



EUROPEAN CENTRAL BANK
EUROSYSTEM

Working Paper Series

Christopher D. Carroll, Edmund Crawley,
Jiri Slacalek, Kiichi Tokuoka, Matthew N. White

Sticky expectations and consumption dynamics

No 2152 / May 2018

Disclaimer: This paper should not be reported as representing the views of the European Central Bank (ECB). The views expressed are those of the authors and do not necessarily reflect those of the ECB.

Abstract

Macroeconomic models often invoke consumption “habits” to explain the substantial persistence of macroeconomic consumption growth. But a large literature has found no evidence of habits in the microeconomic datasets that measure the behavior of individual households. We show that the apparent conflict can be explained by a model in which consumers have accurate knowledge of their personal circumstances but ‘sticky expectations’ about the macroeconomy. In our model, the persistence of aggregate consumption growth 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, Sticky Expectations, Habits, Inattention, Imperfect Information

JEL codes D83, D84, E21, E32

PDF: <http://econ.jhu.edu/people/ccarroll/papers/cAndCwithStickyE.pdf>

Web: <http://econ.jhu.edu/people/ccarroll/papers/cAndCwithStickyE/>

BibTeX: <http://econ.jhu.edu/people/ccarroll/papers/cAndCwithStickyE.bib>

Archive: <http://econ.jhu.edu/people/ccarroll/papers/cAndCwithStickyE.zip>

Slides: Versions to View or Print

Toolkit: <http://github.com/Econ-ARK/HARK>

Non-technical Summary

Starting with Campbell and Deaton (1989), the macroeconomics, finance, and international economics literatures have established a wide variety of evidence that aggregate consumption exhibits ‘excess smoothness’ compared with the benchmark Hall (1978) random walk model of consumption. Over the past two decades many papers in these fields have responded to this problem by incorporating ‘habit formation’ in the utility function of a representative agent.

This literature typically measures excess smoothness with a parameter conventionally labeled as the ‘habit formation coefficient’ (which we denote as χ). A recent comprehensive meta-analysis of 597 published estimates reports that studies based on macro data find that $\chi = 0.6$ on average.

If habits are a true structural characteristic of people’s utility functions, we should see their effects microeconomic data as well as macroeconomic aggregates. But empirical studies using household-level data strongly reject the existence of habits of the magnitude necessary to explain aggregate consumption dynamics. The modal estimate from Havranek, Rusnak and Sokolova’s survey of the micro literature is a ‘habit’ parameter of 0; the mean estimate is about 0.1.

We propose a simple solution to this puzzle. Instead of having consumption habits, microeconomic consumers are subject to a modest informational friction: Not everybody instantaneously notices all *macroeconomic* developments. Instead, households’ macroeconomic expectations are “sticky.” Specifically, while each consumer perfectly (‘frictionlessly’) perceives his own personal circumstances (employment status, wage rate, income received, etc), consumers’ information about macroeconomic quantities like aggregate productivity growth arrives only occasionally (as in the Calvo model of firms’ price updating).

Consumption sluggishness arises as follows. As in the standard (frictionless) setup, sticky-expectations households perfectly observe their market resources (wealth, debt, etc) and income. However, a consumer whose beliefs about the state of 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 (and possibly out-of-date) perception of macroeconomic circumstances. The model thus 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 conventional estimates of the degree of inattention measured using expectations data on other aggregate variables (e.g., inflation expectations), the model's implications for the persistence in aggregate consumption growth match well the estimates of the 'excess smoothness' of consumption growth in the macro literature.

Despite aggregate sluggishness, at the level of individual households, high-frequency consumption growth has little predictability. This can be reconciled with aggregate smoothness because the rationally appropriate contribution of the consumer's perception of the macroeconomic environment to their individual spending choices is swamped by the importance of fluctuations in idiosyncratic components of income which we assume consumers have no difficulty observing (and to which we assume they are perfectly attentive).

Our results are essentially the same in a partial equilibrium model (in which factor prices are constant) and a heterogeneous-agents DSGE model with aggregate shocks (which affect factor prices). Data simulated from our models reproduce what we take to be the main stylized facts about individual and aggregate consumption dynamics.

When estimated on simulated *individual* data (corresponding to microeconomic evidence), regressions in the spirit of Hall (1978) and Campbell and Mankiw (1989) find that consumption growth exhibits little persistence. In this respect (and all others), the micro implications of the model are standard for models of this type (with uninsurable uncertainty as well as precautionary saving and perhaps liquidity constraints), which have been extensively studied in the micro literature.

We then analyze Hall (1978)/Campbell and Mankiw (1989) regressions with simulated *aggregate* data. Thanks to the law of large numbers, the idiosyncratic shocks that dominate the household data cancel out upon aggregation, leaving only the residual

systematic factors, which generate the much greater predictability in aggregate than in idiosyncratic data. Campbell and Mankiw (1989) proposed that such predictability arises because some people just spend all their income, and income growth is predictable. The habit formation literature has argued instead that predictability reflected the sluggishness of consumption growth itself. Horserace regressions that pit these two possibilities against each other produce a clear winner: Almost all of the predictability of consumption growth is explained by its correlation with lagged consumption growth; only a small portion comes from the predictable component of aggregate income growth—both in the data and in our model.

1 Introduction

Starting with Campbell and Deaton (1989), the macroeconomics, finance, and international economics literatures have concluded that aggregate consumption exhibits ‘excess smoothness’ compared with the benchmark Hall (1978) random walk model of consumption.¹ Over the past two decades many papers in these fields have responded to this problem by incorporating ‘habit formation’ in the utility function of a representative agent.

This literature typically measures excess smoothness with a parameter conventionally labeled as the ‘habit formation coefficient’ (which we denote as χ). A recent comprehensive meta-analysis of 597 published estimates (Havranek, Rusnak, and Sokolova (2017)) reports that studies based on macro data find that $\chi = 0.6$ on average; see Figure 1.²

If habits are a true structural characteristic of people’s utility functions, we should see their effects microeconomic data as well as macroeconomic aggregates. But empirical studies using household-level data strongly reject the existence of habits of the magnitude necessary to explain aggregate consumption dynamics. The modal estimate from Havranek, Rusnak, and Sokolova (2017)’s survey of the micro literature is a ‘habit’ parameter of 0; the mean estimate is about 0.1 (see Figure 1).³ Even among studies that have found evidence against the random walk proposition in micro data,⁴ few claim to have found more than a few percentage points’ worth of household-level spending growth to be predictable at any measured horizon. Roughly speaking, the predictability of aggregate spending growth is around an order of magnitude larger than predictability of household-level spending growth (say, 0.30 versus 0.03 in an adjusted R^2 sense).

We propose a simple solution to this puzzle. Instead of having consumption habits,

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

²For examples of such studies, see results and references in Fuhrer (2000) or Christiano, Eichenbaum, and Evans (2005).

³Dynan (2000) is the best known micro study; many others are reported in Havranek, Rusnak, and Sokolova (2017). It is well-known that household-level data on consumption are subject to substantial measurement error. Consequently, some attenuation of the micro estimates of χ might be expected (even though many studies use estimation techniques robust to measurement error). But an extreme amount of noise would be required for the estimated coefficient to be as close to 0 as estimated for the many micro studies in Figure 1.

⁴See the detailed discussion in Section 6.2.

microeconomic consumers experience a modest informational friction: Not everybody instantaneously notices all *macroeconomic* developments. Instead, households' macroeconomic expectations are "sticky," as in Mankiw and Reis (2002) and Carroll (2003). Specifically, while each consumer perfectly ('frictionlessly') perceives his own personal circumstances (employment status, wage rate, income received, etc), consumers' information about macroeconomic quantities like aggregate productivity growth arrives only occasionally (as in the Calvo model of firms' price updating).

Consumption sluggishness *a la* Campbell and Deaton (1989) arises as follows. As in the standard (frictionless) setup, sticky-expectations households perfectly observe their market resources (wealth, debt, etc) and income. However, a consumer whose beliefs about the state of 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 (and possibly out-of-date) perception of macroeconomic circumstances. The model thus 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 conventional estimates of the degree of inattention measured using expectations data on other aggregate variables (e.g., inflation expectations), the model's implications for the persistence in aggregate consumption growth match well the estimates of the 'excess smoothness' of consumption growth in the macro literature.

Despite aggregate sluggishness, at the level of individual households, high-frequency consumption growth has little predictability. This can be reconciled with aggregate smoothness because the rationally appropriate contribution of the consumer's perception of the macroeconomic environment to their individual spending choices is swamped by the importance of fluctuations in idiosyncratic components of income which we assume consumers have no difficulty observing (and to which we assume they are perfectly attentive).⁵

⁵Over long spans of time—say, over a life cycle—models with uncertainty and impatience imply that consumption growth parallels income growth. The theory has bite only when applied to high frequency movements in income—say, at the quarterly or annual frequency.

In our model, the sticky updating of beliefs about the aggregate economy takes the same form (and has the same magnitude) as proposed in Carroll (2003) as a microfoundation for the Mankiw and Reis (2002) model. An advantage compared to those papers is that because we are using an optimizing model, we are able to calculate an explicit utility cost of stickiness. Consistent with a theme in the literature on inattentiveness all the way back to Akerlof and Yellen (1985), we find that the utility penalty from inattention is low, so that, under our calibrated parameters, our consumers would not be willing to pay much for even the most perfect information about the macroeconomic state: They would be willing to pay roughly one two-thousandth of their lifetime income to be perfectly informed in every future period of their lifetime.

Our results are essentially the same in a partial equilibrium model (in which factor prices are constant) and a heterogeneous-agents DSGE model with aggregate shocks (which affect factor prices).⁶ Data simulated from our models reproduce what we take to be the main stylized facts about individual and aggregate consumption dynamics.

When estimated on simulated *individual* data (corresponding to microeconomic evidence), regressions in the spirit of Hall (1978) and Campbell and Mankiw (1989) find that consumption growth exhibits little persistence. This result is essentially identical across all variants of our models: partial or general equilibrium, with or without inattention. It comports well with the conclusions of a micro literature that was already large when Deaton (1992) surveyed it and has remained consistent in finding little persistence. In this respect (and all others), the micro implications of the model are standard for models of this type (with uninsurable uncertainty as well as precautionary saving and perhaps liquidity constraints), which have been extensively studied in the micro literature.⁷

We then analyze Hall (1978)/Campbell and Mankiw (1989) regressions with simulated *aggregate* data. Thanks to the law of large numbers, the idiosyncratic shocks that dominate the household data cancel out upon aggregation, leaving only the residual

⁶In both environments we calibrate households' income processes to be consistent with evidence from both microeconomic and macroeconomic data.

⁷For example, the model reproduces the robust pattern of evidence since Zeldes (1989a) has documented that consumption growth is faster for people with very low cash-on-hand (who may be near a liquidity constraint or below their desired precautionary buffers—or both).

systematic factors, which generate the much greater predictability in aggregate than in idiosyncratic data. Campbell and Mankiw (1989) proposed that such predictability arises because some people just spend all their income, and income growth is predictable. The habit formation literature has argued instead that predictability reflected the sluggishness of consumption growth itself. Horserace regressions that pit these two possibilities against each other produce a clear winner: Almost all of the predictability of consumption growth is explained by its correlation with lagged consumption growth; only a small portion comes from the predictable component of aggregate income growth – both in the data and in our model.

After a brief review of the extensive relevant literature, 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 versions of our models, which abide by the more realistic assumptions (CRRA utility, aggregate as well as individual shocks, time varying factor prices, etc) that have become conventional respectively in the micro and macro literatures.

After calibrating the model (section 5), we describe the stylized facts from both the micro and macro literatures that we argue need to be explained by a good microfounded macroeconomic model of consumption, and show that all of the various versions of our model (partial versus general equilibrium, etc) robustly reproduce those facts (section 6). This robustness indicates that our results are not a fragile implication of any highly specific framework but instead flow from the underlying structure of inattention that is the common element across all versions of our model (including the quadratic utility ‘toy model’ where the consequences can be seen most clearly). We then (section 7) calculate how much a fully informed consumer would be willing to pay at birth to enjoy instantaneous and perfect knowledge of aggregate developments as they live their life (not much, it turns out).

With our model’s quantitative results in hand, we describe the quantitative and qualitative differences between our model and the other ‘imperfect information’ approaches

to explaining aggregate consumption smoothness that have been explored in the prior literature (section 8). Our conclusion suggests directions for future research.

2 Relation to the Literature

No review of the empirical literature is needed; Havranek, Rusnak, and Sokolova (2017) have done an admirable job. Our only critique is that they have followed much of the prior literature in casually referring to the parameter of interest as the ‘habit coefficient.’ A better choice would have been to call it the ‘excess smoothness’ coefficient; ours is not the first paper to suggest that habits are not the only possible explanation for why consumption growth might be too smooth (compared to the Hall (1978) benchmark).

Our ‘sticky expectations’ approach is related to several strands of the burgeoning literature on models of imperfect information processing. A major strand in that literature is models of ‘rational inattention’ in the spirit of Sims (2003), in which agents have a limited ability to pay attention and allocate it optimally, recently embodied (for example) in the work of Maćkowiak and Wiederholt (2015). They 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) with rational inattention. The 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 challenge to this approach has been the extraordinary complexity of solving models that aim to work out the full implications of the fact that everyone else is working out the full implications of the fact that everyone else is rationally inattentive.⁸ In response, 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 rational inattention framework, even the simplified Gabaix approach

⁸For example, the literature on rational inattention has adopted a more stylized setup of idiosyncratic and aggregate income shocks and, to our knowledge, has so far not solved the full Krusell and Smith (1998) framework.

would be formidably difficult to embed in a model with a rich treatment of transitory and persistent income shocks, precautionary motives and other complexities entailed in modern models of microeconomic consumption decisions. It would be similarly challenging to determine how to apply the approaches of Woodford (2002) or Morris and Shin (2006) to our question.⁹

Another way to dial back the complexity of the rational inattention approach is to radically simplify the model's assumptions about decisionmaker's problem. In that spirit Reis (2006a) considers a model in which consumers with a linear consumption function and a conveniently simple environment optimally choose to be inattentive because of explicit (fixed monetary) costs of attention.¹⁰ In this framework, Reis (2006a) is able to calculate an explicit analytical formula for the tradeoff between the disutility from the increase in uncertainty caused by inattention, and the monetary savings due to infrequent payment of the cost of information. Reis shows that in his model, inattention is manifested in the fact the his consumers only gather new information (and therefore only update their consumption) at fixed intervals whose length depends on the cost of obtaining information versus the costs of remaining ignorant.

One of our objectives is to faithfully match microeconomic data. In such 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. It is concave, as the general theory suggests (Carroll and Kimball (1996)), and this concavity matters greatly for matching the main micro facts. 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, nor its prediction that there will be occasional periods in which it moves a lot (at dates of adjustment) and then remains constant at that newer level for some extended period. This critique applies generically to models that incorporate a convex cost of adjustment—whether to the consumer's stock of information (Reis (2006a)) or to the level of consumption as in

⁹Arguably, 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.

¹⁰Reis (2006b) presents a similar setup with a producer setting prices subject to information processing constraints.

Chetty and Szeidl (2016). All such models imply counterfactually ‘jerky’ behavior of spending at the microeconomic level.¹¹

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 substantial recent literature exploring models of this kind; our assumption that macroeconomic expectations are sticky makes no material difference to the solution of the model.¹² 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; its virtue is that the model is quantitatively plausible enough that, for example, it might actually be usable by policymakers who wanted to assess the likely dynamics entailed by alternative fiscal policy options.

Given our choice to embrace the challenge of matching micro data, it was essential to keep the rest of the model as simple as possible, in the spirit of Akerlof and Yellen (1985), Cochrane (1991), Mankiw and Reis (2002) and as forcefully advocated by Browning and Crossley (2001). In pursuit of such simplicity, we adopt the Calvo (1983)-like framework of Carroll (2003) in which updating is a Poisson event.¹³

Inattention is not the only alternative to habits as an explanation for excess smoothness. Information itself can be imperfect, even for a perfectly attentive consumer. The seminal work contemplating this possibility was by Muth (1960), whose most direct descendant in the consumption literature is Pischke (1995) (building also on Lucas

¹¹This 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).

¹²A fascinating study that does not fit neatly into this analysis is Johnson, Parker, and Souleles (2006), who show that spending in the U.S. responded strongly to the random idiosyncratic variation in the timing of tax refunds. Their results might possibly be consistent with a model in which households can see the contents of their bank accounts but do not pay attention to tax policy, which has the flavor of the model proposed here. Similarly, Kueng (2015) argues that excess sensitivity of consumption is consistent with households following near-rational plans and that to model macroeconomic policies, such as economic stimulus programs, near-rational alternatives can perform better than standard consumption models.

¹³Browning and Collado (2001) speculate that the differing results in the micro literature can be resolved if consumers are not perfectly attentive even to all the details of their own personal income processes, an explanation which could easily be interpreted using framework proposed here.

(1973)). 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 that when a model of this kind is calibrated to the actual empirical data, it generates only a modest amount of excess smoothness, far less than exhibited by the empirical data.

Moving from theory to evidence, there is an interesting and growing literature that uses expectations data from surveys in an attempt to directly measure sluggishness in expectations dynamics.¹⁴ 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 (2017a) and Fuhrer (2017b) find that even for professional forecasters, forecast revisions are explainable using lagged information, which would not be the case under perfect information processing.

3 A Quadratic Utility ‘Toy Model’

Here we briefly introduce concepts and notation, and motivate the key result using a simple framework with quadratic utility. We start with the classic Hall (1978) random walk model, with the standard assumption of time separable utility and geometric discounting by factor β . Overall wealth \mathbf{o} (the sum of human and nonhuman wealth) evolves according to the dynamic budget constraint

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

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

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

¹⁴We omit a fuller survey of this interesting literature because expectations data will not be our focus here.

equation:

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

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

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

Consumers spend

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

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

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

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 these are the persons who 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 $\tilde{\mathbf{o}}$ as the consumer's *perception* of wealth:

$$\tilde{\mathbf{o}}_{t+j} \equiv \mathbb{E}_t[\mathbf{o}_{t+j}] = \mathbf{o}_t \quad \text{for } 1 \leq j < n,$$

the consumer spends according to perceived wealth so that

$$\mathbf{c}_{t+j} = (r/R)\tilde{\mathbf{o}}_{t+j} = (r/R)\mathbf{o}_t = \mathbf{c}_t \quad \text{for } 1 \leq j < n.$$

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

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

$$\begin{aligned}
\mathbf{o}_{t+2} &= \mathbf{o}_{t+1} + \zeta_{t+2} = \mathbf{o}_t + \zeta_{t+1}\mathbf{R} + \zeta_{t+2} \\
&\vdots \\
\mathbf{o}_{t+n} &= \mathbf{o}_t + \underbrace{\sum_{s=1}^n \mathbf{R}^{n-s} \zeta_{t+s}}_{\equiv \Delta^n \mathbf{o}_{t+n}},
\end{aligned}$$

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

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

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).

Aggregation

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

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

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:

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

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

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

where per-capita consumption $C_{t+1}^\pi = C_t$ because the nonupdaters at time $t + 1$ are a random subset of the population at time t . The first difference of (2) yields

$$\Delta C_{t+1} = (1 - \Pi)\Delta C_t + \underbrace{\Pi\Delta C_{t+1}^\pi}_{\equiv \varepsilon_{t+1}}$$

and Appendix C.1 shows that ε_{t+1} is approximately mean zero.¹⁵ 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 6. While the details of the informational friction is different in the more realistic models we will set up in Section 4, the same logic and quantitative result holds: the serial correlation of consumption growth approximately equals the proportion of non-updaters.

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

4 Real Models

One of the lessons of the consumption literature after Hall (1978) is that his simplifying assumptions (quadratic utility, perfect capital markets, $R\beta = 1$) 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, $u(\mathbf{c}) = \mathbf{c}^{1-\rho}/(1-\rho)$. When utility is not quadratic, solution of the model requires specification of the exact stochastic structure of the income and transition processes.

¹⁵Intuitively, this term mainly reflects the behavior of updating consumers, and is therefore unpredictable with respect to sufficiently delayed information.

Below, we present two models that will be used to simulate the economy under frictionless and sticky expectations. First, 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. Second, we extend the SOE model to a heterogeneous agents dynamic stochastic general equilibrium (closed-economy) model that endogenizes factor returns, at the cost of a more burdensome computational task.¹⁶

4.1 General Modeling Assumptions

Several features are common across all our models. 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 . These 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 immediately replaced, while their assets are distributed among surviving households in proportion to the recipient’s wealth.

4.1.1 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 $\Upsilon = 1 - \delta$ intact, and as usual the effectiveness of labor depends on the level of aggregate labor productivity.

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

¹⁶Appendix A 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 recognizable representative agent model nor a respectable heterogeneous agents model, which may reduce its appeal to both audiences.

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:

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

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)).¹⁷ We introduce these slow-moving productivity growth rates not just for realism but also because we need to perform, in our simulated data, exercises like those Campbell and Mankiw (1989) performed in 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:

$$p_{t+1,i} = p_{t,i}\psi_{t+1,i}, \quad (4)$$

and like their aggregate counterparts, idiosyncratic permanent productivity shocks are mean one iid ($\mathbb{E}_t[\psi_{t+n,i}] = \mathbb{E}_t[\Psi_{t+n}] = 1 \forall n > 0$).¹⁸ Total labor productivity for

¹⁷We 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.

¹⁸Because we did not have any modeling purpose for which we needed a slow-moving persistent component to idiosyncratic productivity growth, we omitted it. If we wanted a completely parallel representation of the aggregate and idiosyncratic processes, we could introduce a term $\phi_{t,i} \equiv 1 \forall t$.

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, this contributes effective labor equal to:

$$\ell_{t,i} = \underbrace{\theta_{t,i}}_{\equiv \theta_{t,i}} \underbrace{\Theta_t}_{\equiv \Theta_t} \underbrace{p_{t,i} P_t}_{\equiv \mathbf{p}_{t,i}}. \quad (5)$$

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 aggregate transitory shocks, $\theta_{t,i}$ is mean one and iid, so that $\mathbb{E}_t[\theta_{t+n,i}] = \mathbb{E}_t[\Theta_{t+n}] = 1 \forall n > 0$. The idiosyncratic transitory shock has a minimum possible value of 0 (corresponding to an unemployment spell) which occurs with a small finite probability φ . This has the effect of imposing a ‘natural borrowing constraint’ (cf. Zeldes (1989b)) at zero.

4.1.2 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 \mathbf{p} (likewise, the combined permanent shock is boldface $\boldsymbol{\psi}_{t,i} \equiv \psi_{t,i} \Psi_t$). However, a key feature of the models used here is that a household does not necessarily know the true value of the aggregate productivity state variables (P_t, Φ_t), as they might not have (stochastically) observed it in the current period. Instead, each household has *perceptions* about the aggregate state ($\tilde{P}_{t,i}, \tilde{\Phi}_{t,i}$). Our key behavioral assumption is twofold:

1. Household agents always act *as if* their perception of the aggregate state ($\tilde{P}_{t,i}, \tilde{\Phi}_{t,i}$) were the true aggregate state (P_t, Φ_t).
2. Households form their perception of the aggregate state according to the *expectation* of today’s state given the last observed aggregate state.

Given the assumption that productivity growth 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:

$$(\tilde{P}_{t,i}, \tilde{\Phi}_{t,i}) = \mathbb{E}_{t-n} [(P_t, \Phi_t) | (P_{t-n}, \Phi_{t-n})] = (\Phi_{t-n}^n P_{t-n}, \Phi_{t-n}). \quad (6)$$

That is, our assumed random walk in productivity growth means that the household believes that aggregate productivity has grown at the last observed growth rate for the past n periods.¹⁹ For households who observed the true aggregate state this period, $n = 0$ and thus (6) says that $(\tilde{P}_{t,i}, \tilde{\Phi}_{t,i}) = (P_t, \Phi_t)$. The household perceives that their overall permanent productivity level is $\tilde{p}_{t,i} = p_{t,i} \tilde{P}_{t,i}$.

Households in our models always correctly observe the *level* of all real variables—they are able to read their bank statement and paycheck. But (as will be shown below) consumers' optimal behavior in the frictionless model depends on the *ratios* of those real variables to productivity. That is, for some state variable \mathbf{x} (like market wealth), the optimal choice would depend on $x \equiv \mathbf{x}/\mathbf{p}$, 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 $X \equiv \mathbf{X}/P$.

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

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.

4.1.3 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 indefi-

¹⁹Because of the boundedness of the random walk in productivity growth, this holds only approximately.

nately with time. To avoid this inconvenience, we make the Blanchard (1985) assumption: Each consumer faces a constant probability of mortality of D (with complementary survival probability \mathcal{D}). We track death events using a binary indicator:

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

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 $D = \int_0^1 \mathbf{d}_{t,i} di$.

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 $\int_0^1 p_{t,i} di = 1$.²⁰ Our earlier equation (4) for the idiosyncratic productivity transition rule for the inhabitant of location i on the number line is thus adjusted to:²¹

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

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 value of the capital income it yields):

$$\mathbf{m}_{t,i} = \underbrace{W_t \ell_{t,i}}_{\equiv \mathbf{y}_{t,i}} + \underbrace{\mathcal{R}_t}_{\uparrow+r_t} \mathbf{k}_{t,i}. \quad (7)$$

The transition process for \mathbf{m} is broken up, for convenience of analysis, into three steps.

²⁰The dynamics and steady state of the population variance of the idiosyncratic component of permanent income p are derived in Appendix C.2. The variance will exist so long as $\mathcal{D}\mathbb{E}[\psi^2] < 1$, which imposes a loose restriction on the magnitude of the idiosyncratic permanent shocks.

²¹The 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.

‘Assets’ at the end of the period are market resources minus consumption:

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

Next period’s capital is determined from this period’s assets via

$$\mathbf{k}_{t+1,i} = \begin{cases} \mathbf{a}_{t,i}/\mathcal{D} & \text{if } \mathbf{d}_{t+1,i} = 0 \\ 0 & \text{if } \mathbf{d}_{t+1,i} = 1, \end{cases}$$

where the first row’s division of \mathbf{a} by the survival probability \mathcal{D} reflects returns to survivors from the Blanchardian insurance scheme in which the dying agents’ assets are distributed to the survivors. More compactly we can write:

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

4.1.4 Aggregation

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

$$\mathbf{K}_t = \int_0^1 \mathbf{k}_{t,i} di = \int_0^1 ((1 - \mathbf{d}_{t,i})\mathbf{a}_{t-1,i}/\mathcal{D}) di = \int_0^1 \mathbf{a}_{t-1,i} di = \mathbf{A}_{t-1}. \quad (10)$$

The third equality holds because $\mathcal{D}^{-1} \int_0^1 (1 - \mathbf{d}_{t,i}) di = 1$ and $\mathbf{d}_{t,i}$ 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

$$\mathbf{L}_t = \int_0^1 \boldsymbol{\ell}_{t,i} di = \Theta_t P_t. \quad (11)$$

Aggregate market resources can be written as per-capita resources of the survivors times their population mass \mathcal{D} , plus per-capita resources of the newborns times their population mass \mathbf{D} :

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

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

The productivity-normalized version of (12) says that

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

Because the households in our model do not necessarily observe the true aggregate productivity level, their perception of normalized aggregate market resources is

$$\widetilde{M}_{t,i} = \mathbf{M}_t/\widetilde{P}_{t,i} = (P_t/\widetilde{P}_{t,i})M_t. \quad (14)$$

We will sometimes refer to the factor $P_t/\widetilde{P}_{t,i}$ as the household's ‘productivity misperception,’ the scaling factor between actual and perceived market resources. As discussed below, this same misperception factor applies to individual market resources as well.

4.2 Small Open Economy (SOE) Model

Our first realistic model considers a small open economy with perfect international capital mobility, so that factor prices \mathbf{r}_t and \mathbf{W}_t are exogenously determined (at constant values \mathbf{r} and \mathbf{W}). These assumptions permit a partial equilibrium analysis using only the solution to the individual households’ problem. The frictionless consumer’s state variables are simply $(\mathbf{m}_{t,i}, p_{t,i}, P_t, \Phi_t)$. Because 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.

4.2.1 Model and Solution

The household’s problem in levels can be written in Bellman form as:²²

$$\mathbf{v}(\mathbf{m}_{t,i}, p_{t,i}, P_t, \Phi_t) = \max_{\mathbf{c}_{t,i}} \left\{ u(\mathbf{c}_{t,i}) + \beta \mathbb{E}_t \left[(1 - \mathbf{d}_{t+1,i}) \mathbf{v}(\mathbf{m}_{t+1,i}, p_{t+1,i}, P_{t+1}, \Phi_{t+1}) \right] \right\}. \quad (15)$$

Our assumption that the aggregate and idiosyncratic productivity levels both reflect a combination of transitory and purely permanent components now permits us to make

²²Subject to definitions (3), (4), (5), (7), (8) and (9).

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 $\mathbf{p}_{t,i}^{1-\rho} = (p_{t,i}P_t)^{1-\rho}$ while converting to normalized variables as above (e.g., $m_{t,i} = \mathbf{m}_{t,i}/\mathbf{p}_{t,i}$).²³ This yields the normalized form of the problem, which has only $m_{t,i}$ and Φ_t as state variables:

$$\begin{aligned} v(m_{t,i}, \Phi_t) &= \max_{c_{t,i}} \left\{ u(c_{t,i}) + \mathcal{D}\beta \mathbb{E}_t \left[(\Phi_{t+1}\boldsymbol{\psi}_{t+1,i})^{1-\rho} v(m_{t+1,i}, \Phi_{t+1}) \right] \right\} & (16) \\ &\text{s.t.} \\ a_{t,i} &= m_{t,i} - c_{t,i}, \\ k_{t+1,i} &= a_{t,i}/(\mathcal{D}\Phi_{t+1}\boldsymbol{\psi}_{t+1,i}), \\ m_{t+1,i} &= \mathcal{R}k_{t+1,i} + \mathbf{W}\boldsymbol{\theta}_{t+1,i}. \end{aligned}$$

Defining $\mathbf{R} = \mathcal{R}/\mathcal{D}$, the main requirement for this problem to have a solution is an impatience condition:²⁴

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

Designating the converged normalized consumption function that solves (16) as $c(m, \Phi)$, the level of consumption for the frictionless consumer can be obtained²⁵ from

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

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

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

²³This is well understood in the literature; for a full exposition, see, e.g., Carroll (2016).

²⁴Other 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} adjusted for both of the influences (δ and \mathcal{D}) that intervene between the amount of assets with which the consumer ends period t and the amount of productive capital owned when capital income is received in period $t+1$.

²⁵Appendix B.1 provides details for numerically solving the model.

4.2.2 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 period t :²⁶

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

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:

$$(\tilde{P}_{t,i}, \tilde{\Phi}_{t,i}) = \begin{cases} (P_t, \Phi_t) & \text{if } \pi_{t,i} = 1 \\ (\tilde{\Phi}_{t-1,i}, \tilde{P}_{t-1,i}, \tilde{\Phi}_{t-1,i}) & \text{if } \pi_{t,i} = 0. \end{cases} \quad (17)$$

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 $c(m, \Phi)$ are the household's *perceived* normalized market resources $\tilde{m}_{t,i} = \mathbf{m}_{t,i}/\tilde{p}_{t,i} = (P_t/\tilde{P}_{t,i})m_{t,i}$ and perceived aggregate productivity growth $\tilde{\Phi}_{t,i}$. The household chooses their level of consumption by:

$$\mathbf{c}_{t,i} = \tilde{p}_{t,i}c(\tilde{m}_{t,i}, \tilde{\Phi}_{t,i}) = \hat{\mathbf{c}}(\mathbf{m}_{t,i}, p_{t,i}, \tilde{P}_{t,i}, \tilde{\Phi}_{t,i}). \quad (18)$$

The behavior of the 'sticky expectations' consumer converges to that of the frictionless consumer as $\Pi \uparrow 1$; conveniently, we can use the same simulation code for both kinds of consumers by simply setting $\Pi = 1$ to generate the behavior of the frictionless economy.

Because households in our model never misperceive the *level* of their own market resources ($\tilde{\mathbf{m}}_{t,i} = \mathbf{m}_{t,i}$), they can never choose consumption that would violate the budget constraint.²⁷ But their misperceptions of aggregate permanent income do cause

²⁶For 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 5 for details.

²⁷Households 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. Our assumption is simply that households neither perceive nor attempt to extract an estimate of the decomposition of that aggregate state into transitory and permanent components.

them to make systematic errors. See below for calculations showing that for the value of Π that we estimate, those errors are very small.

4.3 Dynamic Stochastic General Equilibrium (HA-DSGE) Model

Our second model relaxes the simplifying assumption 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 growing) class of heterogeneous agent DSGE models.

4.3.1 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:

$$\begin{aligned} W_t &= \frac{\partial \mathbf{Y}_t}{\partial \mathbf{L}_t} = (1 - \gamma)(\mathbf{K}_t / \mathbf{L}_t)^{-\gamma}, \\ r_t &= \frac{\partial \mathbf{Y}_t}{\partial \mathbf{K}_t} = \gamma \underbrace{(\mathbf{K}_t / \mathbf{L}_t)^{\gamma-1}}_{= K_t / \Theta_t}. \end{aligned} \tag{19}$$

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

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:

$$\mathbb{E}[\mathbf{A}_t] = \aleph(\mathbf{M}_t, \Phi_t = \Phi_j) \equiv \exp(\kappa_{j,0} + \kappa_{j,1} \log(\mathbf{M}_t)). \tag{20}$$

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:²⁸

$$\mathbf{v}(\mathbf{m}_{t,i}, \mathbf{M}_t, p_{t,i}, P_t, \Phi_t) = \max_{\mathbf{c}_{t,i}} \left\{ u(\mathbf{c}_{t,i}) + \beta \mathbb{E} \left[(1 - \mathbf{d}_{t,i}) \mathbf{v}(\mathbf{m}_{t+1,i}, \mathbf{M}_{t+1}, p_{t+1,i}, P_{t+1}, \Phi_{t+1}) \right] \right\}. \quad (21)$$

As in the SOE model, the household's problem can be normalized by the combined productivity level $\mathbf{p}_{t,i}$, reducing the state space by two continuous dimensions. Dividing (21) by $\mathbf{p}_{t,i}^{1-\rho}$ and substituting normalized variables, the reduced problem is:

$$\begin{aligned} v(m_{t,i}, M_t, \Phi_t) &= \max_{c_{t,i}} \left\{ u(c_{t,i}) + \beta \mathcal{D} \mathbb{E} \left[(\Phi_{t+1} \boldsymbol{\psi}_{t+1,i})^{1-\rho} v(m_{t+1,i}, M_{t+1}, \Phi_{t+1}) \right] \right\} \\ &\text{s.t.} \\ a_{t,i} &= m_{t,i} - c_{t,i}, \\ k_{t+1,i} &= a_{t,i} / \mathcal{D}, \\ m_{t+1,i} &= \mathcal{R}_{t+1} k_{t+1,i} / (\Phi_{t+1} \boldsymbol{\psi}_{t+1,i}) + \boldsymbol{\theta}_{t+1,i} \mathbf{W}_{t+1}. \end{aligned} \quad (22)$$

Because household beliefs about the aggregate saving rule are linear in logs, (20) 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 $c(m, M, \Phi)$ and an aggregate saving rule \aleph such that when all households believe \aleph , the solution to their individual problem (22) 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).²⁹

4.3.2 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 4.2.2: Because the level of \mathbf{M}_t now affects

²⁸Subject to definitions (3), (4), (5), (7), (8), (9), (10), (11), (12), (19) and (20).

²⁹Details are in Appendix B.1.2.

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

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

Households who misperceive the aggregate productivity state will incorrectly predict aggregate saving at the end of the period, and thus aggregate capital and the distribution of factor prices next period.³⁰

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.

5 Calibration

To calculate the quantitative consequences of sticky expectations, we must calibrate model parameters to match the received wisdom of the literature about empirical magnitudes. We begin by calibrating market-level and preference parameters by standard methods, then specify additional parameters to characterize the idiosyncratic income shock distribution.

5.1 Macroeconomic Calibration

We assume a coefficient of relative risk aversion of 2, in the middle of the range usually considered plausible. The quarterly depreciation rate δ is calibrated by assuming annual

³⁰This incorrect prediction is short-lived: all households will learn the true *levels* of next period's aggregate capital and output.

depreciation of 6%, i.e., $\Upsilon^4 = 0.94$. Capital's share in aggregate output takes its usual value of $\gamma = 0.36$.

We calibrate our aggregate income process as follows. We set the variances of the quarterly transitory and permanent shocks at the approximate values respectively:

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

to match the high persistence in aggregate labor income.³¹ These values are consistent with papers such as Jermann (1998), Boldrin, Christiano, and Fisher (2001), and Chari, Kehoe, and McGrattan (2005), considered standard exercises in the RBC literature. These authors model the state of technology as either a highly persistent AR(1) process or a random walk; the underlying calibrations come from the autocorrelation properties of measured aggregate dynamics which closely correspond to our specification of the income process. We calibrated the process for aggregate productivity growth Φ to match measured U.S. productivity data. A Markov process with eleven states ranging between -3.0% and $+3.0\%$ (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 Staiger, Stock, and Watson (2001), Figure 1.9 and Fernald, Hall, Stock, and Watson (2017), Figure 10.

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 Υ , this choice implies values for the other steady-state characteristics of the PF-DSGE model:³²

$$\check{K} = 12^{1/(1-\gamma)},$$

³¹We measure labor income using U.S. NIPA data as wages and salaries plus transfers minus personal contributions for social insurance.

³²In the SOE model, we fix the interest factor \mathcal{R} and wage rate \mathbf{W} to these PF-DSGE steady state values.

$$\begin{aligned}\check{W} &= (1 - \gamma)\check{K}^\gamma, \\ \check{R} &= \bar{\tau} + \gamma\check{K}_t^{\gamma-1}.\end{aligned}$$

A perfect foresight representative agent would achieve this steady state if his discount factor satisfied $\check{R}\beta = 1$. For the HA-DSGE model, we thus set the discount factor to $\beta_{DSGE} = \check{R}^{-1}$, roughly matching the target capital-to-output ratio.³³ For the SOE model we choose a much lower value of β (0.97). This results in agents with wealth holdings around the median observed in the data.³⁴ The two values of β are chosen to span the rather wide range of calibrations found in the micro and macro literatures. Further alternative calibrations are possible, but experimentation has indicated that results are not sensitive to such choices.

5.2 Calibration of Idiosyncratic Shocks

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

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

These figures are conservative in comparison with standard raw estimates from the micro data;³⁵ 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 job mobility and participation decisions reduces estimates of the permanent variance somewhat; and using very well-measured Danish administrative data, Nielsen and Vissing-

³³The 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.

³⁴The 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.

³⁵See 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>.

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).

Since the variance of the annual permanent innovation is four times the variance of the quarterly innovation, this calibration implies that 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 the 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.³⁶

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 $D = 0.005$ which implies an expected working life of 50 years; results are not sensitive to plausible alternative values of this parameter, so

³⁶For further details of how we construct our quarterly numbers, see Appendix C.3.

long as the life length is short enough to permit a stationary distribution of idiosyncratic permanent income.

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.

5.3 Equilibrium Characteristics

This section briefly characterizes some of the equilibrium characteristics of the solutions to the models under the parameters specified above. Results are reported in Table 2.

Note first the considerable difference between the mean level of assets in the HA-DSGE and SOE models (first row of the table). As indicated above, 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.

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.

6 Results

The calibrated models 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.

6.1 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:

$$\Delta \log \mathbf{C}_{t+1} = \varsigma + \nu \mathbb{E}_t[\mathbf{r}_{t+1}] + \mu X_t + \epsilon_{t+1}, \quad (23)$$

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 Section 4.3), simulation analysis 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 (23) can be reformulated with no loss of statistical power as

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

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 some 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:

$$\Delta \log \mathbf{C}_{t+1} = \varsigma + \chi \Delta \log \mathbf{C}_t + \eta \mathbb{E}_t[\Delta \log \mathbf{Y}_{t+1}] + \alpha A_t + \epsilon_{t+1}. \quad (24)$$

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. Sommer (2007) interprets the serial correlation of consumption growth as reflecting habit formation.³⁷ However, Sommer's choice of instruments, estimation methodology, and tests do not correspond precisely to our purposes here, so we have produced our own estimates using U.S. data.

In Table 3 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.³⁸ Three points are worth emphasizing here.

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.³⁹ 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. 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.

³⁷Weber (2002) makes a similar point using a different methodology.

³⁸Carroll, Sommer, and Slacalek (2011) provide analogous estimates in international data.

³⁹In particular, imputed rent on housing is the largest component of services spending, but even a careful examination of the published documentation on the construction of this data series leaves it unclear what causes quarterly variation in the imputations, since the most important data sources seem to be annual; similar concerns apply to other important types of services spending. See Wilcox (1992) for a detailed discussion of the data measurement issues and their connection to theoretical questions.

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)).⁴⁰

The table 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 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 large point

⁴⁰An 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.

estimate, while the predictable income growth term becomes statistically insignificant (and economically small).

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.

6.2 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 (for a proof that such a target value will exist in models of the class we are using, see Carroll (2016)). Now 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 declining asymptotically 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.

It 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:⁴¹

$$\Delta \log \mathbf{c}_{t+1,i} = \varsigma + \chi \Delta \log \mathbf{c}_{t,i} + \eta \mathbb{E}_{t,i}[\Delta \log \mathbf{y}_{t+1,i}] + \alpha \bar{a}_{t,i} + \epsilon_{t+1,i}. \quad (25)$$

Results for the frictionless model are presented the upper panel of Table 4.⁴² 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

⁴¹Details of the simulation procedure are in Appendix B.2.

⁴²An analogous table for the HA-DSGE model shows similar results and is available upon request.

corresponds to the thought experiment above, in which consumption growth should have *some* positive serial correlation, the magnitude of that correlation is only 0.019.⁴³

The second row of the table presents the results of a Campbell and Mankiw (1989)-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 existing micro literature has typically found much larger Campbell–Mankiw coefficients than ours. However, much of that literature has made little effort to determine the extent to which the predictable component of income growth reflects permanent underlying growth rates like Φ_t versus the extent to which that predictability comes from purely transitory movements. If we were to use instruments that had no power for the transitory component but did have power for Φ_t , our estimated η coefficient would be close to 1 (because consumption growth in models of this kind settles down in the long run to something close to the underlying growth rate of permanent income). Thus, our view is that little can be learned from the magnitude of the η coefficient.

The third row confirms the proposition articulated above: For people with very low

⁴³We 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.

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 final 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.

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.

6.3 Simulated Small Open Economy Empirical Estimation

Table 3 presents the results that an econometrician would obtain from estimating an equation like (24) using aggregate data generated by the same models whose micro results are presented in Table 4. In short, it shows that even though simulated households with sticky expectations do not exhibit any meaningful predictability of idiosyncratic consumption growth, aggregate consumption growth in an economy populated by such consumers exhibits a high degree of serial correlation (similar to that in empirical data).

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 3. Because there is some variation in coefficient estimates depending on the random number generator's seed, we repeat the simulation exercise 100 times. Table 5 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.⁴⁴

Finally, for comparison to empirical results, we take into account Sommer (2007)’s argument (based on Wilcox (1992)) that transitory components of aggregate spending⁴⁵ (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 :

$$\mathbf{C}_t^* = \mathbf{C}_t \times \xi_t. \quad (26)$$

The standard deviation of ξ_t is set to the value that would cause the difference between the OLS and IV estimates of χ in the univariate regression in Table 3 ($\chi^{OLS} = 0.468$ and $\chi^{IV} = 0.830$): $\text{std}(\log(\xi)) = 0.375 \times \text{std}(\Delta \log \mathbf{C}_t)$.

The top panel of Table 5 estimates (24) 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

⁴⁴Specifically, 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 $\Delta_8 \log x_{t-2} \equiv \log x_{t-2} - \log x_{t-10}$.

⁴⁵It 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.

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 (the lower panel of the table), 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 0.261 to 0.263.

6.4 Simulated General Equilibrium Model Estimation

Table 6 reports the results of estimating regression (24) on data generated from the HA-DSGE model of Section 4.3; results are substantially the same as the previous analysis for the SOE model.⁴⁶

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.

7 The Utility Costs of Sticky Expectations

To this point, we have taken Π to be exogenous (though reasonably calibrated), but the probability of updating could depend at least partly on costs and benefits, as in ‘rational’ inattention models. In this section, we briefly examine the tradeoffs by imagining that newborns make a once-and-for-all choice of their idiosyncratic value of Π , yielding an

⁴⁶Essentially similar results are obtained if we assume that households have heterogeneous discount factors, in the style of Carroll, Slacalek, Tokunaka, 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.

intuitive approximating formula for the optimal updating frequency.⁴⁷ We then conduct a numerical exercise to compute the cost of stickiness for the calibrated models. The utility costs of having Π equal to our calibrated value of 0.25, rather than updating every period, are on the order of one two-thousandth of lifetime consumption, so that even small informational costs would justify updating aggregate information only occasionally.

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 $v(W_t, \cdot)$. 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:

$$v(W_t, \cdot) \approx \hat{v}(W_t, \cdot) - \kappa \sigma_{\Psi}^2, \quad (27)$$

where $\hat{v}(W_t, \cdot)$ 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 2 below) that the effect of sticky expectations is approximately to reduce value by an amount proportional to the inverse of the updating probability:

$$\tilde{v}(W_t, \cdot) \approx \hat{v}(W_t, \cdot) - (\kappa/\Pi) \sigma_{\Psi}^2. \quad (28)$$

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

⁴⁷For a more thorough theoretical examination of the tradeoffs in a related model, see Reis (2006a).

- 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.⁴⁸ The newborn's objective is therefore to choose the Π that solves:

$$\max_{\Pi} \dot{v}(W_t, \cdot) - (\kappa/\Pi)\sigma_{\Psi}^2 - \iota\Pi. \quad (29)$$

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:

$$\Pi = (\kappa/\iota)^{0.5}\sigma_{\Psi}. \quad (30)$$

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.⁴⁹ Specifically, we calculate the percentage loss of permanent income that would make a newborn indifferent between being frictionless while taking the loss versus having sticky expectations.⁵⁰

Using notation from the theoretical exercise above, define a newborn's average lifetime

⁴⁸Think 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.

⁴⁹Additional numeric details of the cost of stickiness calculation can be found in Appendix B.3.

⁵⁰The 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.

(normalized) value at birth under frictionless and sticky expectations as respectively:

$$\bar{v}_0 \equiv \mathbb{E} [v(\mathbf{W}_t, \cdot)], \quad \tilde{\bar{v}}_0 \equiv \mathbb{E} [\tilde{v}(\mathbf{W}_t, \cdot)],$$

where the expectation is taken over the distribution of state variables other than $m_{t,i}$ that an agent might be born into (as well as the wage rate, in the HA-DSGE model). 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 expectations can then be calculated as:

$$\omega = 1 - \left(\frac{\tilde{\bar{v}}_0}{\bar{v}_0} \right)^{\frac{1}{1-\rho}}. \quad (31)$$

The bottom row of Table 2 reports the cost of stickiness for the SOE and HA-DSGE models. A newborn in either 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 (2006a), 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.

Now that we have explained how to compute the cost of stickiness numerically, we can test our supposition in equation (28) that the cost of stickiness might have a roughly inverse linear relationship to Π . Figure 2 plots numerically computed ω for various values of Π^{-1} and is close to linear, as we speculated.

8 Muth–Lucas–Pischke and Reis (2006a) Redux

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.

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 (perfectly attentively) 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 the analysis of micro 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 do signal extraction on it.

Pischke calibrates his model with micro data in which he calculates that transitory shocks vastly outweigh permanent shocks.⁵¹ 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. (See below for a more precise formulation of this point).

In its assumption that consumers fail to perceive aggregate shocks immediately and fully, Pischke's model resembles ours. However, few papers in the literature after Pischke (1995) have adopted his assumption that households have no idea, when an idiosyncratic income shock occurs, whether it is transitory or permanent. Especially in the last decade

⁵¹Pischke'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.

or so, the literature instead has almost always assumed that consumers can perfectly perceive the transitory and permanent components of their income.⁵²

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 (and more plausibly than under Pischke’s assumption of complete ignorance), such confusion could generate excess smoothness. To see how, note that in the Muth framework, agents update their estimate of permanent income according to an equation of the form:⁵³

$$\hat{P}_{t+1} = \Pi Y_{t+1} + (1 - \Pi)\hat{P}_t, \quad (32)$$

We can now consider the dynamics of aggregate consumption in response to the arrival of an aggregate shock that (unbeknownst to the consumer) is a permanent shock. 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

⁵²The presumption that permanent idiosyncratic shocks are easily observed is bolstered by the work of Low, Meghir, and Pistaferri (2010), who find that most ‘permanent’ changes to individual income are concentrated in periods when people change jobs or become unemployed or reemployed. It is simply not plausible to assume that consumers are not aware their income has permanently changed when they take a new job or are fired from an existing one. Furthermore, 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 (contra Pischke) 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, it seems hard to accept Pischke’s crucial assumption that consumers treat all income shocks identically because they cannot perceive whether any given shock is transitory or permanent. A further motivation for our choice to assume that consumers perceive their idiosyncratic shocks is that our strategy has been to deviate as little as possible from the well-understood benchmark models in the micro consumption literature, which has become sufficiently widely used that they are the natural base upon which to build. A final piece of evidence comes from some newly collected data from the Federal Reserve Bank of New York’s *Survey of Consumer Expectations*. Karahan, Mihaljevich, and Pillosoph (2017) report on results from an experiment in which consumers were asked to forecast their future income. The actual realizations of the same consumer’s income were collected in a subsequent wave of the survey. The principal result is that consumers’ mean forecast error was very close to zero. The authors of the post (entitled “Understanding Permanent and Temporary Income Shocks”) present the results as a confirmation of modern models’ standard assumption that consumers have accurate perceptions of the dynamics of their income.

⁵³ \hat{P}_t is used to denote that households do an optimal signal-extraction (as opposed to having sticky expectations resulting in \tilde{P}_t).

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 $\Delta \mathbf{C}_{t+n+1} = (1 - \Pi)\Delta \mathbf{C}_{t+n} + \epsilon_{t+n}$.⁵⁴

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 other variables in addition to the lagged history of aggregate income itself (and that those variables have useful predictive power).⁵⁵

We now turn to the question of whether the Muth–Lucas–Pischke story is a good *quantitative* explanation of the *size* of aggregate excess smoothness. Appendix C.4 shows that, defining the signal-to-noise ratio $\tau = \sigma_{\Psi}/\sigma_{\Theta}$, Muth's derivations imply that the

⁵⁴The 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 $(1 - \Pi) \times 0 = 0$.

⁵⁵This 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.

optimal updating coefficient is:⁵⁶

$$\Pi = \varphi\sqrt{1 + \varphi^2/4} - (1/2)\varphi^2 \quad (33)$$

Plugging our calibrations of σ_Ψ and σ_Θ from section 5 into (33), 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.⁵⁷

Considerations similar to the foregoing apply, at least to some degree, to the Reis (2006a) model. Moreover, that model has a further disadvantage relative to any of the other three stories (habits, MLP, or our model). 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, the one-year horizon. But, the habit formation assumption was incorporated into macroeconomic models in large part to explain the fact that consumption growth is forecastable over extended periods—well beyond the one year horizon. A calibration of the Reis model in which consumers update once a year therefore leaves much of the original puzzle in place.⁵⁸

⁵⁶As always in a Muth-type model, the consumer is assumed to know σ_Ψ and σ_Θ .

⁵⁷A further problem with the MLP approach is that it seems implausible to assume that the typical consumer would do a sophisticated Muthian signal extraction problem (really, a special case of the Kalman filter) on a single observed aggregate variable (aggregate income), when they could do much better either by adding a few more variables that are equally easy to observe (interest rates, past consumption growth, etc). If they were really so intently focused on understanding where the aggregate economy is, they could do better yet, and much more easily, just by reading the available news stories reporting on professional forecasters’ forecasts. Discomfort with this somewhat schizophrenic set of assumptions is why in work after Lucas (1973), Lucas moved away from his assumption that microeconomic agents have imperfect information about aggregate data.

⁵⁸A final critique of the Reis model is that, while its simplifying assumptions yield elegant and intuitive results, it is too stylized to be of much practical use in answering many questions beyond its narrow focus on matching aggregate smoothness. A fiscal policymaker who wants to understand how various alternative policies might play out over the course of the business cycle would have a difficult time extracting plausible answers to questions like “how much difference would it make to consumption if our fiscal stimulus took the form of extended unemployment benefits versus a temporary income tax cut.” Our model’s much greater fidelity to the microeconomic data means it might be able to provide at least somewhat plausible answers to these kinds of questions. (Our model is an example of the broader movement in macroeconomics in recent years from the provision of stylized analytical toy models like Reis’s to the construction of models that attempt to be taken seriously for a reasonable range of quantitative as well as qualitative predictions.)

9 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. 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 behavior.

Our argument is 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.

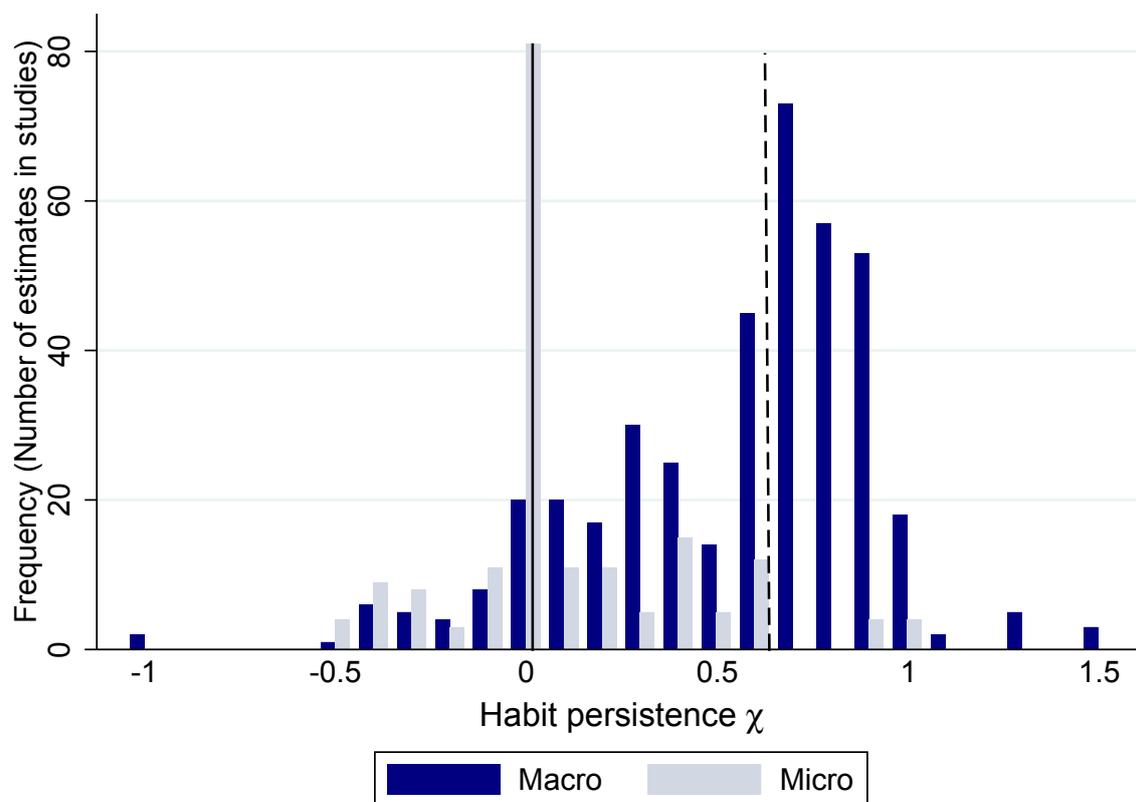
The sticky expectations assumption is more attractive for modeling consumption than for other areas where it has been more widely applied, because in the consumption context there is a well-defined utility-based metric for calculating the cost of sticky expectations (in contrast, say, with models in which households’ inflation expectations are sticky; the cost of misperceiving the inflation rate is unclear). 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 some previous papers have mooted 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 all the main stylized facts from both literatures.

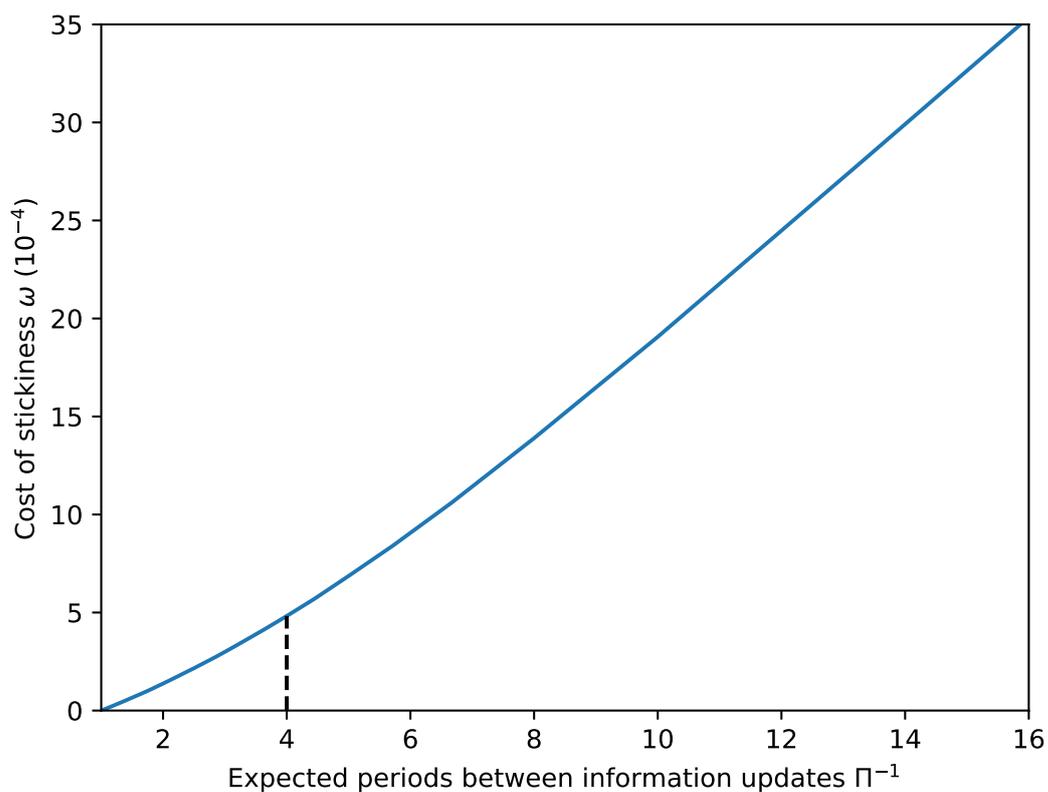
In future work, it would be interesting to enrich the model so that it has plausible implications for how the degree of attention might vary over time or across people, and to connect the model to the available expectations data (for example, measures of consumer sentiment, or measures of uncertainty 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) and crisis times (in which it *is*).

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.

Figure 2 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.

Table 1 Calibration

Macroeconomic Parameters		
γ	0.36	Capital's Share of Income
\daleth	$0.94^{1/4}$	Depreciation Factor
σ_{Θ}^2	0.00001	Variance Aggregate Transitory Shocks
σ_{Ψ}^2	0.00004	Variance Aggregate Permanent Shocks
Steady State of Perfect Foresight DSGE Model		
$(\sigma_{\Psi} = \sigma_{\Theta} = \sigma_{\psi} = \sigma_{\theta} = \wp = \mathsf{D} = 0, \Phi_t = 1)$		
$\check{K}/\check{K}^{\gamma}$	12.0	SS Capital to Output Ratio
\check{K}	48.55	SS Capital to Labor Productivity Ratio ($= 12^{1/(1-\gamma)}$)
\check{W}	2.59	SS Wage Rate ($= (1 - \gamma)\check{K}^{\gamma}$)
\check{r}	0.03	SS Interest Rate ($= \gamma\check{K}^{\gamma-1}$)
$\check{\mathcal{R}}$	1.015	SS Between-Period Return Factor ($= \daleth + \check{r}$)
Preference Parameters		
ρ	2.	Coefficient of Relative Risk Aversion
β_{SOE}	0.970	SOE Discount Factor
β_{DSGE}	0.986	HA-DSGE Discount Factor ($= \check{\mathcal{R}}^{-1}$)
Π	0.25	Probability of Updating Expectations (if Sticky)
Idiosyncratic Shock Parameters		
σ_{θ}^2	0.120	Variance Idiosyncratic Tran Shocks ($= 4 \times$ Annual)
σ_{ψ}^2	0.003	Variance Idiosyncratic Perm Shocks ($= \frac{1}{4} \times$ Annual)
\wp	0.050	Probability of Unemployment Spell
D	0.005	Probability of Mortality

Table 2 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
$\log \mathbf{y} \mathbf{y} > 0$	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 3 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			OLS	2 nd Stage	KP <i>p</i> -val
Independent Variables			or IV	\bar{R}^2	Hansen J <i>p</i> val
Nondurables and Services					
$\Delta \log \mathbf{C}_t$	$\Delta \log \mathbf{Y}_{t+1}$	A_t			
0.468*** (0.076)			OLS	0.216	
0.830*** (0.098)			IV	0.278	0.222
	0.587*** (0.110)		IV	0.203	0.439
		-0.17e-4 (5.71e-4)	IV	-0.005	0.263
0.618*** (0.159)	0.305* (0.161)	-4.96e-4* (2.94e-4)	IV	0.304	0.081
					0.181
					0.415
					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*** (0.058)			OLS	0.036	
0.762*** (0.284)			IV	0.083	0.504
	0.849** (0.357)		IV	0.061	0.727
		9.09e-4 (9.05e-4)	IV	0.008	0.398
0.620** (0.292)	0.313 (0.286)	-3.25e-4 (8.32e-4)	IV	0.077	0.731
					0.118
					0.446
					0.523
					0.821
Memo: For instruments \mathbf{Z}_t , $\Delta \log \mathbf{C}_t = \mathbf{Z}_t \zeta$, $\bar{R}^2 = 0.080$					

Notes: Data source is NIPA, 1960Q1–2016Q. Robust standard errors are in parentheses. Instruments $\mathbf{Z}_t = \{\Delta \log \mathbf{C}_{t-2}, \Delta \log \mathbf{C}_{t-3}, \Delta \log \mathbf{Y}_{t-2}, \Delta \log \mathbf{Y}_{t-3}, A_{t-2}, A_{t-3}, \Delta_8 \log \mathbf{C}_{t-2}, \Delta_8 \log \mathbf{Y}_{t-2}, \text{lags 2 and 3 of differenced Fed funds rate, lags 2 and 3 of the Michigan Index of Consumer Sentiment Expectations}\}$. 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.

Table 4 Micro Consumption Regression on Simulated Data

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

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.

Table 5 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			OLS	2 nd Stage	KP <i>p</i> -val
Independent Variables			or IV	\bar{R}^2	Hansen J <i>p</i> -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 ^{***}			OLS	0.087	
(0.066)					
0.660 ^{**}			IV	0.040	0.237
(0.309)					0.600
	0.457 ^{**}		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 , $\Delta \log \mathbf{C}_t^* = \mathbf{Z}_t \zeta$, $\bar{R}^2 = 0.039$; $\text{var}(\log(\xi_t)) = 5.99\text{e-}6$					
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			
0.508 ^{***}			OLS	0.263	
(0.058)					
0.802 ^{***}			IV	0.260	0.000
(0.104)					0.554
	0.859 ^{***}		IV	0.198	0.060
	(0.182)				0.233
		-8.26e-4 ^{**}	IV	0.066	0.000
		(3.99e-4)			0.002
0.660 ^{***}	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 , $\Delta \log \mathbf{C}_t^* = \mathbf{Z}_t \zeta$, $\bar{R}^2 = 0.260$; $\text{var}(\log(\xi_t)) = 5.99\text{e-}6$					

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 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			OLS	2 nd Stage	KP <i>p</i> -val
Independent Variables			or IV	\bar{R}^2	Hansen J <i>p</i> -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*** (0.072)			OLS	0.036	
0.476 (0.354)			IV	0.020	0.318
	0.368 (0.321)		IV	0.017	0.556
		-0.34e-4 (0.98e-4)	IV	0.015	0.107
0.289 (0.463)	0.214 (0.583)	0.01e-4 (1.87e-4)	IV	0.020	0.457
					0.000
					0.433
					0.572
					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*** (0.061)			OLS	0.223	
0.773*** (0.108)			IV	0.230	0.000
	0.912*** (0.245)		IV	0.145	0.542
		-0.97e-4 (0.56e-4)	IV	0.059	0.105
0.670*** (0.181)	0.171 (0.363)	0.12e-4 (0.86e-4)	IV	0.231	0.187
					0.002
					0.460
					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 $\mathbf{Z}_t = \{\Delta \log \mathbf{C}_{t-2}, \Delta \log \mathbf{C}_{t-3}, \Delta \log \mathbf{Y}_{t-2}, \Delta \log \mathbf{Y}_{t-3}, A_{t-2}, A_{t-3}, \Delta_8 \log \mathbf{C}_{t-2}, \Delta_8 \log \mathbf{Y}_{t-2}\}$.

References

- ABEL, ANDREW B. (1990): “Asset Prices under Habit Formation and Catching Up with the Joneses,” *American Economic Review*, 80(2), 38–42.
- 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 GUIISO, AND FRANCESCO LIPPI (2012): “Durable Consumption and Asset Management with Transaction and Observation Costs,” *American Economic Review*, 102(5), 2272–2300.
- 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.
- 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.
- 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.
- BROWNING, MARTIN, AND THOMAS F. CROSSLEY (2001): “The Life-Cycle Model of Consumption and Savings,” *Journal of Economic Perspectives*, 15(3), 3–22.
- CALVO, GUILLERMO A. (1983): “Staggered Contracts in a Utility-Maximizing Framework,” *Journal of Monetary Economics*, 12(3), 383–98.
- 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/SentAERC CarrollFuhrerWilcox.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. (1991): “The Sensitivity of Tests of the Intertemporal Allocation of Consumption to Near-Rational Alternatives,” *American Economic Review*, 79, 319–37.
- 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.
- DEATON, ANGUS S. (1992): *Understanding Consumption*. Oxford University Press, New York.
- 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>.
- (2017a): “Expectations as a Source of Macroeconomic Persistence: Evidence from Survey Expectations in a Dynamic Macro Model,” *Journal of Monetary Economics*, 86, 22–55.
- (2017b): “Intrinsic Persistence in Expectations: Evidence from Micro Data,” Presentation at NBER Summer Institute, 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.
- 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.
- 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.
- KARAHAN, FATIH, SEAN MIHALJEVICH, AND LAURA PILOSSOPH (2017): “Understanding Permanent and Temporary Income Shocks,” URL link retrieved on 03/02/2018 [here](#).
- KRUSELL, PER, AND ANTHONY A. SMITH (1998): “Income and Wealth Heterogeneity in the Macroeconomy,” *Journal of Political Economy*, 106(5), 867–896.
- KUENG, LORENZ (2015): “Explaining Consumption Excess Sensitivity with Near-Rationality: Evidence from Large Predetermined Payments,” working paper 21772, National Bureau of Economic Research.
- LOW, HAMISH, COSTAS MEGHIR, AND LUIGI PISTAFERRI (2010): “Wage Risk and Employment Over the Life Cycle,” *American Economic Review*, 100(4), 1432–1467.
- LUCAS, ROBERT E. (1973): “Some International Evidence on Output-Inflation Tradeoffs,” *American Economic Review*, 63, 326–334.
- MAĆKOWIAK, BARTOSZ, AND MIRKO WIEDERHOLT (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*.
- PISCHKE, JÖRN-STEFFEN (1995): “Individual Income, Incomplete Information, and Aggregate Consumption,” *Econometrica*, 63(4), 805–40.
- REIS, RICARDO (2006a): “Inattentive Consumers,” *Journal of Monetary Economics*, 53(8), 1761–1800.
- (2006b): “Inattentive Producers,” *Review of Economic Studies*, 73(3), 793–821.
- 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>.
- SOMMER, MARTIN (2007): “Habit Formation and Aggregate Consumption Dynamics,” *Advances in Macroeconomics*, 7(1), Article 21.
- 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.
- 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.

Appendix

A 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 4.1 to formulate the representative agent model, then presents simulated results analogous to Section 6. 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 “representative agent” under sticky expectations, as the representative household’s perception of the aggregate state is “smeared” over the state space. As presented below, the realized level of consumption represents the average level of consumption chosen by the “multiple minds” of the representative household.

A.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:⁵⁹

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

⁵⁹Subject to definitions (3), (10), (11), (12) and (19).

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:⁶⁰

$$\begin{aligned} V(M_t, \Phi_t) &= \max_{C_t} \left\{ u(C_t) + \beta \mathbb{E}_t \left[(\Phi_{t+1} \Psi_{t+1})^{1-\rho} V(M_{t+1}, \Phi_{t+1}) \right] \right\} & (34) \\ \text{s.t.} & \\ A_t &= M_t - C_t. \end{aligned}$$

Noting that the return to (normalized) end-of-period assets for next period's market resources is $\frac{dM_{t+1}}{dA_t} = \mathcal{R}_{t+1}/(\Phi_{t+1}\Psi_{t+1})$, (34) has a single first-order condition that is sufficient to characterize the solution to the normalized problem:

$$\begin{aligned} C_t^{-\rho} - \beta \mathbb{E} \left[\underbrace{\mathcal{R}_{t+1}(\Phi_{t+1}\Psi_{t+1})^{-\rho} V^M(A_t \mathcal{R}_{t+1}/(\Psi_{t+1}\Phi_{t+1}) + \Theta_{t+1} W_{t+1}, \Phi_{t+1})}_{\equiv \mathfrak{A}^A(A_t, \Phi_t)} \right] &= 0 & (35) \\ \implies C_t &= \mathfrak{A}^A(A_t, \Phi_t)^{-1/\rho}. \end{aligned}$$

The representative agent model can be solved using the endogenous grid method, following the same procedure as for the SOE model described in Appendix B.1, yielding normalized consumption function $C(M, \Phi)$.⁶¹

A.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

⁶⁰Subject to definitions (3), (13) and (19).

⁶¹The 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 (35), and the upper bound of \mathbb{A} is 120.

perceive the level of aggregate market resources \mathbf{M}_t , but might 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 \tilde{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_N^j 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:

$$\varphi_{t+1} = (1 - \Pi)\varphi_t + \Pi e_N^j \text{ when } \Phi_{t+1} = \Phi_j. \quad (36)$$

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:

$$\tilde{P}_{t+1} = ((1 - \Pi)\varphi_t \odot \tilde{P}_t + \Pi e_N^j P_{t+1}) \oslash \varphi_{t+1}. \quad (37)$$

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.⁶²

Households who perceive each growth rate Φ act as a partial representative agent, choosing their level of consumption according to their perception of normalized market resources. Defining $\tilde{M}_t^j = \mathbf{M}_t / \tilde{P}_t^j$ as perceived normalized market resources for house-

⁶²The Hadamard operators \odot and \oslash represent element-wise multiplication and division, respectively. As a numeric detail, \tilde{P}_t^j is reset to 1 when $\varphi_t^j = 0$, which would otherwise cause it to be undefined. When no households perceive growth rate Φ_j , the average perception of productivity does not exist for this state and is quantitatively irrelevant, but must exist for (37) to not fail in the next period.

holds who perceive the aggregate growth rate is Φ_j , aggregate consumption is:

$$\mathbf{C}_t = \sum_{\Phi_j \in \{\Phi\}} \tilde{P}_t^j C(\tilde{M}_t^j, \Phi_j) \varphi_t^j. \quad (38)$$

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 (36) and (37) say that $\varphi_t = e_N^j$ and $\tilde{P}_t^j = P_t$ (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.⁶³

A.3 Simulation Results

We calibrate the RA model using the same parameters as for the HA-DSGE model (see Section 5.1 and Table 1), except that there are no idiosyncratic income shocks ($\sigma_\psi^2 = \sigma_\theta^2 = \wp = 0$) and the possibility of death is irrelevant ($D = 0$). After solving the model, we utilize the same simulation procedure described in Section 6, 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 5. Both (instrumented) lagged consumption growth and expected income growth are significant predictors of aggregate consumption growth, but the ‘horse race’ regression reveals that

⁶³An alternative method for modeling sticky expectations with a representative agent would be to track the perceptions of the segments of households who last updated $n = \{0, 1, \dots, 200\}$ 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.

the predictability is dominated by serially correlated consumption growth, confirming the results of the two heterogeneous agents models.

B Numerical Methods

B.1 Solution Methods

B.1.1 Small Open Economy Solution Details

Consider the household's normalized problem in the SOE model, given in (16). 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:

$$c_{t,i}^{-\rho} - \underbrace{\mathbb{R}\beta \mathbb{E}_t \left[(\Phi_{t+1}\psi_{t+1,i})^{-\rho} v^m \left(\mathbb{R} / (\Phi_{t+1}\psi_{t+1,i}) a_{t,i} + \mathbb{W}\theta_{t+1,i}, \Phi_{t+1} \right) \right]}_{\equiv v^a(a_{t,i}, \Phi_t)} = 0 \quad (39)$$

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

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 $\mathbb{A} = \{\mathbf{\blacktriangle} a_j\}_{j=1}^{N_a}$, spanning the range values that an agent might reasonably encounter in a simulated lifetime. We use a triple-exponential grid spanning $\mathbf{\blacktriangle} a \in [10^{-5}, 40]$ 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 $c(m, \cdot) = m$, the solution for an agent who has no future.
4. Define the marginal value function $v^m(\cdot)$ as $u'(c(\cdot))$, as determined by the standard envelope condition.
5. Use the discrete approximations to the shock processes and the Markov transition matrix Ξ to compute $v^a(a_j, \Phi_k)$ for all $(a_j, \Phi_k) \in \mathbb{A} \times \{\Phi\}$.
6. Use (39) to find the level of consumption that would make ending the period with a_j in assets optimal (when aggregate growth is Φ_k): $c_{j,k} = v^a(a_j, \Phi_k)^{-1/\rho}$.
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 $c(m, \Phi_k)$ 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 B.2.

B.1.2 Dynamic Stochastic General Equilibrium Solution Details

Consider the household's normalized problem in the HA-DSGE model, given in (22). Recalling that we are taking the aggregate saving rule \aleph as given, optimal consumption is characterized by the solution to the first-order condition:

$$c_{t,i}^{-\rho} - \beta \mathbb{E} \left[\underbrace{\mathcal{R}_{t+1}(\Phi_{t+1} \boldsymbol{\psi}_{t+1,i})^{-\rho} v^a(\mathcal{R}_t a_{t,i} / (\mathcal{D} \Phi_{t+1} \boldsymbol{\psi}_{t+1,i}) + \boldsymbol{\theta}_{t+1,i} \mathbf{W}_{t+1}, M_{t+1}, \Phi_{t+1})}_{\equiv v^a(a_{t,i}, M_t, \Phi_t)} \right] = 0 \quad (40)$$

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

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 \check{M} , spanning the range $M \in [0.1\check{M}, 5\check{M}]$.
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 B.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 $\epsilon = 10^{-4}$, 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 $\mathbf{v}^a(a_j, M_k, \Phi_\ell)$ for all $(a_j, M_k, \Phi_\ell) \in \mathbb{A} \times \mathbb{M} \times \{\Phi\}$.
6. Use (40) to calculate $c_{j,k,\ell} = \mathbf{v}^a(a_j, M_k, \Phi_\ell)^{-1/\rho}$.
8. For each ℓ , construct $c(m, M, \Phi_\ell)$ by linearly interpolating $c_{j,k,\ell}$ over $m_{j,k,\ell}$ for each k , then interpolating the linear interpolations over \mathbb{M} .

B.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 $c(\cdot)$ (or $C(\cdot)$ in the representative agent model). We first describe the procedure for the SOE and HA-DSGE models from the body of the text, then summarize the simulation method for the representative agent model of Appendix A.

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 $k_{t,i} = \check{K}$ in the HA-DSGE model. Likewise, normalized aggregate capital is set to the perfect foresight steady state $K_t = \check{K}$. 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 (19) (HA-DSGE model) or assigning the constant global values (SOE).
2. Randomly select $DI = 100$ household indices i to die and be replaced: $\mathbf{d}_{i,t} = 1$. 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 III household indices to update their aggregate information: $\pi_{t,i} = 1$. Agents' perceptions $(\tilde{P}_{t,i}, \tilde{\Phi}_{t,i})$ are set according to (17).
4. The economy produces output. All agents draw idiosyncratic shocks $\psi_{t,i}$ and $\theta_{t,i}$, with newborns automatically drawing $\psi_{t,i} = \theta_{t,i} = 1$,⁶⁴ then observe their true $\mathbf{m}_{t,i}$ (and \mathbf{M}_t in the HA-DSGE model).

⁶⁴This 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.

5. Agents compute their perception of normalized idiosyncratic market resources $\tilde{m}_{t,i}$ (and aggregate $\tilde{M}_{t,i}$ in HA-DSGE).
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 2 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 5 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 A, only a few changes are necessary to the procedure above. The vectors of perceptions are initialized to $\tilde{P}_t = \mathbf{1}_{11}$ and $\varphi = e_{11}^6$, 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 (36) and (37), and aggregate consumption is determined using (38).

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.

B.3 Cost of Stickiness Calculation

After simulating a population of households using the procedure in Appendix B.2, we have a history of micro observations $\{\{\mathbf{c}_{t,i}, \mathbf{d}_{t,i}\}_{t=0}^T\}_{i=1}^I$ 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 $\mathbf{d}_{t,i} = 1$. Let $\tau_{i,n}$ be the n -th time t index where $\mathbf{d}_{t,i} = 1$; further define $N_i = \sum_{t=0}^T \mathbf{d}_{t,i}$, the number of replacement events for household index i .

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

$$v_{i,n} = P_{\tau_{i,n}}^{\rho-1} \sum_{t=\tau_{i,n}}^{\tau_{i,n+1}-1} \beta^{t-\tau_{i,n}} u(\mathbf{c}_{t,i}). \quad (41)$$

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

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

$$N_I = \sum_{i=1}^I (N_i - 1). \quad (42)$$

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

$$\bar{v}_0 = N_I^{-1} \sum_{i=1}^I \sum_{n=1}^{N_i-1} v_{i,n}. \quad (43)$$

Because we use $T = 20,000$ and $I = 20,000$, and agents live for 200 periods on average ($D = 0.005$), our simulated history includes about $N_I \approx ITD = 2$ million consumer lifetimes. The standard errors on our numerically calculated \bar{v}_0 and $\tilde{\bar{v}}_0$ 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 \bar{v}_0 and \tilde{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 \tilde{v}_0 in the HA-DSGE model the same as in the SOE model. However, \bar{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 (31) arises from the homotheticity of the household’s problem with respect to $\mathbf{p}_{t,i}$. 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 $m_{t,i} = \mathbf{W}_t$, 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 (31).

C Additional Calculations

C.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,

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

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

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

To see this, define market resources $M_t = Y_t + RA_t$ 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

$$\begin{aligned}
 \mathbf{C}_{t+1}^\pi &= (r/R)(M_{t+1} + H_{t+1}) \\
 \mathbf{C}_t^\pi &= (r/R)(M_t + H_t) \\
 \mathbf{C}_{t+1}^\pi - \mathbf{C}_t^\pi &= (r/R)(M_{t+1} - M_t + H_{t+1} - H_t) \\
 \mathbf{C}_{t+1}^\pi - \mathbf{C}_t^\pi &= (r/R)(R(Y_t + M_t - \mathbf{C}_t) - M_t + H_{t+1} - H_t). \tag{44}
 \end{aligned}$$

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

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

for some white noise ξ_{t+1} . The only difference between this expression and the RHS of (44) is the Π superscript on the \mathbf{C}_t . Thus, substituting, we get

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

So equation (3) can be rewritten as

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

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

$$\Delta \mathbf{C}_{t+1} = (1 - \Pi)\underbrace{(1 + (r/R))}_{\approx 0}\Delta \mathbf{C}_t + \underbrace{\Pi\xi_{t+1}}_{\equiv \epsilon_{t+1}} \quad (45)$$

for a white noise variable ϵ_{t+1} , and $(r/R) \approx 0$ for plausible quarterly interest rates. (45) leads directly to (3).

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

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

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

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

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

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

$$\mathbb{M}[p_{t+1}^2] = \mathcal{D}\mathbb{M}[p_t^2] \mathbb{E}[\psi^2] + \mathbf{D},$$

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

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

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

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

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 $\mathcal{D}\mathbb{E}[\psi^2] < 1$. 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.

C.3 Converting Annual to Quarterly Variances for Idiosyncratic Shocks

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

$$\theta_t^a = \sum_{i=1}^4 \frac{\theta_{t+i}}{4}$$

for $t = 0, 4, 8, \dots$. 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:

$$\psi_t^a = \prod_{i=1}^4 \psi_{t+i}$$

for $t = 0, 4, 8, \dots$. Then the variance of the annual permanent shock is $(1 + \text{var}(\psi))^4 \approx 4 \times \text{var}(\psi)$ for small $\text{var}(\psi)$. Therefore we divide our calibrated annual permanent shock (0.012) by 4 to get a quarterly number.

C.4 Muth (1960) Signal Extraction

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

$$\tilde{P}_t = v_1 Y_t + v_2 Y_{t-1} + v_3 Y_{t-2} + \dots \quad (46)$$

for a sequence of v 's given by

$$v_k = (1 - \lambda_1) \lambda_1^{k-1} \quad (47)$$

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

$$\tilde{P}_t = (1 - \lambda_1)(Y_t + \lambda_1 Y_{t-1} + \lambda_1^2 Y_{t-2} \dots) \quad (48)$$

$$\tilde{P}_{t+1} = (1 - \lambda_1)(Y_{t+1} + \lambda_1 Y_t + \lambda_1^2 Y_{t-1} + \lambda_1^3 Y_{t-2} \dots) \quad (49)$$

$$= (1 - \lambda_1) Y_{t+1} + \lambda_1 \underbrace{(1 - \lambda_1)(Y_t + \lambda_1^2 Y_{t-1} + \lambda_1^3 Y_{t-2} \dots)}_{\tilde{P}_t} \quad (50)$$

$$= (1 - \lambda_1) Y_{t+1} + \lambda_1 \tilde{P}_t \quad (51)$$

This compares with (32) in the main text

$$\tilde{P}_{t+1} = \Pi Y_{t+1} + (1 - \Pi) \tilde{P}_t \quad (52)$$

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

$$\lambda_1 = 1 - \Pi \quad (53)$$

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

(1960) we have

$$\begin{aligned}\lambda_1 &= 1 + (1/2)\varphi^2 - \varphi\sqrt{1 + \varphi^2/4} \\ (1 - \Pi) &= 1 + (1/2)\varphi^2 - \varphi\sqrt{1 + \varphi^2/4} \\ -\Pi &= (1/2)\varphi^2 - \varphi\sqrt{1 + \varphi^2/4}\end{aligned}\tag{54}$$

yielding equation (33) in the main text.

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			OLS	2 nd Stage	KP <i>p</i> -val
Independent Variables			or IV	\bar{R}^2	Hansen J <i>p</i> -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 , $\Delta \log \mathbf{C}_t^* = \mathbf{Z}_t \zeta$, $\bar{R}^2 = 0.018$; $\text{var}(\log(\xi_t)) = 3.33\text{e-}6$					
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			
0.412 ^{•••}			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 , $\Delta \log \mathbf{C}_t^* = \mathbf{Z}_t \zeta$, $\bar{R}^2 = 0.186$; $\text{var}(\log(\xi_t)) = 3.33\text{e-}6$					

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 $\mathbf{Z}_t = \{\Delta \log \mathbf{C}_{t-2}, \Delta \log \mathbf{C}_{t-3}, \Delta \log \mathbf{Y}_{t-2}, \Delta \log \mathbf{Y}_{t-3}, A_{t-2}, A_{t-3}, \Delta_8 \log \mathbf{C}_{t-2}, \Delta_8 \log \mathbf{Y}_{t-2}\}$.

Acknowledgements

The computational results in this paper were constructed using tools in the Econ-ARK/HARK toolkit, and the code to produce the paper's results is available there. The toolkit 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, Bartosz Maćkowiak, Kathrin Schlafmann, Gianluca Violante 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.

Christopher D. Carroll

Johns Hopkins University, Baltimore, United States; email: ccarroll@jhu.edu

Edmund Crawley

Johns Hopkins University, Baltimore, United States; email: ecrawle2@jhu.edu

Jiri Slacalek

European Central Bank, Frankfurt am Main, Germany; email: jiri.slacalek@ecb.europa.eu

Kiichi Tokuoka

Ministry of Finance, Tokyo, Japan; email: kiichi.tokuoka@mof.go.jp

Matthew N. White

University of Delaware, Newark, United States; email: mnwecon@udel.edu

© European Central Bank, 2018

Postal address 60640 Frankfurt am Main, Germany
Telephone +49 69 1344 0
Website www.ecb.europa.eu

All rights reserved. Any reproduction, publication and reprint in the form of a different publication, whether printed or produced electronically, in whole or in part, is permitted only with the explicit written authorisation of the ECB or the authors.

This paper can be downloaded without charge from www.ecb.europa.eu, from the [Social Science Research Network electronic library](#) or from [RePEc: Research Papers in Economics](#). Information on all of the papers published in the ECB Working Paper Series can be found on the [ECB's website](#).

ISSN	1725-2806 (pdf)	DOI	10.2866/509 (pdf)
ISBN	978-92-899-3257-8 (pdf)	EU catalogue No	QB-AR-18-032-EN-N (pdf)