Sticky Expectations and Consumption Dynamics — Online Appendix

March 1, 2019

Christopher D. Carroll[1] JHU

Edmund Crawley[2] JHU

Jiri Slacalek[3] ECB

Kiichi Tokuoka[4] MoF Japan

Matthew N. White[5] UDel

Abstract

To match empirical evidence about aggregate consumption dynamics, representative agent macroeconomic models require some mechanism to generate 'excess smoothness' in consumption growth. But a large literature has found no evidence of corresponding smoothness in microeconomic data; consequently, existing microfounded models constructed to match the microdata fail to match the macro smoothness facts. We show that the micro and macro evidence can be reconciled by a microfounded model in which consumers have accurate knowledge of their personal circumstances but 'sticky expectations' about the macroeconomy, and 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, Expectations, Habits, Inattention

JEL codes D83, D84, E21, E32

`Repo:` https://github.com/llorracc/cAndCwithStickyE
`PDF:` http://econ.jhu.edu/people/ccarroll/papers/cAndCwithStickyE.pdf
`Web:` http://econ.jhu.edu/people/ccarroll/papers/cAndCwithStickyE/
`Slides:` Versions to View or Print
`Toolkit:` http://github.com/Econ-ARK/HARK

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

A Calibration

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

A.1 Macroeconomic Calibration

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

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

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

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

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

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

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

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

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

data;[2] the value of β satisfying $\mathcal{R}\beta = 1$ is used in the closed economy models presented in the online appendix, allowing those models to fit the *mean* observed wealth.

A.2 Calibration of Idiosyncratic Shocks

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

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

Our calibration for the sizes of the idiosyncratic shocks are conservative relative to the literature;[3] using data from the *Panel Study of Income Dynamics*, for example, Carroll and Samwick (1997) estimate $\sigma_\psi^2 = 0.0217$ and $\sigma_\theta^2 = 0.0440$; Storesletten, Telmer, and Yaron (2004) estimate $\sigma_\psi^2 \approx 0.017$, with varying estimates of the transitory component. But recent work by Low, Meghir, and Pistaferri (2010) suggests that controlling for participation decisions reduces estimates of the permanent variance somewhat; and using very well-measured Danish administrative data, Nielsen and Vissing-Jorgensen (2006) estimate $\sigma_\psi^2 \approx 0.005$ and $\sigma_\theta^2 \approx 0.015$, which presumably constitute lower bounds for plausible values for the truth in the U.S. (given the comparative generosity of the Danish welfare state).

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

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

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

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

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

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

For the HA-DSGE model, we set the discount factor to $\beta = \mathcal{R}^{-1} = 0.986$, roughly matching the target capital-to-output ratio.[4] Households in the HA-DSGE model thus hold significantly more wealth than their counterparts in the baseline SOE model, who were calibrated to approximate the *median* observed wealth-to-income ratio. This reflects our goal of presenting results that span the full range of calibrations in the micro and macro literatures; the micro literature has often focused on trying to explain the wealth holdings of the median household, which are much smaller than average wealth holdings. Experimentation has indicated that our results are not sensitive to such choices.

B.1 Model and Solution

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

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

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

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

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

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

The growth-rate-conditional parameters $\kappa_{j,0}$ and $\kappa_{j,1}$ are exogenous to the individual's (partial equilibrium) optimization problem, but are endogenous to the general equilibrium of the economy. Taking the aggregate saving rule $\aleph$ as given, the household's problem can be written in Bellman form as:[5]

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

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

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

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

The equilibrium of the HA-DSGE model is characterized by a (normalized) consumption function $\mathrm{c}(m, M, \Phi)$ and an aggregate saving rule $\aleph$ such that when all households believe $\aleph$, the solution to their individual problem (24) is c; and when all agents act according to c, the best log-linear fit of A_t on M_t (conditional on Φ_t) is $\aleph$. The model is solved using a method similar to Krusell and Smith (1998).[6]

B.2 Frictionless vs Sticky Expectations

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

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

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

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

Households who misperceive the aggregate productivity state will incorrectly predict

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

[6] Details are in Appendix D.1.2.

aggregate saving at the end of the period, and thus aggregate capital and the distribution of factor prices next period.[7]

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

B.3 Results

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

Table 6 reports the results of estimating regression (17) on data generated from the HA-DSGE model. The results are substantially the same as the previous analysis for the SOE model (in Table 3).[8]

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

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

C Representative Agent (RA) Model

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

C.1 Model and Solution

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

The representative agent's problem can be written in Bellman form as:[9]

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

Normalizing the representative agent's problem by the productivity level P_t as in the SOE and HA-DSGE models, the problem's state space can be reduced to:[10]

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

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

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

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

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

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

The representative agent model can be solved using the endogenous grid method, following the same procedure as for the SOE model described in Appendix D.1, yielding normalized consumption function $\mathrm{C}(M, \Phi)$.[11]

C.2 Frictionless vs Sticky Expectations

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

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

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

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

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

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

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

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

That is, the average perception of productivity in each growth state is the weighted average of updaters and non-updaters who perceive that growth rate.[12]

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

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

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

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

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

When the representative agent frictionlessly updates its information every period ($\Pi = 1$), equations (27) and (28) say that $\varphi_t = e_N^j$ and $\widetilde{P}_t^j = P_t$ (with irrelevant values in the other vector elements), so that the representative agent is truly representative. When expectations are sticky ($\Pi < 1$), the representative agent's perceptions of the growth rate become "smeared" across its past realizations; its perceptions the productivity level likewise deviate from the true value, even for the part of the representative agent who perceives the true growth rate.[13]

C.3 Simulation Results

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

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

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

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

D Numerical Methods

D.1 Solution Methods

D.1.1 Small Open Economy Solution Details

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

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

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

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

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

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

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

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

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

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

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

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

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

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

D.1.2 Dynamic Stochastic General Equilibrium Solution Details

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

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

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

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

1. Construct a grid of (normalized) aggregate market resources $\mathbb{M} = \{M_j\}_{j=1}^{N_M}$. We use a $N_M = 19$ point grid based on the steady state of the perfect foresight DSGE model, spanning the range of 10 percent to 500 percent of this value.

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

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

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

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

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

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

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

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

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

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

D.2 Simulation Procedures

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

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

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

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

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

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

4. The economy produces output. All agents draw idiosyncratic shocks $\psi_{t,i}$ and $\theta_{t,i}$, with newborns automatically drawing $\psi_{t,i} = \theta_{t,i} = 1$,[14] then observe their true $\mathbf{m}_{t,i}$ (and $\mathbf{M}_t$ in the HA-DSGE model).

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

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

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

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

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

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

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

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

D.3 Cost of Stickiness Calculation

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

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

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

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

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

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

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

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

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

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

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

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

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

E Muth–Lucas–Pischke

To see how the Muth–Lucas–Pischke model can generate smoothness, note that in the Muth framework, agents update their estimate of permanent income according to an equation of the form:[15]

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

We can now consider the dynamics of aggregate consumption in response to the arrival of an aggregate shock that (unbeknownst to the consumer) is permanent. The consumer spends Π of the shock in the first period, leaving $(1-\Pi)$ unspent because that reflects the average transitory component of an undifferentiated shock. However, since the shock really was permanent, income next period does not fall back as the consumer guessed it would on the basis of the mistaken belief that $(1-\Pi)$ of the shock was transitory. The next-period consumer treats this surprise as a positive shock relative to expected income, and spends the same proportion Π out of the perceived new shock. These dynamics continue indefinitely, but with each successive perceived shock (and therefore each consumption increment) being smaller than the last by the proportion $(1-\Pi)$. Thus, after a true permanent shock received in period t, the full-information prediction of the expected dynamics of future consumption changes would be $\Delta \mathbf{C}_{t+n+1} = (1-\Pi)\Delta \mathbf{C}_{t+n} + \epsilon_{t+n}$.[16]

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

The in-principle reconciliation of econometric evidence of predictability/excess smoothness in consumption growth, and the random walk proposition, is therefore that the econometricians who are making their forecasts of aggregate consumption growth use additional variables (beyond the lagged history of aggregate income itself), and that those variables have useful predictive power.[17]

---

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

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

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

F Alternate Belief Specification

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

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

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

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

$$\Theta_t^{\not\Pi} \sim \mathcal{N}(-(\sigma_\Theta^2 + \sigma_\Psi^2/\Pi)/2, \sigma_\Theta^2 + \sigma_\Psi^2/\Pi), \qquad \Psi_t^{\not\Pi} = 1, \qquad \Xi^{\not\Pi} \sim I.$$

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

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

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

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

The key result is that households' optimal consumption function barely changes from baseline when the alternate beliefs are introduced: across states actually attained during simulation, normalized consumption differs by no more than 0.2 percent, and the difference is less than 0.02 percent in the vast majority of states. More importantly, the macroeconomic dynamics generated by sticky expectations households' collective behavior is nearly identical between the bottom panels of Table 8 below and Table 3 in the main text.[18] This experiment represents a more general proposition that our main results should be robust to the details of the precise specification of households' understanding of their inattention, so long as the key feature remains that agents' idiosyncratic errors are *systematically correlated* due to the lag in information.

G Additional Calculations

G.1 Quadratic Utility Consumption Dynamics

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

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

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

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

The text asserts (equation (3)) that

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

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

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

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

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

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

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

So equation (3) can be rewritten as

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

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

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

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

G.2 Population Variance of Idiosyncratic Permanent Income

This appendix follows closely Appendix A in the ECB working paper version of Carroll, Slacalek, and Tokuoka (2015).[19] It computes dynamics and steady state of the square of the idiosyncratic component of permanent income (from which the variance can be

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

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

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

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

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

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

$$\mathbb{M}\left[p_{t+1}^2\right] = (1-\mathrm{d})\mathbb{M}[p_t^2]\,\mathbb{E}[\psi^2] + \mathrm{d},$$

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

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

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 $(1-\mathrm{d})\,\mathbb{E}[\psi^2] < 1$. This requires that income does not spread out so quickly among survivors as to overcome the compression of the distribution that arises because of death.

G.3 Converting Annual to Quarterly Variances for Idiosyncratic Shocks

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

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

for $t = 0, 4, 8, \ldots$ Then the variance of the annual transitory shock is $\frac{1}{4}$ of the variance of the quarterly transitory shock: $\mathrm{var}(\theta^a) = \frac{4}{16}\,\mathrm{var}(\theta) = \frac{1}{4}\,\mathrm{var}\,\theta$. 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 $t = 0, 4, 8, \ldots$ Then the variance of the annual permanent shock is $\big(1+\mathrm{var}(\psi)\big)^4 \approx$

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

G.4 Muth (1960) Signal Extraction

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

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

for a sequence of v's given by

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

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

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

This compares with (32) in the main text

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

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

$$\lambda_1 = 1 - \Pi$$

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

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

yielding equation (20) in the main text.

References

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.

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.

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.

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.

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

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

Low, Hamish, Costas Meghir, and Luigi Pistaferri (2010): "Wage risk and employment risk over the life cycle," *The American economic review*, 100(4), 1432–1467.

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*.

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.

Table 5 Equilibrium Statistics

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

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

Table 6 Aggregate Consumption Dynamics in HA-DSGE Model

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

Expectations : Dep Var Independent Variables			OLS or IV	2^{nd} Stage $\bar{R}^2$	Hansen J p-val
Frictionless : $\Delta \log \mathbf{C}^*_{t+1}$ (with measurement error $\mathbf{C}^*_t = \mathbf{C}_t \times \xi_t$);					
$\Delta \log \mathbf{C}^*_t$	$\Delta \log \mathbf{Y}_{t+1}$	A_t			
0.189			OLS	0.036	
(0.072)					
0.476			IV	0.020	0.556
(0.354)					
	0.368		IV	0.017	0.457
	(0.321)				
		−0.34e–4	IV	0.015	0.433
		(0.98e–4)			
0.289	0.214	0.01e–4	IV	0.020	0.531
(0.463)	(0.583)	(1.87e–4)			
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			OLS	0.223	
(0.061)					
0.773			IV	0.230	0.542
(0.108)					
	0.912		IV	0.145	0.187
	(0.245)				
		−0.97e–4	IV	0.059	0.002
		(0.56e–4)			
0.670	0.171	0.12e–4	IV	0.231	0.551
(0.181)	(0.363)	(0.86e–4)			
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. Instruments $\mathbf{Z}_t = \{\Delta \log \mathbf{C}_{t-2}, \Delta \log \mathbf{C}_{t-3}, \Delta \log \mathbf{Y}_{t-2}, \Delta \log \mathbf{Y}_{t-3}, A_{t-2}, A_{t-3}, \Delta_8 \log \mathbf{C}_{t-2}, \Delta_8 \log \mathbf{Y}_{t-2}\}$.

Table 7 Aggregate Consumption Dynamics in RA Model

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

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

Notes: Reported statistics are the average values for 100 samples of 200 simulated quarters each. Instruments $\mathbf{Z}_t = \{\Delta \log \mathbf{C}_{t-2}, \Delta \log \mathbf{C}_{t-3}, \Delta \log \mathbf{Y}_{t-2}, \Delta \log \mathbf{Y}_{t-3}, A_{t-2}, A_{t-3}, \Delta_8 \log \mathbf{C}_{t-2}, \Delta_8 \log \mathbf{Y}_{t-2}\}$.

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

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

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

Notes: Reported statistics are the average values for 100 samples of 200 simulated quarters each. Instruments $\mathbf{Z}_t = \{\Delta \log \mathbf{C}_{t-2}, \Delta \log \mathbf{C}_{t-3}, \Delta \log \mathbf{Y}_{t-2}, \Delta \log \mathbf{Y}_{t-3}, A_{t-2}, A_{t-3}, \Delta_8 \log \mathbf{C}_{t-2}, \Delta_8 \log \mathbf{Y}_{t-2}\}$.