous dimensions. Dividing (23) by $\boldsymbol{p}_{t,i}^{1-\rho}$ and substituting normalized variables, the reduced problem is:

$$\begin{aligned}
\mathrm{v}(m_{t,i}, M_t, \Phi_t) &= \max_{c_{t,i}} \left\{ \mathrm{u}(c_{t,i}) + \beta(1-\mathsf{D})\, \mathbb{E}\left[(\Phi_{t+1}\boldsymbol{\psi}_{t+1,i})^{1-\rho}\mathrm{v}(m_{t+1,i}, M_{t+1}, \Phi_{t+1})\right]\right\} \\
& \text{s.t.} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (24) \\
a_{t,i} &= m_{t,i} - c_{t,i}, \\
k_{t+1,i} &= a_{t,i}/(1-\mathsf{D}), \\
m_{t+1,i} &= \mathcal{R}_{t+1} k_{t+1,i}/(\Phi_{t+1}\boldsymbol{\psi}_{t+1,i}) + \boldsymbol{\theta}_{t+1,i}\mathsf{W}_{t+1}.
\end{aligned}$$

Because household beliefs about the aggregate saving rule are linear in logs, (22) holds with normalized market resources and aggregate assets as well as in levels.

The equilibrium of the HA-DSGE model is characterized by a (normalized) consumption function $\mathrm{c}(m, M, \Phi)$ and an aggregate saving rule $\aleph$ such that when all households believe $\aleph$, the solution to their individual problem (24) is c; and when all agents act according to c, the best log-linear fit of A_t on M_t (conditional on Φ_t) is $\aleph$. The model is solved using a method similar to Krusell and Smith (1998).[52]

B.2 Frictionless vs Sticky Expectations

The treatment of sticky beliefs in the HA-DSGE model is the natural extension of what we did in the SOE model presented in section 4.6: Because the level of $\mathbf{M}_t$ now affects future wages and interest rates, a consumer's perceptions of that variable $\widetilde{M}_{t,i} = \mathbf{M}_t/\widetilde{P}_{t,i}$ now matter. As households in our model do not necessarily observe the true aggregate productivity level, their perception of normalized aggregate market resources is

$$\widetilde{M}_{t,i} = \mathbf{M}_t/\widetilde{P}_{t,i} = \left(P_t/\widetilde{P}_{t,i}\right) M_t.$$

Households in the DSGE model choose their level of consumption using their perception of their normalized state variables:

$$\mathbf{c}_{t,i} = \widetilde{\boldsymbol{p}}_{t,i}\mathrm{c}(\widetilde{m}_{t,i}, \widetilde{M}_{t,i}, \widetilde{\Phi}_{t,i}) = \mathbf{c}(\mathbf{m}_{t,i}, \mathbf{M}_t, p_{t,i}, \widetilde{P}_{t,i}, \widetilde{\Phi}_{t,i}).$$

Households who misperceive the aggregate productivity state will incorrectly predict

[51] Subject to definitions (3), (4), (5), (7), (8), (9), (10), (11), (12), (21) and (22).

[52] Details are in Appendix D.1.2.

aggregate saving at the end of the period, and thus aggregate capital and the distribution of factor prices next period.[53]

Because households who misperceive the aggregate productivity state will make (slightly) different consumption–saving decisions than they would have if fully informed, *aggregate* saving behavior will be different under sticky than under frictionless expectations. Consequently, the equilibrium aggregate saving rule $\aleph$ will be slightly different under sticky vs frictionless expectations. When the HA-DSGE model is solved under sticky expectations, we implicitly assume that all households understand that all other households also have sticky expectations, and the equilibrium aggregate saving rule is the one that emerges from this belief structure.

B.3 Results

We report some of the equilibrium characteristics of the SOE and HA-DSGE models in Table 5, to highlight their qualitatively similar patterns. The table suggests a broad generalization that we have confirmed with extensive experimentation: With respect to either cross section statistics, mean outcomes, or idiosyncratic consumption dynamics, the frictionless expectations and sticky expectations models are virtually indistinguishable using microeconomic data, and very similar in most aggregate implications aside from the dynamics of aggregate consumption.

Table 6 reports the results of estimating regression (17) on data generated from the HA-DSGE model. The results are substantially the same as the previous analysis for the SOE model (in Table 3).[54]

The model with frictionless expectations (top panel) implies aggregate consumption growth that is moderately (but not statistically significantly) serially correlated when examined in isolation (second row), but the effect "washes out" when expected income growth and the aggregate wealth to income ratio are included in the horse race regression (fourth row). As expected in a closed economy model, the aggregate wealth-to-income ratio A_t is negatively correlated with consumption growth, but its predictive power is so slight that it is statistically insignificant in samples of only 200 quarters.

The model with sticky expectations (bottom panel) again implies a serial correlation coefficient of consumption growth not far from 0.75 in the univariate IV regression (second row). As in the SOE simulation, the horserace regression (fifth row) indicates that the apparent success of the Campbell–Mankiw specification (third row) reflects the correlation of predicted current income growth with instrumented lagged consumption growth.

[53] This incorrect prediction is short-lived: all households will learn the true *levels* of next period's aggregate capital and output.

[54] Essentially similar results are obtained if we assume that households have heterogeneous discount factors, in the style of Carroll, Slacalek, Tokuoka, and White (2017). Using a calibration of the distribution of β that approximately matches the distribution of net worth in the U.S., the results presented in Table 6 are effectively unchanged (table available upon request). The main results hold whether β is chosen to match aggregate asset holdings, the wealth of the median household, or the entire distribution of wealth; it is not sensitive to the particular calibration of the model.

C Representative Agent (RA) Model

This appendix presents a representative agent model for analyzing the consequences of sticky expectations in a DSGE framework while abstracting from idiosyncratic income shocks and the death (and replacement) of households. It builds upon the modeling assumptions in section 4 to formulate the representative agent model, then presents simulated results analogous to section 5. The primary advantage of this model is that it allows fast analysis of sticky expectations in a closed economy, yielding very similar results to the heterogeneous agents DSGE model with less than a minute of computation, rather than a few hours. However, the model is not truly a "representative agent" model under sticky expectations; instead it is as though there is an agent whose beliefs about the aggregate state are "smeared" over the state space with a probability distribution that reflects the distribution of perceptual delay implied by the Calvo updating probability. That is, the ealized level of consumption represents the weighted average level of consumption chosen by the "many minds" of the representative household, with weights reflecting the likelihood of each possible degree of perceptual delay.

C.1 Model and Solution

The representative agent's state variables at the time of its consumption decision are the level of market resources $\mathbf{M}_t$, the productivity of labor P_t, and the growth rate of productivity Φ_t. Idiosyncratic productivity shocks ψ and θ do not exist, and the possibility of death is irrelevant; aggregate permanent and transitory productivity shocks Ψ and Θ are distributed as usual.

The representative agent's problem can be written in Bellman form as:[55]

$$\begin{aligned}\mathbf{V}(\mathbf{M}_t, P_t, \Phi_t) &= \max_{\mathbf{C}_t} \left\{ \mathrm{u}(\mathbf{C}_t) + \mathbb{E}\left[\mathbf{V}(\mathbf{M}_{t+1}, P_{t+1}, \Phi_{t+1})\right]\right\} \\ &\text{s.t.} \\ \mathbf{A}_t &= \mathbf{M}_t - \mathbf{C}_t.\end{aligned}$$

Normalizing the representative agent's problem by the productivity level P_t as in the SOE and HA-DSGE models, the problem's state space can be reduced to:[56]

$$\begin{aligned}\mathrm{V}(M_t, \Phi_t) &= \max_{C_t} \left\{ \mathrm{u}(C_t) + \beta\, \mathbb{E}_t \left[(\Phi_{t+1}\Psi_{t+1})^{1-\rho}\mathrm{V}(M_{t+1}, \Phi_{t+1})\right]\right\} \\ &\text{s.t.} \\ A_t &= M_t - C_t.\end{aligned} \tag{25}$$

Noting that the return to (normalized) end-of-period assets for next period's market resources is $\frac{\mathrm{d}M_{t+1}}{\mathrm{d}A_t} = \mathcal{R}_{t+1}/(\Phi_{t+1}\Psi_{t+1})$, (25) has a single first-order condition that is sufficient to characterize the solution to the normalized problem:

$$C_t^{-\rho} - \underbrace{\beta\, \mathbb{E}\left[\mathcal{R}_{t+1}(\Phi_{t+1}\Psi_{t+1})^{-\rho}\mathrm{V}^M(A_t\mathcal{R}_{t+1}/(\Psi_{t+1}\Phi_{t+1}) + \Theta_{t+1}\mathsf{W}_{t+1}, \Phi_{t+1})\right]}_{\equiv \mathfrak{V}^A(A_t, \Phi_t)} = 0 \tag{26}$$

[55]Subject to definitions (3), (10), (11), (12) and (21).

[56]Subject to definitions (3), (13) and (21).

$$\Longrightarrow C_t = \mathfrak{V}^A(A_t, \Phi_t)^{-1/\rho}.$$

The representative agent model can be solved using the endogenous grid method, following the same procedure as for the SOE model described in Appendix D.1, yielding normalized consumption function $\mathrm{C}(M, \Phi)$.[57]

C.2 Frictionless vs Sticky Expectations

The typical interpretation of a representative agent model is that it represents a continuum of households that face no idiosyncratic shocks, and thus all find themselves with the same state variables; idiosyncratic decisions are equivalent to aggregate, representative agent decisions. Once we introduce sticky expectations of aggregate productivity, this no longer holds: different households will have different perceptions of productivity, and thus make different consumption decisions.

To handle this departure from the usual representative agent framework, we take a "multiple minds" or quasi-representative agent approach. That is, we model the representative agent as being made up of a continuum of households who all correctly perceive the level of aggregate market resources $\mathbf{M}_t$, but have different perceptions of the aggregate productivity state. Each household chooses their level of consumption based on their perception of the productivity state; the realized level of aggregate consumption is simply the sum across all households.

Formally, we track the *distribution* of perceptions about the aggregate productivity state as a stochastic vector φ_t over the current growth rate $\Phi_t \in \{\Phi\}$, representing the fraction of households who perceive each value of Φ, and a vector $\widetilde{P}_t$ representing the average perceived productivity level among households who perceive each Φ. As in our other models, agents update their perception of the true aggregate productivity state (P_t, Φ_t) with probability Π; likewise, the distinction between frictionless and sticky expectations is simply whether $\Pi = 1$ or $\Pi < 1$.

Defining e^j_N as the N-length vector with zeros in all elements but the j-th, which has a one, the distribution of population perceptions of growth rate Φ_t evolves according to:

$$\varphi_{t+1} = (1 - \Pi)\varphi_t + \Pi e^j_N \text{ when } \Phi_{t+1} = \Phi_j. \tag{27}$$

That is, a Π proportion of households who perceive each growth rate update their perception to the true state $\Phi_{t+1} = \Phi_j$, while the other $(1 - \Pi)$ proportion of households maintain their prior belief (which might already be Φ_j).

The vector of average perceptions of aggregate productivity for each growth rate can then be calculated as:

$$\widetilde{P}_{t+1} = \big((1 - \Pi)\varphi_t \odot \widetilde{P}_t + \Pi e^j_N P_{t+1}\big) \oslash \varphi_{t+1}. \tag{28}$$

That is, the average perception of productivity in each growth state is the weighted average of updaters and non-updaters who perceive that growth rate.[58]

[57] The only differences in solution method are that the RA model uses $N_\Psi = N_\Theta = 7$ point approximations to the aggregate shock distribution, expected marginal value of assets is calculated using (26), and the upper bound of $\mathbb{A}$ is 120.

[58] The Hadamard operators $\odot$ and $\oslash$ represent element-wise multiplication and division, respectively. As a numeric detail, $\widetilde{P}^j_t$ is reset to 1 when $\varphi^j_t = 0$, which would otherwise cause it to be undefined. When no households perceive

Households who perceive each growth rate Φ choose their level of consumption according to their perception of normalized market resources, as though they knew their perception to be the truth. Defining $\widetilde{M}_t^j = \mathbf{M}_t/\widetilde{P}_t^j$ as perceived normalized market resources for households who perceive the aggregate growth rate is Φ_j, aggregate consumption is:

$$\mathbf{C}_t = \sum_{\Phi_j \in \{\Phi\}} \widetilde{P}_t^j \mathrm{C}(\widetilde{M}_t^j, \Phi_j)\varphi_t^j. \tag{29}$$

This represents the weighted average of per-state consumption levels of the partial representative agents.

When the representative agent frictionlessly updates its information every period ($\Pi = 1$), equations (27) and (28) say that $\varphi_t = e_N^j$ and $\widetilde{P}_t^j = P_t$ (with irrelevant values in the other vector elements), so that the representative agent is truly representative. When expectations are sticky ($\Pi < 1$), the representative agent's perceptions of the growth rate become "smeared" across its past realizations; its perceptions the productivity level likewise deviate from the true value, even for the part of the representative agent who perceives the true growth rate.[59]

C.3 Simulation Results

We calibrate the RA model using the same parameters as for the HA-DSGE model (see section A.1, Table 1, and Appendix B.3), except that there are no idiosyncratic income shocks ($\sigma_\psi^2 = \sigma_\theta^2 = \wp = 0$) and the possibility of death is irrelevant ($\mathsf{D} = 0$). After solving the model, we utilize the same simulation procedure described in section 5, taking 100 samples of 200 quarters each; average coefficients and standard errors across the samples are reported in Table 7.

The upper panel of Table 7 shows that under frictionless expectations, consumption growth in the representative agent model cannot be predicted to any statistically significant degree under any specification. The lower panel, under sticky expectations, yields results that are strikingly similar to the SOE model in Table 3. Both (instrumented) lagged consumption growth and expected income growth are significant predictors of aggregate consumption growth, but the 'horse race' regression reveals that the predictability is dominated by serially correlated consumption growth, confirming the results of the two heterogeneous agents models.

growth rate Φ_j, the average perception of productivity does not exist for this state and is quantitatively irrelevant, but must exist for (28) to not fail in the next period.

[59] An alternative method for modeling sticky expectations with a representative agent would be to track the perceptions of the segments of households who last updated $n = \{0, 1, \ldots, 200\}$ periods ago, compute consumption for each segment, and take the weighted average across the segments to yield aggregate consumption. This approach would only be slightly more complicated to implement, and we believe it would yield quantitatively similar results.

D Numerical Methods

D.1 Solution Methods

D.1.1 Small Open Economy Solution Details

Consider the household's normalized problem in the SOE model, given in (14). Substituting the latter two constraints into the maximand, this problem has one first order condition (with respect to $c_{t,i}$), which is sufficient to characterize the solution:

$$c_{t,i}^{-\rho} - \underbrace{\mathsf{R}(1-\mathsf{D})\beta\,\mathbb{E}_t\left[(\Phi_{t+1}\boldsymbol{\psi}_{t+1,i})^{-\rho}\mathrm{v}^m\big(\mathsf{R}/(\Phi_{t+1}\boldsymbol{\psi}_{t+1,i})a_{t,i} + \mathsf{W}\boldsymbol{\theta}_{t+1,i}, \Phi_{t+1}\big)\right]}_{\equiv \mathfrak{v}^a(a_{t,i},\Phi_t)} = 0 \qquad (30)$$

$$\implies c_{t,i} = \mathfrak{v}^a(a_{t,i}, \Phi_t)^{-1/\rho}.$$

We use the endogenous grid method to solve the model by iterating on the first order condition. Eliding some uninteresting complications, our procedure is straightforward:

1. Construct discrete approximations to the lognormal distributions of θ, Θ, ψ, and Ψ, adjusting for the point mass at 0 for θ with probability $\wp$. We use equiprobable $N_\psi = N_\theta = 7$ point approximations for the (lognormal portion of) the idiosyncratic shocks and $N_\Psi = N_\Theta = 5$ point approximations for the aggregate shocks.

2. Choose an exogenous grid of end-of-period normalized assets-above-natural-borrowing-constraint $\mathbb{A} = \{\blacktriangle a_j\}_{j=1}^{N_a}$, spanning the range values that an agent might reasonably encounter in a simulated lifetime. We use a triple-exponential grid spanning $\blacktriangle a \in [10^{-5}, 40]$ with $N_a = 48$ gridpoints. The natural borrowing constraint is zero because of the possibility of $\theta = 0$, so assets-above-natural-borrowing-constraint is simply assets a.

3. Initialize the guess of the consumption function to $\mathrm{c}(m, \cdot) = m$, the solution for an agent who has no future.

4. Define the marginal value function $\mathrm{v}^m(\cdot)$ as $\mathrm{u}'(\mathrm{c}(\cdot))$, as determined by the standard envelope condition.

5. Use the discrete approximations to the shock processes and the Markov transition matrix Ξ to compute $\mathfrak{v}^a(a_j, \Phi_k)$ for all $(a_j, \Phi_k) \in \mathbb{A} \times \{\Phi\}$.

6. Use (30) to find the level of consumption that would make ending the period with a_j in assets optimal (when aggregate growth is Φ_k): $c_{j,k} = \mathfrak{v}^a(a_j, \Phi_k)^{-1/\rho}$.

7. Calculate beginning of period market resources $m_{j,k} = a_{j,k} + c_{j,k}$ for all j, k.

8. For each k, construct $\mathrm{c}(m, \Phi_k)$ by linearly interpolating $c_{j,k}$ over $m_{j,k}$, with an additional point at $(m = 0, c = 0)$.

9. Calculate the supnorm distance between the newly constructed c and the previous guess, evaluated at the $N_a \times ||\{\Phi\}||$ gridpoints. If the distance is less than $\epsilon = 10^{-6}$, STOP; else go to step 4.

The numerically computed consumption function can then be used to simulate a population of households, as described in Appendix D.2.

D.1.2 Dynamic Stochastic General Equilibrium Solution Details

Consider the household's normalized problem in the HA-DSGE model, given in (24). Recalling that we are taking the aggregate saving rule $\aleph$ as given, optimal consumption is characterized by the solution to the first-order condition:

$$c_{t,i}^{-\rho} - \underbrace{\beta\,\mathbb{E}\left[\mathcal{R}_{t+1}(\Phi_{t+1}\boldsymbol{\psi}_{t+1,i})^{-\rho}\mathrm{v}\big(\mathcal{R}_t a_{t,i}/((1-\mathsf{D})\Phi_{t+1}\boldsymbol{\psi}_{t+1,i}) + \boldsymbol{\theta}_{t+1,i}\mathsf{W}_{t+1}, M_{t+1}, \Phi_{t+1}\big)\right]}_{\equiv \mathfrak{v}^a(a_{t,i}, M_t, \Phi_t)} = 0 \tag{31}$$

$$\Longrightarrow c_{t,i} = \mathfrak{v}^a(a_{t,i}, M_t, \Phi_t)^{-1/\rho}.$$

Solving the HA-DSGE model requires a nested loop procedure in the style of Krusell and Smith (1998), as the equilibrium of the model is a fixed point in the space of household beliefs about the aggregate saving rule. For the outer loop, searching for the equilibrium $\aleph$, we use the following procedure:

1. Construct a grid of (normalized) aggregate market resources $\mathbb{M} = \{M_j\}_{j=1}^{N_M}$. We use a $N_M = 19$ point grid centered around the steady state of the perfect foresight DSGE model $\breve{M}$, spanning the range $M \in [0.1\breve{M}, 5\breve{M}]$.

2. For each $\Phi_k \in \{\Phi\}$, initialize the aggregate saving rule to arbitrary values. We use $\kappa_{k,0} = 0$ and $\kappa_{k,1} = 1$; there exist more efficient initial guesses.

3. In the inner loop, solve the household's optimization problem for the current guess of $\aleph$, using the procedure described below.

4. Simulate many households for many periods, using the procedure described in Appendix D.2, yielding a long *history* of aggregate market resources, productivity growth, and assets $\mathfrak{H} = \{(M_t, \Phi_t, A_t)\}_{t=0}^{T}$.

5. For each k, define $\mathfrak{H}_k \equiv \{\mathfrak{H} | \Phi_t = \Phi_k\}$. Regress A_t on M_t on the set $\mathfrak{H}_k$, yielding coefficients that provide updated values of $\kappa_{k,0}$ and $\kappa_{k,1}$ for $\aleph$.

6. Calculate the supnorm distance between the new and previous values of aggregate saving rule coefficients κ. If it is less than $\grave{\epsilon} = 10^{-4}$, STOP; else go to step 3.

The inner solution loop (step 3) proceeds very similarly to the SOE solution method above, with differences in the following steps:

2. The set $\mathbb{A}$ spans $[10^{-5}, 120]$ because of the higher β in the HA-DSGE model.

5. End-of-period marginal value of assets is calculated as $\mathfrak{v}^a(a_j, M_k, \Phi_\ell)$ for all $(a_j, M_k, \Phi_\ell) \in \mathbb{A} \times \mathbb{M} \times \{\Phi\}$.

6. Use (31) to calculate $c_{j,k,\ell} = \mathfrak{v}^a(a_j, M_k, \Phi_\ell)^{-1/\rho}$.

8. For each ℓ, construct $\mathrm{c}(m, M, \Phi_\ell)$ by linearly interpolating $c_{j,k,\ell}$ over $m_{j,k,\ell}$ for each k, then interpolating the linear interpolations over $\mathbb{M}$.

D.2 Simulation Procedures

This appendix describes the procedure for generating a *history* of simulated outcomes once the household's optimization problem has been solved to yield consumption function $\mathrm{c}(\cdot)$ (or $\mathrm{C}(\cdot)$ in the representative agent model). We first describe the procedure for the SOE and HA-DSGE models, then summarize the simulation method for the representative agent model of Appendix C.

In any given period t, there are exactly $I = 20,000$ households in the simulated population. At the very beginning of the simulation, all households are given an initial level of capital: $k_{t,i} = 0$ in the SOE model (as if they were newborns) and $k_{t,i} = \breve{K}$ in the HA-DSGE model. Likewise, normalized aggregate capital is set to the perfect foresight steady state $K_t = \breve{K}$. At the beginning of time, all households have $p_{t,i} = 1$ and correct perceptions of the aggregate state. We initialize $P_t = 1$ and $\Phi_t = 1$, average growth.

Time begins in period $t = -1000$, but the reported *history* begins at $t = 0$ following a 1000 period "burn in" phase to allow the population distribution of $p_{t,i}$ and $a_{t,i}$ to reach its long run distribution. In each simulated period t, we execute the following steps:

1. Draw aggregate shocks Θ_t and Ψ_t and productivity growth Φ_t, then calculate the new level of aggregate permanent productivity P_t and factor returns W_t and R_t using (21) (HA-DSGE model) or assigning the constant global values (SOE).

2. Randomly select $\mathsf{D}I = 100$ household indices i to die and be replaced: $\mathsf{d}_{i,t} = 1$. Newborns get $p_{t,i} = 1$, $k_{t,i} = 0$, and a correct perception of the aggregate state. Survivors receive the capital of the dead via the Blanchardian scheme.

3. Randomly select ΠI household indices to update their aggregate information: $\pi_{t,i} = 1$. Agents' perceptions $(\widetilde{P}_{t,i}, \widetilde{\Phi}_{t,i})$ are set according to (15).

4. The economy produces output. All agents draw idiosyncratic shocks $\psi_{t,i}$ and $\theta_{t,i}$, with newborns automatically drawing $\psi_{t,i} = \theta_{t,i} = 1$,[60] then observe their true $\mathbf{m}_{t,i}$ (and $\mathbf{M}_t$ in the HA-DSGE model).

5. Agents compute their perception of normalized idiosyncratic market resources $\widetilde{m}_{t,i}$ (and aggregate $\widetilde{M}_{t,i}$ in HA-DSGE).

[60] This prevents newborns from being unemployed in their first period of life and thus getting $\mathbf{c}_{t,i} = 0$. It also simplifies the calculation of the cost of stickiness.

6. Agents choose their level of consumption $\mathbf{c}_{t,i}$ according to their consumption function and their perceived state, and end the period with $\mathbf{a}_{t,i} = \mathbf{m}_{t,i} - \mathbf{c}_{t,i}$ in assets.

7. Aggregate assets $\mathbf{A}_t$ and consumption $\mathbf{C}_t$ are calculated by taking population averages across the I households. This period's assets become next period's aggregate capital $\mathbf{K}_{t+1}$, and the next period begins.

We simulate a total of about 21,000 periods, so that the final period is indexed by $t = T = 20,000$. The time series values reported in Table 5 are calculated on the span of the history, $t = 0$ to $t = T$; the cross sectional values in this table are averaged across all within-period cross sections. The time series regressions in Tables 3 and 6 partition the history into 200 samples of 100 quarters each; the tables report average coefficients and statistics across 100 sample regressions.

When simulating the representative agent model of Appendix C, only a few changes are necessary to the procedure above. The vectors of perceptions are initialized to $\widetilde{P}_t = \mathbf{1}_{11}$ and $\varphi = e^6_{11}$, so the "entire" representative agent has correct perceptions of the aggregate state. No households are ever "replaced" in the RA simulation, idiosyncratic shocks do not exist; only aggregate market resources are relevant. The vectors of perceptions evolve according to (27) and (28), and aggregate consumption is determined using (29).

The microeconomic (or cross sectional) regressions in Table 4 are generated using a single 4000 period sample of the history, from $t = 0$ to $t = 4000$, using 5000 of the 20,000 households. After dropping observations with $\mathbf{y}_{t,i} = 0$, this leaves about 19 million observations, far larger than any consumption panel dataset that we know of. Standard errors are thus vanishingly small, and have little meaning in any case, which is why we do not report them in the table summarizing our microsimulation results.

When making their forecasts of expected income growth, households are assumed to forecast that the transitory component of income will grow by the factor $1/\theta_{t,i}$, which is the forecast implied by their observation of the idiosyncratic transitory component of income. Substantively, this assumption reflects the real-world fact that essentially all of the predictable variation in income growth at the household level comes from idiosyncratic components of income.

D.3 Cost of Stickiness Calculation

After simulating a population of households using the procedure in Appendix D.2, we have a history of micro observations $\left\{\{\mathbf{c}_{t,i}, \mathsf{d}_{t,i}\}_{t=0}^{T}\right\}_{i=1}^{I}$ and a history of aggregate permanent productivity levels $\{P_t\}_{t=0}^{T}$. Each household index i contains the history of many agents, as the agent at i dies and is replaced at the beginning of any period with $\mathsf{d}_{t,i} = 1$. Let $\tau_{i,n}$ be the n-th time t index where $\mathsf{d}_{t,i} = 1$; further define $N_i = \sum_{t=0}^{T} \mathsf{d}_{t,i}$, the number of replacement events for household index i.

A single consumer's (normalized) discounted sum of lifetime utility is then:

$$\mathrm{v}_{i,n} = P_{\tau_{i,n}}^{\rho-1} \sum_{t=\tau_{i,n}}^{\tau_{i,n+1}-1} \beta^{t-\tau_{i,n}} \mathrm{u}(\mathbf{c}_{t,i}).$$

Normalizing by aggregate productivity at birth P_t is equivalent to normalizing by the consumer's total productivity at birth $\boldsymbol{p}_{t,i}$ because $p_{t,i} = 1$ at birth by assumption.

The total number of households who are born and die in the history is:

$$N_I = \sum_{i=1}^{I} (N_i - 1).$$

The overall expected lifetime value at birth can then be computed as:

$$\overline{\mathrm{v}}_0 = N_I^{-1} \sum_{i=1}^{I} \sum_{n=1}^{N_i - 1} \mathrm{v}_{i,n}.$$

Because we use $T = 20,000$ and $I = 20,000$, and agents live for 200 periods on average ($\mathsf{D} = 0.005$), our simulated history includes about $N_I \approx IT\mathsf{D} = 2$ million consumer lifetimes. The standard errors on our numerically calculated $\overline{\mathrm{v}}_0$ and $\overline{\overline{\mathrm{v}}}_0$ are thus negligible and not reported.

In the SOE model, we use the same random seed for the frictionless and sticky specifications, so the same sequence of replacement events and income shocks occurs in both. With no externalities or general equilibrium effects, the distribution of states that consumers are born into is likewise identical, so the "value ratio" calculation is valid.

The cost of stickiness in the HA-DSGE model is slightly more complicated. If we used the generated histories of the frictionless and sticky specifications to compute $\overline{\mathrm{v}}_0$ and $\overline{\overline{\mathrm{v}}}_0$, the calculated ω would represent a newborn's willingness-to-pay for *everyone* to be frictionless rather than sticky. We are interested in the utility cost of *just one agent* having sticky expectations, so an alternate procedure is required.

We compute $\overline{\overline{\mathrm{v}}}_0$ in the HA-DSGE model the same as in the SOE model. However, $\overline{\mathrm{v}}_0$ is calculated as the expected lifetime (normalized) value of a newborn who is frictionless *but lives in a world otherwise populated by sticky consumers*. To do this, we simulate a new history of micro observations using the consumption function for the sticky HA-DSGE economy, but with all I households updating their knowledge of the aggregate state frictionlessly. Critically, we *do not* actually calculate $\mathbf{A}_t = \mathbf{K}_{t+1}$ each period; instead, we use the *same sequence* of $\mathbf{A}_t$ that occurred in the ordinary sticky simulation. Thus our simulated population of I households represents an infinitesimally small portion of an economy made up (almost) entirely of consumers with sticky expectations. The calculated ω is thus the willingness-to-pay to be the very first agent to "wake up."

The formula for willingness-to-pay (19) arises from the homotheticity of the household's problem with respect to $\boldsymbol{p}_{t,i}$. If a consumer gives up an ω portion of their permanent income at the moment they are "born", before receiving income that period, then his normalized market resources will still be $m_{t,i} = \mathsf{W}_t$, and he will make the same normalized consumption choice that he would have, had he not lost any permanent income. In fact, he will make the *exact same* sequence of normalized consumption choices for his entire life; the *level* of his consumption will be scaled by the factor $(1-\omega)$ in every period. With CRRA utility, this means that utility is scaled by $(1-\omega)^{1-\rho}$ in every period of life, which can be factored out of the lifetime summation. The indifference condition between being frictionless and losing an ω fraction of permanent income versus having sticky expectations (and not losing) can be easily rearranged into (19).

E Muth–Lucas–Pischke

To see how the Muth–Lucas–Pischke model can generate smoothness, note that in the Muth framework, agents update their estimate of permanent income according to an equation of the form:[61]

$$\hat{P}_{t+1} \quad = \quad \Pi \mathbf{Y}_{t+1} + (1-\Pi)\hat{P}_t.$$

We can now consider the dynamics of aggregate consumption in response to the arrival of an aggregate shock that (unbeknownst to the consumer) is permanent. The consumer spends Π of the shock in the first period, leaving $(1-\Pi)$ unspent because that reflects the average transitory component of an undifferentiated shock. However, since the shock really was permanent, income next period does not fall back as the consumer guessed it would on the basis of the mistaken belief that $(1-\Pi)$ of the shock was transitory. The next-period consumer treats this surprise as a positive shock relative to expected income, and spends the same proportion Π out of the perceived new shock. These dynamics continue indefinitely, but with each successive perceived shock (and therefore each consumption increment) being smaller than the last by the proportion $(1-\Pi)$. Thus, after a true permanent shock received in period t, the full-information prediction of the expected dynamics of future consumption changes would be $\Delta \mathbf{C}_{t+n+1} = (1-\Pi)\Delta \mathbf{C}_{t+n} + \epsilon_{t+n}$.[62]

At first blush, this predictability in consumption growth would appear to be a violation of Hall (1978)'s proof that, for consumers who make rational estimates of their permanent income, consumption must be a random walk. The reconciliation is that what Hall proves is that consumption must be a random walk *with respect to the knowledge the consumer has.* The random walk proposition remains true for consumers whose knowledge base contains only the perceived level of aggregate income. Our thought experiment was to ask how much predictability would be found by an econometrician *who knows more than the consumer* about the level of aggregate permanent income.

The in-principle reconciliation of econometric evidence of predictability/excess smoothness in consumption growth, and the random walk proposition, is therefore that the econometricians who are making their forecasts of aggregate consumption growth use additional variables (beyond the lagged history of aggregate income itself), and that those variables have useful predictive power.[63]

[61] $\hat{P}_t$ is used to denote that households do an optimal signal-extraction (as opposed to having sticky expectations resulting in $\widetilde{P}_t$).

[62] The reciprocal logic would apply in the case of a shock that was known by the econometrician to be perfectly transitory, generating the same serial correlation in predictable consumption growth as in the case of the known-to-be-permanent shock. The only circumstance under which this serial correlation does *not* arise is when the econometrician has exactly the same beliefs as the consumer about the breakdown of the shock between transitory components. More precisely, it is still the case that the serial correlation coefficient on the predictable component of consumption growth is $(1-\Pi)$. But that predictable component itself is now zero, and $(1-\Pi) \times 0 = 0$.

[63] This is logically identical to Pischke's analysis of the case where the macroeconometrician knows that aggregate shocks are permanent, but the microeconomic consumers do not perceive those aggregate permanent shocks.

F Alternate Belief Specification

In the model presented in the main text, households with sticky expectations use the same consumption function as households who frictionlessly observe macroeconomic information in all periods. They treat their *perceptions* of macroeconomic states as if they were the true values, and do not account for their inattention when optimizing. In this appendix, we present an alternate specification in which households with sticky expectations partially account for their inattention by optimizing as if the flow of macroeconomic information they will receive is the true aggregate shock process. Simulated results analogous to Table 3 in the main text are presented below in Table 8.

Sticky expectations households do not update their macroeconomic information a $1-\Pi$ fraction of the time. In these periods, they perceive that there was no permanent aggregate shock Ψ_t and no innovation to the aggregate growth rate Φ_t. When they do update, they learn of the accumulation of permanent aggregate shocks since their last update (compounded with deviations from the last observed aggregate growth rate), as well as the new growth rate. In the "alternate beliefs" specification, households solve for their optimal consumption rule by treating their *perceived flow* of macroeconomic information as the true aggregate process. In this way, they *partially* account for their inattention by recognizing that the macroeconomic news they will perceive is leptokurtic relative to frictionless households.

The perceived aggregate shock process on which sticky households optimize is a linear combination of the shocks they perceive in non-updating periods (with weight $1-\Pi$) and the shocks they perceive when they do update (with weight Π). In periods in which they do and don't update, households treat the distribution of aggregate shocks as respectively:

$$\Theta_t^{\Pi} \sim \mathcal{N}(-\sigma_{\Theta}^2/2, \sigma_{\Theta}^2), \qquad \Psi_t^{\Pi} \sim \mathcal{N}(-\sigma_{\Psi}^2/(2\Pi), \sigma_{\Psi}^2/\Pi), \qquad \Xi^{\Pi} \sim \Xi^{\lfloor 1/\Pi \rceil}.$$

$$\Theta_t^{\cancel{\Pi}} \sim \mathcal{N}(-(\sigma_{\Theta}^2 + \sigma_{\Psi}^2/\Pi)/2, \sigma_{\Theta}^2 + \sigma_{\Psi}^2/\Pi), \qquad \Psi_t^{\cancel{\Pi}} = 1, \qquad \Xi^{\cancel{\Pi}} \sim I.$$

Here, Ξ represents the transition matrix among discrete Markov states for Φ_t in the true aggregate shock process. Under sticky expectations, households optimize under the assumption that in the Π fraction of periods in which Φ_t is observed, the true transition process has transpired an average of $\lfloor 1/\Pi \rceil$ times since the last update (four, under our calibration); they anticipate no Markov dynamics in the periods when they do not update (identity matrix I). Likewise, aggregate permanent shocks are interpreted to be degenerate in non-updating periods, but to make up for the fact that updating periods are one quarter as common, when an update occurs its variance is four times as large as in the baseline model.

In non-updating periods, households interpret all deviations from expected P_t as transitory aggregate shocks, so their perceived variance of Θ_t includes both transitory aggregate variance and a geometric series of permanent aggregate variance, decaying at rate $(1-\Pi)$:

$$\sigma_{\Theta}^2 + \sigma_{\Psi}^2 + (1-\Pi)\sigma_{\Psi}^2 + (1-\Pi)^2\sigma_{\Psi}^2 + \cdots = \sigma_{\Theta}^2 + \sigma_{\Psi}^2/\Pi.$$

This alternate belief specification does not have sticky expectations households fully and correctly adjust for their inattention. They do not track the *number* of periods since their last macroeconomic update, instead treating all non-updating periods alike from the perspective of perceived transitory shocks. Households act according to the same consumption function whether or not they just updated; the more sophisticated shock structure is used only to better approximate the perceived arrival of macroeconomic news when solving the problem. Moreover, households do not account for the positive covariance between accumulated permanent aggregate shocks and the innovation to Φ_t in periods when they *do* update. Incorporating these calculations would be extremely computationally burdensome, while changing the optimal consumption policy by very little. To the extent that our model represents an abstraction from households choosing the frequency of updating to balance the marginal cost and benefit of obtaining macroeconomic news (see section 6), it seems unlikely that agents would then adopt a vastly more complicated view of the world to offset the mild consequences of their inattention.

The key result is that households' optimal consumption function barely changes from baseline when the alternate beliefs are introduced: across states actually attained during simulation, normalized consumption differs by no more than 0.2 percent, and the difference is less than 0.02 percent in the vast majority of states. More importantly, the macroeconomic dynamics generated by sticky expectations households' collective behavior is nearly identical between the bottom panels of Table 8 below and Table 3 in the main text.[64] This experiment represents a more general proposition that our main results should be robust to the details of the precise specification of households' understanding of their inattention, so long as the key feature remains that agents' idiosyncratic errors are *systematically correlated* due to the lag in information.

G Additional Calculations

G.1 Quadratic Utility Consumption Dynamics

This appendix derives the equation (3) asserted in the main text. Start with the definition of consumption for the updaters,

$$\begin{aligned}
\mathbf{C}_t^{\pi} &\equiv \Pi^{-1}\int_0^1 \pi_{t,i}\mathbf{c}_{t,i}\,\mathrm{d}i \\
&= \Pi^{-1}\int_0^1 \pi_{t,i}(\mathsf{r}/\mathsf{R})\mathbf{o}_{t,i}\,\mathrm{d}i \\
&= \Pi^{-1}(\mathsf{r}/\mathsf{R})\int_0^1 \pi_{t,i}\mathbf{o}_{t,i}\,\mathrm{d}i \\
&= \Pi^{-1}(\mathsf{r}/\mathsf{R})\Pi\mathbf{O}_t \\
&= (\mathsf{r}/\mathsf{R})\mathbf{O}_t,
\end{aligned}$$

[64] The top panels are literally identical, as they report the same model.

where the penultimate line follows from the fact that the updaters are chosen randomly among members of the population so that the average per capita value of $\mathbf{o}$ among updaters is equal to the average per capita value of $\mathbf{o}$ for the population as a whole.

The text asserts (equation (3)) that

$$\begin{aligned}
\mathbf{C}_{t+1} &= \Pi\Delta\mathbf{C}^{\pi}_{t+1} + (1-\Pi)\Delta\mathbf{C}_t \\
&\approx (1-\Pi)\Delta\mathbf{C}_t + \xi_{t+1}.
\end{aligned}$$

To see this, define market resources $M_t = Y_t + \mathsf{R}A_t$ where Y_t is noncapital income in period t and A_t is the level of nonhuman assets with which the consumer ended the previous period; and define H_t as 'human wealth,' the present discounted value of future noncapital income. Then write

$$\begin{aligned}
\mathbf{C}^{\pi}_{t+1} &= (\mathsf{r}/\mathsf{R})\big(M_{t+1} + H_{t+1}\big) \\
\mathbf{C}^{\pi}_{t} &= (\mathsf{r}/\mathsf{R})\big(M_{t} + H_{t}\big) \\
\mathbf{C}^{\pi}_{t+1} - \mathbf{C}^{\pi}_{t} &= (\mathsf{r}/\mathsf{R})\big(M_{t+1} - M_t + H_{t+1} - H_t\big) \\
\mathbf{C}^{\pi}_{t+1} - \mathbf{C}^{\pi}_{t} &= (\mathsf{r}/\mathsf{R})\big(\mathsf{R}(Y_t + M_t - \mathbf{C}_t) - M_t + H_{t+1} - H_t\big).
\end{aligned} \tag{32}$$

What theory tells us is that *if aggregate consumption were chosen frictionlessly in period* t, then this expression would be white noise; that is, we know that

$$(\mathsf{r}/\mathsf{R})\big(\mathsf{R}(Y_t + M_t - \mathbf{C}^{\pi}_t) - M_t + H_{t+1} - H_t\big) = \xi_{t+1}$$

for some white noise ξ_{t+1}. The only difference between this expression and the RHS of (32) is the Π superscript on the $\mathbf{C}_t$. Thus, substituting, we get

$$\begin{aligned}
\mathbf{C}^{\pi}_{t+1} - \mathbf{C}^{\pi}_{t} &= (\mathsf{r}/\mathsf{R})\big(\mathsf{R}\,(Y_t + M_t - (\mathbf{C}_t + \mathbf{C}^{\pi}_t - \mathbf{C}^{\pi}_t)) - M_t + H_{t+1} - H_t\big) \\
\mathbf{C}^{\pi}_{t+1} - \mathbf{C}^{\pi}_{t} &= (\mathsf{r}/\mathsf{R})\big(\mathsf{R}(Y_t + M_t - \mathbf{C}^{\pi}_t) - M_t + H_{t+1} - H_t\big) + (\mathsf{r}/\mathsf{R})(\mathbf{C}^{\pi}_t - \mathbf{C}_t) \\
&= \xi_{t+1} + (\mathsf{r}/\mathsf{R})(\mathbf{C}^{\pi}_t - \mathbf{C}_t).
\end{aligned}$$

So equation (3) can be rewritten as

$$\Delta\mathbf{C}_{t+1} = (1-\Pi)\Delta\mathbf{C}_t + \Pi\big((\mathsf{r}/\mathsf{R})(\mathbf{C}^{\pi}_t - \mathbf{C}_t) + \xi_{t+1}\big)$$

where ξ_{t+1} is a white noise variable. Thus,

$$\Delta\mathbf{C}_{t+1} = (1-\Pi)\big(1 + \underbrace{(\mathsf{r}/\mathsf{R})}_{\approx 0}\big)\Delta\mathbf{C}_t + \underbrace{\Pi\xi_{t+1}}_{\equiv \epsilon_{t+1}} \tag{33}$$

for a white noise variable ϵ_{t+1}, and $(\mathsf{r}/\mathsf{R}) \approx 0$ for plausible quarterly interest rates. (33) leads directly to (3).

G.2 Population Variance of Idiosyncratic Permanent Income

This appendix follows closely Appendix A in the ECB working paper version of Carroll, Slacalek, and Tokuoka (2015).[65] It computes dynamics and steady state of the square of the idiosyncratic component of permanent income (from which the variance can be

[65] Carroll, Christopher D., Jiri Slacalek, and Kiichi Tokuoka (2014): "Buffer-Stock Saving in a Krusell–Smith World," working paper 1633, European Central Bank, https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1633.pdf.

derived). Recalling that consumers are born with $p_{t,i} = 1$:

$$\begin{aligned} p_{t+1,i} &= (1-\mathsf{d}_{t+1,i})p_{t,i}\psi_{t+1,i} + \mathsf{d}_{t+1,i} \\ p^2_{t+1,i} &= \big((1-\mathsf{d}_{t+1,i})p_{t,i}\psi_{t+1,i}\big)^2 + \underbrace{(1-\mathsf{d}_{t+1,i})\mathsf{d}_{t+1,i}}_{=0} 2p_{t,i}\psi_{t+1,i} + \mathsf{d}^2_{t+1,i} \end{aligned}$$

and because $\mathbb{E}_t[\mathsf{d}^2_{t+1,i}] = \mathsf{D}$ we have

$$\begin{aligned} \mathbb{E}_t[p^2_{t+1,i}] &= \mathbb{E}_t[((1-\mathsf{d}_{t+1,i})p_{t,i}\psi_{t+1,i})^2] + \mathsf{D} \\ &= (1-\mathsf{D})p^2_{t,i}\,\mathbb{E}[\psi^2] + \mathsf{D}. \end{aligned}$$

Defining the mean operator $\mathbb{M}[\bullet_t] = \int_0^1 \bullet_{t,\iota} \mathrm{d}\iota$, we have

$$\mathbb{M}\left[p^2_{t+1}\right] = (1-\mathsf{D})\mathbb{M}[p^2_t]\,\mathbb{E}[\psi^2] + \mathsf{D},$$

so that the steady state expected level of $\mathbb{M}[p^2] \equiv \lim_{t\to\infty} \mathbb{M}[p^2_t]$ can be found from

$$\begin{aligned} \mathbb{M}[p^2] &= (1-\mathsf{D})\,\mathbb{E}[\psi^2]\mathbb{M}[p^2] + \mathsf{D} \\ &= \frac{\mathsf{D}}{1-(1-\mathsf{D})\,\mathbb{E}[\psi^2]}. \end{aligned}$$

Finally, note the relation between p^2 and the variance of p:

$$\begin{aligned} \sigma^2_p &= \mathbb{M}[(p-\mathbb{M}[p])^2] \\ &= \mathbb{M}[(p^2 - 2p\mathbb{M}[p] + (\mathbb{M}[p])^2)] \\ &= \mathbb{M}[p^2] - 1, \end{aligned}$$

where the last line follows because under the other assumptions we have made, $\mathbb{M}[p] = 1$.

For the preceding derivations to be valid, it is necessary to impose the parameter restriction $(1-\mathsf{D})\,\mathbb{E}[\psi^2] < 1$. This requires that income does not spread out so quickly among survivors as to overcome the compression of the distribution that arises because of death.

G.3 Converting Annual to Quarterly Variances for Idiosyncratic Shocks

If the quarterly transitory shock is θ_t, define the annual transitory shock as:

$$\theta^a_t = \sum_{i=1}^{4} \frac{\theta_{t+i}}{4}$$

for $t = 0, 4, 8, \ldots$ Then the variance of the annual transitory shock is $\frac{1}{4}$ of the variance of the quarterly transitory shock: $\text{var}(\theta^a) = \frac{4}{16}\,\text{var}(\theta) = \frac{1}{4}\,\text{var}\,\theta$. We therefore multiply our calibrated annual transitory shock (0.03) by 4 to get a quarterly number.

Let ψ_t be the quarterly permanent shock. Define the annual permanent shock as:

$$\psi^a_t = \prod_{i=1}^{4} \psi_{t+i}$$

for $t = 0, 4, 8, \ldots$ Then the variance of the annual permanent shock is $\big(1 + \text{var}(\psi)\big)^4 \approx$

$4 \times \text{var}(\psi)$ for small $\text{var}(\psi)$. Therefore we divide our calibrated annual permanent shock (0.012) by 4 to get a quarterly number.

G.4 Muth (1960) Signal Extraction

Muth (1960), pp. 303–304, shows that the signal-extracted estimate of permanent income is

$$\widetilde{P}_t = v_1 Y_t + v_2 Y_{t-1} + v_3 Y_{t-2} + ...$$

for a sequence of v's given by

$$v_k = (1-\lambda_1)\lambda_1^{k-1}$$

for $k = 1, 2, 3, ...$ So:

$$\begin{aligned}
\widetilde{P}_t &= (1-\lambda_1)(\qquad Y_t + \lambda_1 Y_{t-1} + \lambda_1^2 Y_{t-2} ...) \\
\widetilde{P}_{t+1} &= (1-\lambda_1)(Y_{t+1} + \lambda_1 Y_t + \lambda_1^2 Y_{t-1} + \lambda_1^3 Y_{t-2} ...) \\
&= (1-\lambda_1)\ Y_{t+1} + \lambda_1 \underbrace{(1-\lambda_1)(Y_t + \lambda_1^2 Y_{t-1} + \lambda_1^3 Y_{t-2} ...)}_{\widetilde{P}_t} \\
&= (1-\lambda_1) Y_{t+1} + \lambda_1 \widetilde{P}_t
\end{aligned}$$

This compares with (32) in the main text

$$\widetilde{P}_{t+1} = \Pi Y_{t+1} + (1-\Pi)\widetilde{P}_t$$

so the relationship between our Π and Muth's λ_1 is:

$$\lambda_1 = 1 - \Pi$$

Defining the signal-to-noise ratio $\varphi = \sigma_{\boldsymbol{\psi}}/\sigma_{\boldsymbol{\theta}}$, starting with equation (3.10) in Muth (1960) we have

$$\begin{aligned}
\lambda_1 &= 1 + (1/2)\varphi^2 - \varphi\sqrt{1 + \varphi^2/4} \\
(1-\Pi) &= 1 + (1/2)\varphi^2 - \varphi\sqrt{1 + \varphi^2/4} \\
-\Pi &= (1/2)\varphi^2 - \varphi\sqrt{1 + \varphi^2/4}
\end{aligned}$$

yielding equation (20) in the main text.

Table 5 Equilibrium Statistics

	SOE Model		HA-DSGE Model	
	Frictionless	Sticky	Frictionless	Sticky
Means				
A	7.49	7.43	56.85	56.72
C	2.71	2.71	3.44	3.44
Standard Deviations				
Aggregate Time Series ('Macro')				
$\log A$	0.332	0.321	0.276	0.272
$\Delta \log \mathbf{C}$	0.010	0.007	0.010	0.005
$\Delta \log \mathbf{Y}$	0.010	0.010	0.007	0.007
Individual Cross Sectional ('Micro')				
$\log \mathbf{a}$	0.926	0.927	1.015	1.014
$\log \mathbf{c}$	0.790	0.791	0.598	0.599
$\log p$	0.796	0.796	0.796	0.796
$\log \mathbf{y} \vert \mathbf{y} > 0$	0.863	0.863	0.863	0.863
$\Delta \log \mathbf{c}$	0.098	0.098	0.054	0.055
Cost of Stickiness	4.82e–4		4.51e–4	

Notes: The cost of stickiness is calculated as the proportion by which the permanent income of a newborn frictionless consumer would need to be reduced in order to achieve the same reduction of expected value associated with forcing them to become a sticky expectations consumer.

Table 6 Aggregate Consumption Dynamics in HA-DSGE Model

$$\Delta \log \mathbf{C}_{t+1} = \varsigma + \chi \Delta \log \mathbf{C}_t + \eta \mathbb{E}_t[\Delta \log \mathbf{Y}_{t+1}] + \alpha A_t + \epsilon_{t+1}$$

Expectations : Dep Var Independent Variables			OLS or IV	2^{nd} Stage $\bar{R}^2$	KP p-val Hansen J p-val
Frictionless : $\Delta \log \mathbf{C}^*_{t+1}$ (with measurement error $\mathbf{C}^*_t = \mathbf{C}_t \times \xi_t$);					
$\Delta \log \mathbf{C}^*_t$	$\Delta \log \mathbf{Y}_{t+1}$	A_t			
0.189$^{\bullet\bullet\bullet}$			OLS	0.036	
(0.072)					
0.476			IV	0.020	0.318
(0.354)					0.556
	0.368		IV	0.017	0.107
	(0.321)				0.457
		−0.34e–4	IV	0.015	0.000
		(0.98e–4)			0.433
0.289	0.214	0.01e–4	IV	0.020	0.572
(0.463)	(0.583)	(1.87e–4)			0.531
Memo: For instruments $\mathbf{Z}_t$, $\Delta \log \mathbf{C}^*_t = \mathbf{Z}_t \zeta$, $\bar{R}^2 = 0.023$; $\text{var}(\log(\xi_t)) = 4.16\text{e–}6$					
Sticky : $\Delta \log \mathbf{C}^*_{t+1}$ (with measurement error $\mathbf{C}^*_t = \mathbf{C}_t \times \xi_t$);					
$\Delta \log \mathbf{C}^*_t$	$\Delta \log \mathbf{Y}_{t+1}$	A_t			
0.467$^{\bullet\bullet\bullet}$			OLS	0.223	
(0.061)					
0.773$^{\bullet\bullet\bullet}$			IV	0.230	0.000
(0.108)					0.542
	0.912$^{\bullet\bullet\bullet}$		IV	0.145	0.105
	(0.245)				0.187
		−0.97e–4$^{\bullet}$	IV	0.059	0.000
		(0.56e–4)			0.002
0.670$^{\bullet\bullet\bullet}$	0.171	0.12e–4	IV	0.231	0.460
(0.181)	(0.363)	(0.86e–4)			0.551
Memo: For instruments $\mathbf{Z}_t$, $\Delta \log \mathbf{C}^*_t = \mathbf{Z}_t \zeta$, $\bar{R}^2 = 0.232$; $\text{var}(\log(\xi_t)) = 4.16\text{e–}6$					

Notes: Reported statistics are the average values for 100 samples of 200 simulated quarters each. Bullets indicate that the average sample coefficient divided by average sample standard error is outside of the inner 90%, 95%, and 99% of the standard normal distribution. Instruments $\mathbf{Z}_t = \{\Delta \log \mathbf{C}_{t-2}, \Delta \log \mathbf{C}_{t-3}, \Delta \log \mathbf{Y}_{t-2}, \Delta \log \mathbf{Y}_{t-3}, A_{t-2}, A_{t-3}, \Delta_8 \log \mathbf{C}_{t-2}, \Delta_8 \log \mathbf{Y}_{t-2}\}$.

Table 7 Aggregate Consumption Dynamics in RA Model

$$\Delta \log \mathbf{C}_{t+1} = \varsigma + \chi \Delta \log \mathbf{C}_t + \eta \mathbb{E}_t[\Delta \log \mathbf{Y}_{t+1}] + \alpha A_t + \epsilon_{t+1}$$

Expectations : Dep Var Independent Variables			OLS or IV	2$^{\text{nd}}$ Stage $\bar{R}^2$	KP p-val Hansen J p-val
Frictionless : $\Delta \log \mathbf{C}^*_{t+1}$ (with measurement error $\mathbf{C}^*_t = \mathbf{C}_t \times \xi_t$);					
$\Delta \log \mathbf{C}^*_t$	$\Delta \log \mathbf{Y}_{t+1}$	A_t			
-0.015			OLS	0.002	
(0.077)					
0.387			IV	0.014	0.367
(0.390)					0.570
	0.390		IV	0.016	0.084
	(0.311)				0.475
		-0.26e–4	IV	0.016	0.000
		(1.11e–4)			0.493
0.122	0.267	0.16e–4	IV	0.018	0.547
(0.519)	(0.575)	(2.12e–4)			0.572
Memo: For instruments $\mathbf{Z}_t$, $\Delta \log \mathbf{C}^*_t = \mathbf{Z}_t \zeta$, $\bar{R}^2 = 0.018$; var(log(ξ_t)) = 3.33e–6					
Sticky : $\Delta \log \mathbf{C}^*_{t+1}$ (with measurement error $\mathbf{C}^*_t = \mathbf{C}_t \times \xi_t$);					
$\Delta \log \mathbf{C}^*_t$	$\Delta \log \mathbf{Y}_{t+1}$	A_t			
0.412•••			OLS	0.179	
(0.063)					
0.788•••			IV	0.183	0.001
(0.138)					0.532
	0.641•••		IV	0.128	0.085
	(0.163)				0.171
		-0.47e–4	IV	0.075	0.000
		(0.52e–4)			0.027
0.632•••	0.118	0.10e–4	IV	0.184	0.321
(0.223)	(0.280)	(0.79e–4)			0.480
Memo: For instruments $\mathbf{Z}_t$, $\Delta \log \mathbf{C}^*_t = \mathbf{Z}_t \zeta$, $\bar{R}^2 = 0.186$; var(log(ξ_t)) = 3.33e–6					

Notes: Reported statistics are the average values for 100 samples of 200 simulated quarters each. Bullets indicate that the average sample coefficient divided by average sample standard error is outside of the inner 90%, 95%, and 99% of the standard normal distribution. Instruments $\mathbf{Z}_t = \{\Delta \log \mathbf{C}_{t-2}, \Delta \log \mathbf{C}_{t-3}, \Delta \log \mathbf{Y}_{t-2}, \Delta \log \mathbf{Y}_{t-3}, A_{t-2}, A_{t-3}, \Delta_8 \log \mathbf{C}_{t-2}, \Delta_8 \log \mathbf{Y}_{t-2}\}$.

Table 8 Aggregate Consumption Dynamics in SOE Model (Alternate Beliefs)

$$\Delta \log \mathbf{C}_{t+1} = \varsigma + \chi \Delta \log \mathbf{C}_t + \eta \mathbb{E}_t[\Delta \log \mathbf{Y}_{t+1}] + \alpha A_t + \epsilon_{t+1}$$

Expectations : Dep Var Independent Variables			OLS or IV	2nd Stage $\bar{R}^2$	KP p-val Hansen J p-val
Frictionless : $\Delta \log \mathbf{C}^*_{t+1}$ (with measurement error $\mathbf{C}^*_t = \mathbf{C}_t \times \xi_t$);					
$\Delta \log \mathbf{C}^*_t$	$\Delta \log \mathbf{Y}_{t+1}$	A_t			
$0.295^{\bullet\bullet\bullet}$			OLS	0.087	
(0.066)					
$0.660^{\bullet\bullet}$			IV	0.040	0.237
(0.309)					0.600
	$0.457^{\bullet\bullet}$		IV	0.035	0.059
	(0.209)				0.421
		−6.92e–4	IV	0.026	0.000
		(5.87e–4)			0.365
0.420	0.258	0.45e–4	IV	0.041	0.516
(0.428)	(0.365)	(9.51e–4)			0.529
Memo: For instruments $\mathbf{Z}_t$, $\Delta \log \mathbf{C}^*_t = \mathbf{Z}_t \zeta$, $\bar{R}^2 = 0.039$; var(log(ξ_t)) = 5.99e–6					
Sticky : $\Delta \log \mathbf{C}^*_{t+1}$ (with measurement error $\mathbf{C}^*_t = \mathbf{C}_t \times \xi_t$);					
$\Delta \log \mathbf{C}^*_t$	$\Delta \log \mathbf{Y}_{t+1}$	A_t			
$0.508^{\bullet\bullet\bullet}$			OLS	0.263	
(0.058)					
$0.800^{\bullet\bullet\bullet}$			IV	0.257	0.000
(0.104)					0.552
	$0.857^{\bullet\bullet\bullet}$		IV	0.195	0.060
	(0.182)				0.229
		$-8.12\text{e–}4^{\bullet\bullet}$	IV	0.065	0.000
		(3.97e–4)			0.002
$0.659^{\bullet\bullet\bullet}$	0.191	0.60e–4	IV	0.259	0.355
(0.187)	(0.277)	(5.01e–4)			0.544
Memo: For instruments $\mathbf{Z}_t$, $\Delta \log \mathbf{C}^*_t = \mathbf{Z}_t \zeta$, $\bar{R}^2 = 0.257$; var(log(ξ_t)) = 6.03e–6					

Notes: Reported statistics are the average values for 100 samples of 200 simulated quarters each. Bullets indicate that the average sample coefficient divided by average sample standard error is outside of the inner 90%, 95%, and 99% of the standard normal distribution. Instruments $\mathbf{Z}_t = \{\Delta \log \mathbf{C}_{t-2}, \Delta \log \mathbf{C}_{t-3}, \Delta \log \mathbf{Y}_{t-2}, \Delta \log \mathbf{Y}_{t-3}, A_{t-2}, A_{t-3}, \Delta_8 \log \mathbf{C}_{t-2}, \Delta_8 \log \mathbf{Y}_{t-2}\}$.