

Slow Information Diffusion and the Inertial Behavior of Durable Consumption^{*}

Yulei Luo[†]

Jun Nie[‡]

The University of Hong Kong

Federal Reserve Bank of Kansas City

Eric R. Young[§]

University of Virginia

Forthcoming in *Journal of the European Economic Association*

Abstract

This paper studies the aggregate dynamics of durable and nondurable consumption under slow information diffusion (SID) due to noisy observations and learning within the permanent income framework. We show that SID can significantly improve the model's predictions on the joint behavior of income, durable consumption, and nondurable consumption at the aggregate level. Specifically, we find that SID can significantly improve the model's predictions for: (i) smoothness in durable and nondurable consumption, (ii) autocorrelation of durable consumption, and (iii) contemporaneous correlation between durable and nondurable consumption.

JEL Classification Numbers: *C61, D81, E21.*

Keywords: Durability, Slow Learning, Slow Information Diffusion, Infrequent Adjustments, Consumption Stickiness.

*We are grateful for Marios Angeletos (Editor) and three anonymous referees for many constructive suggestions and comments. We also thank Edward Knotek II, Geng Li, Jordan Rappaport, Tom Sargent, Chris Sims, Yi-Chang Tsai, Jonathan Willis, and seminar and conference participants at the Federal Reserve Bank of Kansas City, Federal Reserve Board, The University of Tokyo, The Hong Kong University of Science and Technology, Washington and Lee University, the Midwest Macroeconomics Meetings, and the Shanghai Macroeconomic Workshop for helpful discussions and comments, and Wei Li and Lisa Taylor for valuable research assistance. Luo thanks the General Research Fund (GRF#: HKU749711 and HKU748209) in Hong Kong for financial support. Young thanks the Bankard Fund for Political Economy at Virginia for financial support. The views expressed here are the opinions of the authors only and do not necessarily represent those of the Federal Reserve Bank of Kansas City or the Federal Reserve System. All remaining errors are our responsibility.

[†]School of Economics and Finance, Faculty of Business and Economics, The University of Hong Kong, Hong Kong. Email address: yluo@econ.hku.hk.

[‡]Research Department, Federal Reserve Bank of Kansas City. E-mail: jun.nie@kc.frb.org.

[§]Department of Economics, University of Virginia, Charlottesville, VA 22904. Email address: ey2d@virginia.edu.

1. Introduction

Representing more than two-thirds of real GDP, personal consumption expenditures is by far the largest component of the US economy, highlighting the importance of understanding consumption dynamics. Within the general consumption category, durable consumption is worth particular attention because expenditures on durable goods are highly volatile and the dynamics of durable spending differ significantly from those of nondurable spending.¹ The standard approach to studying the dynamics of consumption begins with the permanent income hypothesis (PIH) and adapts the basic model (Hall 1978) in response to various deviations of theory from data, including the celebrated “excess sensitivity” and “excess smoothness” puzzles.

Though the implications of the PIH framework on total consumption expenditures are well understood, little attention has been directed toward whether the theory also generates data-consistent predictions for expenditures on durable consumption goods. In this paper, we first document the discrepancies between the predictions of the PIH theory and the data, focusing on different dimensions of durable goods dynamics. We then show how incorporating slow information diffusion (SID) due to noisy observations in a standard PIH model can significantly improve the model’s performance at explaining the joint dynamics of nondurable consumption, durables expenditures, and income in the U.S. economy.

The key implication of Hall’s PIH model is that nondurable consumption is a random walk, i.e., the changes in nondurable consumption are unpredictable. Translating this basic insight to the context of durable goods, Mankiw (1982) notes that the stock of durable goods must follow a random walk and, by implication, the change in the shock of durables follows an MA(1) process. Mankiw then proceeds to show that this prediction is grossly rejected in the data: using quarterly US data to estimate the above equation, he finds that in Hall’s (1978) PIH model in which utility is a quadratic function of the stock of durable goods, the stock of durable goods is a random walk and the change in durable goods, Δe , follows an MA(1) process, with the MA coefficient equal to the negative of one minus the depreciation rate:

$$\Delta e_t = \zeta_t - (1 - \delta) \zeta_{t-1}, \quad (1)$$

¹Broadly speaking, durable consumption consists of consumer spending in four categories: motor vehicles and parts, recreational goods and vehicles, furnishings and durable household equipment, and other durables (which includes jewelry, luggage, books, and telephone equipment). In total, durable consumption accounts for about 10 percent of total personal consumption. In general, quantitative work has assigned durable expenditures to investment rather than consumption, as the dynamics are more similar to investment.

where ζ_t is a white-noise innovation and $\delta \in [0, 1]$ is the depreciation rate.² Using quarterly US data to estimate this equation, Mankiw finds that the change in the stock of durables has positive serial correlation in post-war US quarterly data and the depreciation rate would need to be roughly 100 percent to make the model fit the data (that is, durables are not in fact durable). This finding is called the “Mankiw puzzle” in the literature. In addition, Caballero (1994) shows in a PIH model with both nondurable and durable goods that the rejection of the martingale property of durable goods is an order of magnitude larger than that for nondurables and the finding is robust across categories of durable goods. The Mankiw puzzle is not an isolated phenomenon; Caballero (1990, 1994) and Adda and Cooper (2006) find that the puzzle is robust across different time periods, different frequencies, and different countries.

Bernanke (1985) studies the joint behavior of nondurable and durable consumption in the presence of adjustment costs of changing durables stocks within a simple representative agent PIH framework. He finds that the costs of adjusting durables stocks are substantial and help improve the model’s prediction for the joint behavior of aggregate consumption and income.³ The main prediction of Bernanke’s model is that with adjustment costs households always adjust their stock gradually to the desired level, as determined by their permanent income; in other words, in the presence of income shocks, households engage in purchases and resales on a continuous basis in the sense that they will purchase successively better durable goods over several consecutive periods. However, this prediction is inconsistent with an important feature of the micro-level data on durables (e.g., automobile expenditures) that households adjust their durables stocks infrequently.⁴ Bar-Ilan and Blinder (1991) show that consumers facing lumpy transaction costs either fully adjust by replacing their old durable good or do not adjust at all; in other words, people purchase durable goods infrequently and, when they do, the additions to their stocks are significant. In addition, Bertola and Caballero (1990) show that intermittent large adjustments can be explained by the observation that microeconomic adjustment cost functions are often kinked at the no-adjustment point.

In this paper, we take an alternative approach to the Mankiw puzzle, one based on informational frictions at the micro-level. As argued in many studies, informational frictions can be very important: households, firms, individual investors, and even the government may have heterogeneous beliefs and observations about the current state of the economy. This could be due to many

²Hall (1978) shows that under the PIH, the change in nondurable consumption is unpredictable.

³We will discuss Bernanke’s adjustment cost model in Section 6.

⁴Lam (1991) reports that households only occasionally adjust their stock of durables.

reasons. For example, it could arise from segmented market interactions (Lucas 1972; Angelos and La'O 2013), from difficulty in distinguishing different components in the income process (Muth 1960; Wang 2004), from infrequent information updating (Mankiw and Reis 2002; Reis 2006); or from rational inattention due to finite information-processing capacity (Sims 2003). Specifically, in this paper we study a permanent income model with durable goods and examine implications of slow information diffusion (SID) for the joint dynamics of nondurable and durable consumption at both the micro- and macro-levels. SID is induced by the assumption that noisy signals about the true state(s) have to be learned slowly due to signal extraction. One microfoundation of noisy observations and slow learning is rational inattention (RI), a consequence of finite information-processing constraints. RI was first proposed by Sims (2003) as a tool to capture the observed sluggishness, randomness, and delays in the responses of economic variables to shocks.⁵ Under RI, agents only have finite information-processing capacity and thus cannot observe the state of the economy without errors; consequently, they react to exogenous shocks gradually and with delay.⁶ In Section 4, we will show that in our setting RI and signal extraction due to measurement error (or any other exogenously-generated noise) are observationally equivalent in the sense that they lead to the same model dynamics.

Intuitively, the SID model we propose can resolve the Mankiw puzzle because it breaks the link between the MA coefficient on durable expenditures and the depreciation rate. With sluggish adjustment, there are internal dynamics to durable expenditures that are not present under full-information rational expectations (FI-RE). As households gradually learn about a change in the state, their stock of durables will slowly adjust.⁷ Indeed, Caballero (1990a) explicitly suggests that slow diffusion of information could account for the particular adjustment process he posits. Thus, our SID model provides a simple microfoundation for the slow adjustment mechanism used in that paper.

After solving our model explicitly, we analytically prove that SID improves the model's performance in the following key aspects of the joint behavior of income, nondurable consumption,

⁵Luo (2008), Luo and Young (2010), and Tutino (2013) use the RI framework to examine the dynamics of nondurable consumption.

⁶Reis (2006) uses "inattentiveness" to characterize the inertial behavior of consumers. In this paper, to avoid the confusion between "rational inattention" and "inattentiveness," we use the terminology "sticky expectations" instead of "inattentiveness."

⁷The effect of habit formation in consumption on the joint dynamics of nondurable and durable consumption depends on how we model habit formation in the utility function. In some studies (e.g., Deaton 1992; Dynan 2000), they consider habit formation and durability separately and model both as the "consumption stocks" having opposite impacts on utility. For example, Dynan (2000) argues that durability makes expenditure growth lumpy whereas habit formation smoothes it out. In addition, her results yield no evidence of habit formation at the annual frequency.

and durable consumption: (i) it reduces the relative volatility of aggregate nondurable consumption to aggregate income, which helps resolve the excess smoothness puzzle in the literature on nondurable consumption; (ii) it reduces the relative volatility of aggregate durable consumption to nondurable consumption, (iii) it increases the first-order serial correlation of expenditures on aggregate durable consumption, which helps resolve the Mankiw puzzle; and (iv) it reduces the contemporaneous correlation between nondurable and durable consumption.⁸ The mechanisms through which SID can improve these dimensions are as follows. First, as households cannot fully observe the true state under SID, they adjust their consumption more gradually in response to income shocks. This helps reduce the volatility of both nondurable and durable consumption. Second, as durable consumption measures the changes in the stock of durables, it tends to respond even more gradually than nondurable consumption. The main reason for this is due to the interaction of the depreciation channel and the SID channel. Given that the MA coefficient in (1), $1 - \delta$, is greater than 0, the change in durable consumption is actually more volatile than that in the stock of durables. In other words, the depreciation channel ($\delta < 1$) has the potential to increase the relative volatility of the change in durable consumption to the change in nondurable consumption. In contrast, the SID channel offsets the depreciation channel and thus reduces the relative volatility of durable consumption to nondurable consumption.⁹ Third, as durable consumption responds more gradually to income shocks, the persistence tends to increase, which is a typical dynamic of consumption under imperfect state observation.¹⁰ Finally, as durable consumption adjusts more gradually than nondurable consumption, the correlation between them tends to decrease. Using the explicit solutions, we show that although consumers can devote much more capacity to processing economic information and then improve their optimal consumption decisions, it is rational for them not to do so because the welfare improvement is tiny.¹¹

It is clear that the benchmark SID model cannot capture the observed inertial behavior at the individual level, i.e., infrequent and lumpy purchases on durables in the micro-level data. The reason that SID cannot capture this key feature of the behavior of individual consumers is that consumers who extract useful information from learning noisy signals adjust their durable stock

⁸As far as we know, the contemporaneous correlation between nondurable and durable consumption has not been studied in the PIH framework. In addition, several recent papers have pointed out that the standard New Keynesian model cannot produce the positive co-movements between durable consumption and nondurable consumption. (See Monacelli (2009) for a discussion.)

⁹Note that when δ is 1, the processes of nondurable and durable consumption are essentially the same and thus SID has no impact on the relative volatility of the change in durable consumption to that in nondurable consumption.

¹⁰See Luo (2008) for a discussion.

¹¹This result is consistent with that obtained in the models without durable consumption, e.g., Pischke (1995) and Luo and Young (2010).

gradually in response to income shocks. We then show that introducing fixed adjustment costs into the benchmark SID model can capture both infrequent adjustments at the individual level and gradual adjustments at the aggregate level.

The rest of the paper is organized as follows. Section 2 presents key facts on durable and non-durable consumption. Section 3 proposes a stylized permanent income model with durable goods and discusses the model's predictions on the dynamics of nondurable and durable consumption. Section 4 solves the permanent income model with durable goods and SID due to imperfect state observation and examines the welfare implications of SID. Section 5 studies the empirical implications of SID for the stochastic properties of the joint dynamics of aggregate nondurable and durable consumption. Section 6 includes some extensions and a discussion on how fixed adjustment costs lead to infrequent adjustments and can thus potentially better explain both micro- and macro-level data. Section 7 concludes.

2. Facts

This section documents key aspects of durable and nondurable consumption. Because this paper studies whether information frictions can help explain the dynamics of durables and nondurables, we follow closely the literature in constructing the data and the key moments.

We follow Galí's (1993) definition of durable and nondurable consumption, where nondurables are defined as personal consumption expenditures less durable goods.¹² The data covers the period of 1955 – 2007.¹³ The data is taken from the database of Forecasting, Analysis, and Modeling Environment (FAME) and the Archival Federal Reserve Economic Data (ALFRED). As in Galí (1993), we use seasonally-adjusted quarterly real variables and focus on the quarterly change of durables and nondurables.¹⁴ Income is constructed as real GDP minus investment (i.e., Gross Fixed Capital Formation) and government expenditures (i.e., General Government Final Consumption Expenditure).¹⁵ All data are real, with the base year being 2005. The data is detrended using the Hodrick-Prescott (HP) filter (with a smoothing parameter of 1600). The reported stan-

¹²This means nondurables consumption includes both nondurable goods and services.

¹³We follow Mankiw (1982) to exclude the Korean war period as he argued that the permanent income hypothesis (PIH) may not hold in that period. Similarly, we also exclude the period surrounding the 2007 – 2009 Great Recession. For curiosity of readers, we also report statistics using the full sample, i.e., 1955 – 2012. (See Table 2.)

¹⁴Notice that Galí (1993) uses per-capita variables, while we focus on aggregate variables. Using per-capita variables has little effect on the studied statistics (as many of them are ratios).

¹⁵To be consistent with the welfare analysis in Section 4.2 which is based on individual consumption dynamics, we used per-capita income in the estimation of the income process in the next section. Using aggregate income does not alter the results qualitatively.

dard errors in the parentheses are the GMM-corrected standard errors of the statistics.

We briefly list the facts we focus on in Table 1. As all variables are measured in changes, we'll simply omit "changes" in the remainder of this section.¹⁶ First, nondurable consumption is less volatile than income. The ratio of the standard deviation of nondurable expenditures to the standard deviation of income is 0.66.¹⁷ Second, durable expenditures are less volatile than nondurable consumption. The ratio of the standard deviation of durable expenditures to the standard deviation of nondurable consumption is 0.62. Third, the autocorrelation of durable expenditures is -0.3 for the 1955 – 2007 period. (It is -0.03 for the 1955 – 2012 period, which is not statistically different from zero.) Fourth, the correlation between durable expenditures and nondurable consumption is positive but not very large: 0.46.

3. A Stylized Permanent Income Model with Durable Goods

In this section we present a standard full-information rational expectations (FI-RE) version of the permanent income model with durable goods, and discuss the main empirical shortcomings of the model. We will then examine how incorporating slow information diffusion due to noisy signals and slow learning affects the joint behavior of nondurables and durable consumption in the next section. All model economies will be populated by a continuum of infinitely-lived consumers and prices will be assumed exogenous and constant.¹⁸

3.1. The Model

Following Mankiw (1982), Bernanke (1985), and Galí (1993), we consider an FI-RE version of the PIH model which integrates both durable and nondurable consumption, where the latter includes both nondurable goods and services.¹⁹ The optimizing decisions of a representative consumer in

¹⁶The literature on PIH is usually focused on changes in variables rather than growth rates. See Hall (1978), Mankiw (1982), and Galí (1993).

¹⁷As will be clear in later sections, both the standard PIH model and the PIH model with imperfect state observation imply $E[\Delta C] = E[\Delta E] = 0$. We therefore detrend both durable and nondurable consumption data to make the data and the model comparable. Notice that in the PIH model durable and nondurable consumption are not stationary while changes in durable and nondurable consumption are stationary. We also find that using the raw data to compute the key second moments only slightly changes the main results in this paper.

¹⁸The benchmark model presented in this paper is usually interpreted as a partial equilibrium PIH model. However, as noted in Hansen (1987), they can also be interpreted as a general equilibrium model with a linear production technology and an exogenous income process. Specifically, given the expression of optimal consumption derived from the benchmark model, we can price assets by treating optimal consumption as though it were an endowment process. In this setup, equilibrium prices are shadow prices that leave the agent content with that endowment process.

¹⁹Although the original Mankiw (1982) model only considers durable consumption, including nondurables consumption in preferences does not change his main conclusion provided they enter in a separable manner.

the RE-PIH model with durable goods can be formulated as

$$\max_{\{c_t, k_{t+1}\}} \left\{ E_0 \left[\sum_{t=0}^{\infty} \beta^t u(c_t, k_t) \right] \right\}, \quad (2)$$

subject to the budget constraint

$$a_{t+1} = Ra_t + y_t - c_t - e_t, \quad (3)$$

and the accumulation equation for durables

$$k_t = (1 - \delta) k_{t-1} + e_t, \quad (4)$$

where $u(c_t, k_t) = -(\bar{c} - c_t)^2 / 2 - \rho (\bar{k} - k_t)^2 / 2$ is the utility function, \bar{c} and \bar{k} are the bliss points, c_t is consumption of nondurables, k_t is the stock of durable goods, y_t is labor income, e_t is the purchase of durable goods, δ is the depreciation rate of durable goods, β is the discount factor, R is the constant gross interest rate, and $\beta R = 1$ (an assumption typically imposed in the literature to guarantee a stochastic steady state).²⁰ Combining (3) and (4) gives the period-to-period finance constraint of the consumer:

$$a_{t+1} = Ra_t + (1 - \delta) k_{t-1} - k_t + y_t - c_t. \quad (5)$$

We define

$$s_t = a_t + \frac{1 - \delta}{R} k_{t-1} + \frac{1}{R} \sum_{j=0}^{\infty} R^{-j} E_t [y_{t+j}]; \quad (6)$$

s_t is the expected present value of lifetime resources, consisting of financial wealth (the risk free foreign bond), existing stock of durable goods, plus human wealth. Solving this optimization problem gives optimal decisions for nondurable and durable consumption:

$$c_t = H_c s_t, \quad (7)$$

$$k_t = \frac{R + \delta - 1}{\rho} c_t = H_k s_t, \quad (8)$$

where the marginal propensity to consume out of permanent income, H_c , is

$$H_c = (R - 1) \left(1 + \frac{(1 - \beta(1 - \delta))^2}{\beta \rho} \right)^{-1} \quad (9)$$

²⁰For simplicity, we assume that the price of durable goods in terms of nondurable consumption is 1.

and $H_k = (R + \delta - 1) H_c / \varrho$ (see Appendix 8.1 for derivations).

As shown in Luo (2008), to facilitate the introduction of signal extraction (or rational inattention), we reduce the above multivariate PIH model to a univariate one in which the unique state variable is permanent income s_t that can be solved in closed-form under noisy signals and slow learning.²¹ Specifically, if s_t is defined as a new state variable, the original finance constraint can be rewritten as

$$s_{t+1} = R s_t - c_t - (1 - \beta(1 - \delta)) k_t + \zeta_{t+1}, \quad (10)$$

where the time $(t + 1)$ innovation to permanent income, ζ_{t+1} , is

$$\zeta_{t+1} = \frac{1}{R} \sum_{j=0}^{\infty} \left(\frac{1}{R} \right)^j (E_{t+1} - E_t) [y_{t+1+j}]. \quad (11)$$

We complete the model description by specifying the income process. Following Quah (1990) and Deaton (1993), we assume that aggregate labor income includes a unit root and the whole income process has two kinds of structural shocks to labor income: One has a permanent impact on the level of labor income and the other has only transitory impact. Specifically, the income process can be written as:

$$y_{t+1} = y_{t+1}^p + y_{t+1}^i, \quad (12)$$

$$y_{t+1}^p = y_t^p + \varepsilon_{t+1}, \quad (13)$$

$$y_{t+1}^i = \bar{y} + \epsilon_{t+1}, \quad (14)$$

where y_{t+1}^p and y_{t+1}^i are permanent and transitory income components, respectively, ε_{t+1} and ϵ_{t+1} are orthogonal permanent and transitory iid shocks with mean 0 and variance ω^2 and ω_ϵ^2 , respectively. As shown in Quah (1990), this two-component income specification provides a potential resolution to Deaton's puzzle (i.e., the excess smoothness puzzle) in the standard permanent income model if the relative importance of transitory to permanent components is large. Here we estimate the income process using the U.S. data from the period of 1955 – 2007, and find that $\omega^2 = 125.7^2$ and $\omega_\epsilon^2 = 2.4^2$.²² This is consistent with Quah's (1990) finding that the volatility

²¹Reduction of the state space to univariate is particularly convenient for the rational inattention (RI) problem, as it is well-known that multi-dimensional RI problems are significantly less tractable. In particular, while the optimal distribution chosen by the RI agent is still Gaussian, it cannot in general be computed analytically; the problem is a form of the classic water-filling problem but the weighting scheme differs in the utility function and the information constraint, rendering the problem intractable.

²²If we use the data set from the 1955 – 2012 period, and find that $\omega^2 = 131.7^2$ and $\omega_\epsilon^2 = 0.8^2$. The estimation is

of consumption is mainly due to the variations of the permanent component in the income process.²³ In the permanent income model with durables we presented above, it is straightforward to show that the income specification can affect the relative volatility of nondurable consumption to income, but has no impact on examining how SID affects the stochastic properties of the joint dynamics of nondurable and durable consumption if consumers can distinguish between the two components in income.²⁴

In this case, permanent income s_t , can be written as

$$s_t = a_t + \frac{1 - \delta}{R} k_{t-1} + \frac{1}{R - 1} y_t; \quad (15)$$

that is, s_t is a linear combination of three state variables, financial wealth, the stock of durable goods, and labor income. Given the specification of the income process, (12)-(14), the innovation to permanent income can be written as $\zeta_{t+1} = \varepsilon_{t+1} / (R - 1) + \epsilon_{t+1} \sim N(0, \omega_\zeta^2)$, where $\omega_\zeta^2 = \omega^2 / (R - 1)^2 + \omega_\epsilon^2 \cong \omega^2 / (R - 1)^2$.²⁵ Combining (7), (8), and (10) gives the expressions for the changes in nondurable goods and the stock of durable goods:

$$\Delta c_t = H_c \zeta_t, \quad (16)$$

$$\Delta k_t = \frac{R + \delta - 1}{\rho} H_c \zeta_t. \quad (17)$$

Equation (16) is just the random walk result of Hall (1978), and the expenditure on durable goods follows the ARMA(1, 1) process

$$e_t = e_{t-1} + \zeta_t - (1 - \delta) \zeta_{t-1}, \quad (18)$$

where $\zeta_t = [(R + \delta - 1) / \rho] [1 + (1 - \beta(1 - \delta))^2 / (\beta\rho)]^{-1} \varepsilon_t$ is an unforecastable innovation to consumption at time t . The MA coefficient is determined entirely by the depreciation rate, δ . In estimating the above equation using US quarterly data, Mankiw (1982) finds that empirically δ is quite close to 1. In other words, durables do not look very durable at all and the stochastic behavior of durables purchases seems to be too similar to that of nondurables consumption to be consistent with the standard PIH's predictions. Specifically, (18) implies that the first-order autocorrelation

implemented using the Matlab toolbox: SSMMATLAB.

²³As Tables 1-9 in Quah (1990) show, across different specifications, the variation of the transitory component in the income process only accounts about 1% – 2% of total variation of consumption.

²⁴In Section 4.3, we consider an extension in which consumers cannot distinguish the two components in income.

²⁵Given that $\omega^2 = 125.7^2$, $\omega_\epsilon^2 = 2.4^2$, and $R = 1.01$, $\omega^2 / (R - 1)^2 \gg \omega_\epsilon^2$.

of Δe_t is

$$\rho_{\Delta e_t} \equiv \text{corr}(\Delta e_t, \Delta e_{t-1}) = \frac{\delta - 1}{1 + (1 - \delta)^2} < 0$$

because the depreciation rate is less than 1 in the data. For example, if $\delta = 0.05$ (a value that roughly produces the observed ratio of durables to producer capital in a standard growth model), $\rho_{\Delta e_t} = -0.499$. However, the estimated value of $\rho_{\Delta e_t}$ is far from this number: using the same data set that Mankiw used, the correlation is 0.06, which implies that the depreciation rate should be 1.07 to make the model fit the data, and more recent data generates similar results (a correlation of -0.04 implies $\delta = 0.99527$). Tables 1 and 2 report our new estimates of $\rho_{\Delta e_t}$ using the U.S. data from 1955 – 2007 and 1955 – 2012, respectively.²⁶ In the two samples, $\rho_{\Delta e_t}$ is equal to -0.3 (s.d. 0.07) and -0.03 (s.d. 0.12), respectively, which require $\delta = 0.67$ and $\delta = 0.97$. It is clear that the Mankiw puzzle still exists. Obviously, a model with this property is going to be difficult to calibrate to observed aggregate data on investment and stocks of durables.

4. Permanent Income Models with Durable Goods and Slow Information Diffusion

In this section, we incorporate slow information diffusion (SID) due to imperfect state observation and slow learning into the otherwise standard permanent income model with durable goods and explore how slow information diffusion due to imperfect observations and slow learning affect the dynamic impacts of income shocks on the joint behavior of nondurables and durables consumption.

4.1. Imperfect State Observation and Slow Learning

We assume that consumers in the model economy cannot perfectly observe the true state, permanent income (s_t), and can only observe a noisy signal

$$s_t^* = s_t + \zeta_t, \tag{19}$$

when making decisions, where s_t follows (10) and ζ_t is the iid Gaussian noise due to imperfect observations. The specification in (19) is standard in the signal extraction (SE) literature and captures the situation where consumers happen or choose to have imperfect knowledge of the idiosyncratic or underlying common shocks.²⁷ It is worth noting that this assumption is also consistent with the rational inattention (RI) hypothesis proposed by Sims (2003) that ordinary people only devote

²⁶We study the sample 1955 – 2007 because we want to exclude the Great Recession period after the financial crisis.

²⁷For example, Muth (1960), Lucas (1972), Lorenzoni (2009), and Angeletos and La'O (2010, 2012).

finite information-processing capacity to processing financial information and thus cannot observe the states perfectly. One could think of the noise in the signal as being due to rational inattention. However, it is unclear whether a signal of the form: s_t plus iid noise ξ_t is optimal from a rational inattention perspective. It may be the case that the household prefers a signal with different dynamic properties. In this paper, we focus on the iid noise case and leave this to future research.²⁸ It is worth noting that in the traditional SE problem, we do not have such restriction on the stochastic properties of the noises, and the fundamental variable could be correlated with the exogenous noise.²⁹

Since imperfect observations on the state lead to welfare losses, households use the processed information to estimate the true state.³⁰ Specifically, we assume that households use the Kalman filter to update the perceived state $\hat{s}_t = E_t[s_t]$ after observing new signals in the steady state in which the conditional variance of s_t , $\Sigma_t = \text{var}_t(s_t)$, has converged to a constant Σ :

$$\hat{s}_{t+1} = (1 - \theta) [R\hat{s}_t - c_t - (1 - \beta(1 - \delta))k_t] + \theta s_{t+1}^*, \quad (20)$$

where θ is the Kalman gain (i.e., the optimal weight on any new observation).³¹

In the signal extraction problem, the Kalman gain can be written as

$$\theta = \Sigma \Lambda^{-1}, \quad (21)$$

where Σ is the steady state value of the conditional variance of s_{t+1} , $\text{var}_{t+1}(s_{t+1})$, and $\Lambda = \text{var}_t(\xi_{t+1})$ is the variance of the noise. Σ and Λ are linked by the following updating equation for the conditional variance in the steady state:

$$\Lambda^{-1} = \Sigma^{-1} - \Psi^{-1}, \quad (22)$$

where Ψ is the steady state value of the ex ante conditional variance of s_{t+1} , $\Psi_t = \text{var}_t(s_{t+1})$. Multiplying ω_ξ^2 (the variance of the innovation to s) on both sides of (22) and using the fact that

²⁸For an analytical solution to an RI tracking problem of this form, please see Propositions 3-4 in Mackowiak and Wiederholt (2009).

²⁹See Luo and Young (2014) for a discussion on this issue.

³⁰See Section 4.2 for details about the welfare implication under SID.

³¹Note that θ measures how much uncertainty about the state can be removed upon receiving the new signals about the state.

$\Psi = R^2\Sigma + \omega_\zeta^2$, we have

$$\omega_\zeta^2\Lambda^{-1} = \omega_\zeta^2\Sigma^{-1} - \left[R^2 \left(\omega_\zeta^2\Sigma^{-1} \right)^{-1} + 1 \right]^{-1}, \quad (23)$$

where $\omega_\zeta^2\Sigma^{-1} = \left(\omega_\zeta^2\Lambda^{-1} \right) \left(\Lambda\Sigma^{-1} \right)$.

Define the signal-to-noise ratio (SNR) as $\pi = \omega_\zeta^2\Lambda^{-1}$. We obtain the following equality linking π and the Kalman gain (θ):

$$\pi = \theta \left(\frac{1}{1-\theta} - R^2 \right). \quad (24)$$

Solving for θ yields

$$\theta = \frac{-(1+\pi) + \sqrt{(1+\pi)^2 + 4R^2(\pi + R^2)}}{2R^2}, \quad (25)$$

where we omit the negative values of θ because both Σ and Λ must be positive. It is straightforward to show that θ and π have one-to-one monotonic relationship. Note that given π , we can pin down Λ using $\pi = \omega_\zeta^2\Lambda^{-1}$ and Σ using (21) and (25).

Notice that this signal extraction problem with exogenously specified noises is observationally equivalent to the RI model with endogenous noises and fixed (or elastic) capacity. Specifically, consumers under RI face both the usual flow budget constraint as well as an information-processing constraint due to finite Shannon capacity. Following Sims (2003), the typical consumer's information-processing constraint can be characterized by the inequality

$$\mathcal{H}(s_{t+1}|\mathcal{I}_t) - \mathcal{H}(s_{t+1}|\mathcal{I}_{t+1}) \leq \kappa, \quad (26)$$

where \mathcal{I}_t is the consumer's currently processed information, κ is the consumer's channel capacity, $\mathcal{H}(s_{t+1}|\mathcal{I}_t)$ denotes the entropy of the state prior to observing the new signal at $t+1$, and $\mathcal{H}(s_{t+1}|\mathcal{I}_{t+1})$ is the entropy after observing the new signal. (26) implies that the reduction in the uncertainty about the state variable gained from observing a new signal is bounded by κ . As shown in Sims (2003), within the linear-quadratic-Gaussian setting, \mathcal{D}_t is a normal distribution $N(\hat{s}_t, \Sigma_t)$; as a result, (26) can be reduced to

$$\ln |\Psi_t| - \ln |\Sigma_{t+1}| = 2\kappa, \quad (27)$$

where $\Sigma_{t+1} = \text{var}(s_{t+1}|\mathcal{I}_{t+1})$ and $\Psi_t = \text{var}(s_{t+1}|\mathcal{I}_t)$ are the posterior and prior variances of the

state variable, respectively. In this univariate case, (27) has the steady state $\Sigma = \omega_{\xi}^2 / (\exp(2k) - R^2)$, and the consumer behaves as if observing a noisy measurement of permanent income $s_{t+1}^* = s_{t+1} + \xi_{t+1}$, where ξ_{t+1} is the endogenous noise with mean 0 and variance $\Lambda_t = \text{var}(\xi_{t+1} | \mathcal{I}_t)$; in the steady state $\Lambda = (\Sigma^{-1} - \Psi^{-1})^{-1}$ and the Kalman gain:

$$\theta = \Sigma \Lambda^{-1} = 1 - \frac{1}{\exp(2\kappa)}, \quad (28)$$

is the optimal weight on any new observation. Comparing (25) with (28), it is clear that if the SNR and capacity satisfy:

$$\pi = \left(1 - \frac{1}{\exp(2\kappa)}\right) (\exp(2\kappa) - R^2), \quad (29)$$

the SE and RI problems are observationally equivalent in the sense that they lead to the same model dynamics governed by θ . (29) clearly shows that the SNR is an increasing function of channel capacity. In addition, as argued in Sims (2010), instead of assuming that channel capacity is fixed, it is also reasonable to assume that the marginal cost of information processing is constant such that capacity can be elastic in response to a change in environment. In other words, the Lagrange multiplier on (27), λ , is constant. As these two modeling strategies are also observationally equivalent in the sense that they lead to the same model dynamics, here we just use the Kalman gain θ to characterize the degree of SID.³²

4.2. Individual Dynamics and Welfare Implications under SID

Combining (19) with (20), we obtain the following proposition about the dynamic behavior of the perceived state \hat{s}_t :

Proposition 1. *Under SID, \hat{s}_t follows:*

$$\hat{s}_{t+1} = R\hat{s}_t - c_t - (1 - \beta(1 - \delta))k_t + \eta_{t+1}, \quad (30)$$

where

$$\eta_{t+1} = \theta \left[\left(\frac{\xi_{t+1}}{1 - (1 - \theta)R \cdot L} \right) + \left(\xi_{t+1} - \frac{\theta R \xi_t}{1 - (1 - \theta)R \cdot L} \right) \right], \quad (31)$$

$$\omega_{\eta}^2 = \text{var}(\eta_{t+1}) = \frac{\theta}{1 - (1 - \theta)R^2} \omega_{\xi}^2 > \omega_{\xi}^2, \text{ for } \theta < 1, \quad (32)$$

³²See Section III of Luo and Young (2014) for the derivation of the observational equivalence between these two assumptions.

and we use the facts that $\omega_{\zeta}^2 = \text{var}(\zeta_{t+1}) = \frac{1}{\theta} \frac{1}{1/(1-\theta) - R^2} \omega_{\zeta}^2$ and the estimation error, $s_t - \hat{s}_t$, can be written as

$$s_t - \hat{s}_t = \frac{(1-\theta)\zeta_t}{1 - (1-\theta)R \cdot L} - \frac{\theta\tilde{\zeta}_t}{1 - (1-\theta)R \cdot L}. \quad (33)$$

Expression (33) shows that the estimation error reacts to the fundamental shock positively, while it reacts to the noise shock negatively. In addition, the importance of the estimation error is decreasing with θ . More specifically, as θ increases, the first term in (33) becomes less important because $(1-\theta)\zeta_t$ in the numerator decreases, and the second term also becomes less important because the variance of $\tilde{\zeta}_t$ decreases as θ increases.

The optimization problem for the typical household facing state uncertainty can thus be reformulated as

$$v(\hat{s}_t) = \max_{\{c_t, k_t\}} \{E_t[u(c_t, k_t) + \beta v(\hat{s}_{t+1})]\}, \quad (34)$$

subject to (30)-(32), and given \hat{s}_0 . Solving this Bellman equation yields the following proposition:

Proposition 2. *Under SID, the consumption and durable accumulation functions are*

$$c_t = H_c \hat{s}_t, \quad (35)$$

$$k_t = H_k \hat{s}_t, \quad (36)$$

where H_c is defined in (9) and $H_k = \frac{1-\beta(1-\delta)}{\beta\varrho} H_c$, and the value function is

$$v(\hat{s}_t) = A_0 + A_1 \hat{s}_t + A_2 \hat{s}_t^2, \quad (37)$$

where $A_2 = -R(R-1)/(2\Theta)$, $A_1 = R\bar{c}$, and $A_0 = -R\Theta\bar{c}^2/(2(R-1)) - R\text{var}(\eta_{t+1})/(2\Theta)$.

Proof. See Appendix 8.2. ■

When $\theta = 1$, i.e., in the FI-RE case, the consumption functions reduce to (35) and (36), respectively, and the value function reduces to

$$\tilde{v}(s_t) = A_0 + A_1 s_t + A_2 s_t^2,$$

where $A_2 = -R(R-1)/(2\Omega)$, $A_1 = R\bar{c}$, and $A_0 = -R\Theta\bar{c}^2/(2(R-1)) - R\text{var}(\zeta_{t+1})/(2\Omega)$. Since imperfect-state-observation cannot help in individuals' optimization – consumers with finite

capacity cannot observe the state perfectly when making optimal decisions – the average welfare difference between the RI and FI-RE economies is greater than 0. We examine here the welfare cost of RI – how much utility does a consumer lose if the actual consumption path he chooses under SID deviates from the first-best FI-RE path? Specifically, following Barro (2007) and Luo and Young (2010), given the initial value of the state, \hat{s}_0 , the marginal welfare costs (mwc) due to SID in our benchmark model can be written as

$$mwc \equiv -\frac{\partial v(\hat{s}_0) / \partial \theta}{(\partial v(\hat{s}_0) / \partial \hat{s}_0) \hat{s}_0} = -\frac{1}{2[(R-1)\hat{s}_0 - \Theta \bar{c}] \hat{s}_0} \frac{\partial \omega_\eta^2}{\partial \theta}.$$

This expression gives the proportionate reduction in the initial level of the perceived state (\hat{s}_0) that compensates, at the margin, for a decrease in θ (i.e., stronger SID) — in the sense of preserving the same effect on welfare for a given \hat{s}_0 . To do quantitative welfare analysis we need to know the levels of \hat{s}_0 and \bar{c} . First, denote by γ the local coefficient of relative risk aversion, which equals $\gamma = E[y] / (\bar{c} - E[y])$ for the utility function $u(\cdot)$ evaluated at mean income $E[y]$. Using the U.S. data from the 1955 – 2007 period, we have $E[y_t] = 16798$, $\omega = 125.7$, and $\omega_\epsilon = 2.4$ (all in 2005 U.S. dollars), and then find the value of the bliss point \bar{c} that generates reasonable relative risk aversion γ . For example, if γ is equal to 1.5, $\bar{c} = 1.5E[y_t]$. Furthermore, assume that the ratio of the initial level of financial wealth (\hat{a}_0) to mean income ($\hat{y}_0 \equiv E[y_t]$) is 5, that is, $\hat{a}_0 / \hat{y}_0 = 5$.³³ Given that $\hat{s}_0 = \hat{a}_0 + (1 - \delta) \hat{k}_0 / R + \frac{1}{R-1} \hat{y}_0$, we can calculate that $mwc = 9.89 \times 10^{-4}$ when $\theta = 0.5$, using the fact that $\partial \omega_\eta^2 / \partial \theta = (1 - R^2) \omega_\xi^2 / [1 - (1 - \theta) R^2]^2 < 0$. Therefore, to maintain the level of the value function, an increase in θ of 100% (from 0.5 to 1) requires a reduction in the initial level of \hat{s}_0 by approximately 0.049%.³⁴ This result thus provides some evidence that it is reasonable for consumers to learn the true state slowly due to finite capacity because the welfare improvement from increasing learning capacity is trivial. In other words, although consumers can devote much more capacity to processing economic information and then improve their optimal consumption decisions, it is rational for them not to do so because the welfare improvement is tiny. This result is consistent with that obtained in the models without durable consumption, e.g., Pischke (1995) and Luo and Young (2010).

³³This number varies largely for different individuals, from 2 to 20. 5 is the average wealth/income ratio in the Survey of Consumer Finances 2001. We find that changing the value of this ratio only has minor effects on the welfare implication.

³⁴This result is robust to the change in the value of θ . For example, when $\theta = 0.6$, $mwc = 6.775 \times 10^{-4}$, and an increase in θ by 66.7% (from 0.6 to 1) requires a reduction in the initial level of \hat{s}_0 by approximately 0.027% in order to maintain the level of the value function.

Using (35) and (36), straightforward calculations imply that

$$\Delta c_t = \theta H_c \left[\frac{\zeta_t}{1 - (1 - \theta)R \cdot L} + \left(\xi_t - \frac{\theta R \xi_{t-1}}{1 - (1 - \theta)R \cdot L} \right) \right], \quad (38)$$

$$\Delta k_t = \theta H_k \left[\frac{\zeta_t}{1 - (1 - \theta)R \cdot L} + \left(\xi_t - \frac{\theta R \xi_{t-1}}{1 - (1 - \theta)R \cdot L} \right) \right], \quad (39)$$

where we use the fact that $\Delta \widehat{s}_t = \theta \left[\frac{\zeta_t}{1 - (1 - \theta)R \cdot L} + \left(\xi_t - \frac{\theta R \xi_{t-1}}{1 - (1 - \theta)R \cdot L} \right) \right]$. Hence, under the innocuous assumption that $(1 - \theta)R < 1$, both consumption processes follow MA(∞) processes.³⁵ Expenditure on durable goods follows the process

$$\Delta e_t = \theta H_k \left[\begin{array}{c} \left(\zeta_t + \frac{((1 - \theta)R - (1 - \delta))\zeta_{t-1}}{1 - (1 - \theta)R \cdot L} \right) + \\ \left(\xi_t - (1 - \delta + \theta R) \xi_{t-1} + \frac{\theta R((1 - \delta) - (1 - \theta)R)\xi_{t-2}}{1 - (1 - \theta)R \cdot L} \right) \end{array} \right], \quad (40)$$

which reduces to $\Delta e_t = H_k \zeta_t$ when $\theta = 1$.

4.3. Aggregation

Our model economy is now populated by a continuum of ex ante identical but ex post heterogeneous consumers because consumers face the idiosyncratic noise shock. Sun (2006) presents a law of large numbers for this type of economic models and then characterizes the cancellation of individual risk via aggregation. In this paper, following Uhlig (1996) and Zaffaroni (2004), we show that the idiosyncratic RI-induced noises can be exactly cancelled out after aggregating across all agents if they converge in mean square to the population mean (0). Specifically, after aggregating over all consumers under an assumption of identical θ , we obtain the expressions for changes in aggregate nondurables and durables:

$$\Delta C_t = \theta H_c \left[\frac{\zeta_t}{1 - (1 - \theta)R \cdot L} \right], \quad (41)$$

$$\Delta K_t = \frac{1 - \beta(1 - \delta)}{\beta \varrho} \Delta C_t, \quad (42)$$

and

$$\Delta E_t = \theta H_k \left[\zeta_t + \frac{((1 - \theta)R - (1 - \delta))\zeta_{t-1}}{1 - (1 - \theta)R \cdot L} \right], \quad (43)$$

respectively (see Appendix 8.3 for the proof). It is worth noting that the assumption of convergence in mean square helps resolve the impossibility result discussed in Judd (1985) and Feldman and

³⁵This assumption only has bite when θ is very close to 0 where the absence of precautionary motives due to quadratic utility is likely to be problematic.

Gilles (1985).³⁶

Equations (41)-(43) clearly show that SID can help generate the smooth and hump-shaped impulse responses of nondurables and durables consumption to the income shock. More specifically, we explore how SID affects the stochastic properties of the joint dynamics of income, nondurables, and durables along the following key dimensions: (i) the relative volatility of nondurables to income, (ii) the relative volatility of expenditures on durables and nondurables, (iii) the first-order autocorrelation of changes in durables expenditures, and (iv) the contemporaneous correlation between nondurables and durable expenditures. After inspecting the third aspect above, we easily determine whether the Mankiw puzzle can be resolved by breaking the tight link between the MA coefficient and the depreciation rate implied by the FI-RE assumption.

5. Empirical Implications

5.1. Stochastic Properties of Nondurable and Durable Consumption

5.1.1. The Relative Volatility of ΔC_t to ΔY_t

Given (41) and (43), the relative volatility of the changes in nondurable consumption to income can be written as:

$$\mu \equiv \frac{\text{sd}(\Delta C_t)}{\text{sd}(\Delta Y_t)} = \left(1 + \frac{(1 - \beta(1 - \delta))^2}{\beta\varrho}\right)^{-1} \sqrt{\frac{\theta^2}{1 - ((1 - \theta)R)^2}}, \quad (44)$$

where $\Theta \equiv \sqrt{\frac{\theta^2}{1 - ((1 - \theta)R)^2}}$.

Proposition 3.

$$\frac{\partial \mu}{\partial \theta} > 0. \quad (45)$$

Proof. It's straightforward to show $\frac{\partial \Theta}{\partial \theta} > 0$. Thus, $\frac{\partial \mu}{\partial \theta} > 0$. ■

The above proof shows that slow learning reduces the relative volatility μ via an additional factor due to SID, Θ . Figure 1 illustrates how θ affects Θ . It clearly shows that slow learning due to noisy state observations increases the excess smoothness of nondurables relative to income.

³⁶The impossibility result says that if agents in a continuum population face idiosyncratic uncertainty, the strong law of large numbers that assures the exact cancellation of the idiosyncratic uncertainty does not apply because the sets of agents obtaining a certain realization may not be measurable or they do not have the appropriate measure even if they are measurable.

As shown in Table 3, when $\theta = 0.62$, i.e., 62% of any new information is transmitted each period (equivalently 62% of the uncertainty is removed upon the receipt of a new signal), $\mu = 0.66$, exactly what it is in the data. It is not difficult to understand why SID reduces the relative volatility of nondurable consumption. As (41) shows, the nondurable consumption changes, ΔC_t , becomes an $MA(\infty)$ process, meaning that it not only depends on the current innovation but also is influenced by innovations in previous periods. This makes nondurable consumption change more gradually, and therefore has a lower volatility. In addition, as well documented in the consumption literature (e.g., Deaton (1992) and Reis (2006)), the impulse response of aggregate nondurable consumption to aggregate income takes a hump-shaped form, which means that aggregate consumption reacts to income shocks gradually and with delay.

5.1.2. The Relative Volatility of ΔE_t to ΔC_t

Given (41) and (43), the relative volatility of the changes in durable to nondurable consumption can be defined as follows

$$rv \equiv \frac{sd(\Delta E_t)}{sd(\Delta C_t)} = \left(\frac{R + \delta - 1}{\varrho} \right) \sqrt{1 + (1 - \delta)^2 - 2(1 - \delta)(1 - \theta)R}, \quad (46)$$

where $sd(\Delta C_t)$ and $sd(\Delta E_t)$ are standard deviation of ΔC_t and ΔE_t , respectively.

Proposition 4.

$$\frac{\partial (rv)}{\partial \theta} > 0.$$

Proof. This can easily be proved by simply illustrating that the second term on the right-hand side of (46) monotonically increases with the degree of slow learning, θ . ■

This proposition is very interesting and probably requires more explanations. First of all, it says that changes in durable consumption become less volatile than changes in nondurable consumption under SID.³⁷ Given that Proposition 3 shows that SID can reduce the volatility of nondurable consumption changes, it is not surprising to see that it can also reduce the volatility of durable consumption changes. The question is why SID reduces the volatility of durable consumption changes more than that of nondurable consumption changes. The key reason that the change in durable consumption is actually more volatile than that in the stock of durables when $\delta < 1$ is due to the MA representation of ΔE_t . (Note that in this case, $1 + (1 - \delta)^2 > 1$, and SID measured by

³⁷As Table 4 shows, the standard FI-RE model predicts a relative volatility of durable to nondurable larger than 1. Similar evidence has been documented in Galí (1993) as well.

$\theta < 1$ smooths the process for the stock of durables and nondurables in a similar fashion as shown by (41) and (43). In other words, the depreciation channel ($\delta < 1$) in this case has the potential to increase the relative volatility of ΔE_t to ΔC_t . When the depreciation rate is 100 percent, the C_t and K_t processes are essentially the same and thus SID has no impact on the relative volatility of the change in durable consumption to that in nondurable consumption. In contrast, in the presence of SID, i.e., $\theta < 1$, (41) and (43) clearly show that the SID channel offsets the depreciation channel and thus reduces the relative volatility of ΔE_t to ΔC_t .³⁸

Second, it is worth noting that $E[\Delta C_t] = E[\Delta K_t] = 0$ in the model with SID.³⁹ This means that the variations of ΔC_t and ΔK_{t+1} are not influenced by their levels (as both are zero, on average). Therefore, it excludes the possibility that SID reduces $\text{sd}(\Delta C_t) / \text{sd}(\Delta E_t)$ by altering the relative size of ΔC to ΔE .

Another way to examine how SID affects the relative variability is to define

$$\Pi \equiv \frac{\text{rv}(\theta = 1)}{\text{rv}(\theta < 1)} = \sqrt{\frac{1 + (1 - \delta)^2}{1 + (1 - \delta)^2 - 2(1 - \delta)(1 - \theta)R}}. \quad (47)$$

Figure 2 clearly shows that the presence of SID governed by θ can improve the model's prediction for the observed variability ratio for different values of R . An example is when $\delta = 0.05$ and $\theta = 0.1$, $\Pi = 3.7$, that is, if 10% of the uncertainty is removed upon the receipt of a new signal, the predicted relative variability can be reduced by about 4 times.

5.1.3. The First-order Autocorrelation of ΔE_t

By construction, (43) can be rewritten as

$$\Delta E_t = \zeta_t + \frac{((1 - \theta)R - (1 - \delta))\zeta_{t-1}}{1 - (1 - \theta)R \cdot L}, \quad (48)$$

where $\zeta_t = [(R + \delta - 1) / \varrho] \theta H_c \zeta_t$ with $\text{var}(\zeta_t) = [(R + \delta - 1) / \varrho]^2 (\theta H_c)^2 \omega_\zeta^2$. Given (48), the first-order autocorrelation of ΔE_t can be written as follows:

$$\rho \equiv \frac{\text{cov}(\Delta E_{t+1}, \Delta E_t)}{\text{var}(\Delta E_t)} = (1 - \theta)R - (1 - \delta) \frac{1 - ((1 - \theta)R)^2}{1 + (1 - \delta)^2 - 2(1 - \delta)(1 - \theta)R}. \quad (49)$$

³⁸Note that in the expression for ΔE_t , the $1 - (1 - \theta)R \cdot L$ term makes the process smoother and $(1 - \theta)R - (1 - \delta)$ also reduces the initial impact of the depreciation channel.

³⁹This is also true in the standard FI-RE model.

Based on (49), the following proposition shows how the combination of (θ, δ) affects the first-order autocorrelation of the expenditure on durables.⁴⁰

Proposition 5.

$$\frac{\partial \rho}{\partial \theta} < 0, \frac{\partial \rho}{\partial \delta} > 0.$$

Proof. Given (49), it is straightforward to show that $\partial \rho / \partial \theta < 0$ because $\theta, (1 - \theta)R \in (0, 1)$, and

$$\frac{\partial \rho}{\partial \delta} = \frac{(1 - (1 - \delta)^2) [1 - ((1 - \theta)R)^2]}{[1 + (1 - \delta)^2 - 2(1 - \delta)(1 - \theta)R]^2} > 0.$$

■

This proposition shows that SID increases the first-order autocorrelation of ΔE_t , i.e., the less the value of θ , the larger ρ . Figure 3 illustrates how ρ increases with the degree of SID. In addition, the higher the depreciation rate, the larger ρ . For example, given $R = 1.01$ and $\delta = 0.05$, $\rho = 0.5$ when $\theta = 1$, $\rho = -0.25$ when $\theta = 0.5$, and $\rho = -0.03$ when $\theta = 0.1$. These results suggest that SID has the potential to resolve the Mankiw puzzle.

5.1.4. The Correlation between ΔE_t and ΔC_t

Given (41) and (43), the contemporaneous correlation between the changes in durable and non-durable consumption can be written as:

$$\text{corr}(\Delta E_t, \Delta C_t) \equiv \frac{\text{cov}(\Delta E_t, \Delta C_t)}{\text{sd}(\Delta C_t) \text{sd}(\Delta E_t)} = \frac{1 - (1 - \delta)(1 - \theta)R}{\sqrt{1 + (1 - \delta)^2 - 2(1 - \delta)(1 - \theta)R}}, \quad (50)$$

We then have the next result.

Proposition 6.

$$\frac{\partial \text{corr}(\Delta E_t, \Delta C_t)}{\partial \theta} > 0. \quad (51)$$

Proof. The proof is straightforward. ■

Here is some intuition on why SID reduces the contemporaneous correlation between the changes in nondurable and durable consumption. From Proposition 4 we see that SID reduces

⁴⁰Note that when $\theta = 1$, $\rho = \frac{-(1-\delta)}{1+(1-\delta)^2}$.

relative volatility of durable consumption changes to nondurable consumption changes, meaning that SID makes durable consumption respond more gradually to income shocks than nondurable consumption. Figure 4 illustrates how the contemporaneous correlation between durable and nondurable consumption decreases with the degree of SID (i.e., less θ). Intuitively, as durable consumption and nondurable consumption respond to income shocks in increasingly different ways, the correlation between them also declines.

5.2. Quantitative Results

The previous section provides qualitative results based on the closed-form solutions which show that introducing SID can help a standard PIH model with both durable and nondurable goods better explain the four dimensions of durable and nondurable consumption. In particular, Propositions 3-6 have shown that all these improvements are driven by the change of one single parameter, θ , which is the optimal weight on any new observation (or the Kalman gain). That is, based on a standard framework used in the literature, our analysis highlights the effects of information frictions on the model implications for the joint dynamics of durable and nondurable consumption.

This section quantifies the improvement in model predictions through assigning values for this key parameter, θ . Generally speaking, there are multiple ways we can choose a value for θ . For instance, we can set different values for θ to match each of the four dimensions we studied in the previous section. However, as the focus of this analysis is on how SID helps explain the behavior of durable goods, we will not use the moments involving durable consumption to calibrate θ . So, in the calibration, θ is chosen to match the observed relative volatility of nondurable consumption to income in the data.⁴¹

Before going to the results, Table 3 reports the values for other parameters used to generate the quantitative results. In choosing values for these parameters, we closely follow the literature, which allows us to focus on the effects of our key SID parameter (θ) on changing the model predictions. The preference parameter ρ is chosen from Bernanke (1985). The (quarterly) depreciation rate for durable goods is set to be 1.5% which lies well in the range used in the literature. For example, Bernanke (1985) uses 2.5% and Monacelli (2009) uses 1% (i.e., annually 4%).

Table 4 reports the quantitative results using data from the period 1955 – 2007 and the results for the period 1955 – 2012 are reported in Table 5. These tables clearly show that SID can

⁴¹Estimating θ without using a model is difficult; estimates in the literature exist for the amount of information that humans can process (Landauer 1986), but it is difficult to map these numbers into the amount of attention that is actually allocated to monitoring the economic situation of the household.

significantly improve the model's quantitative predictions on the joint dynamics of nondurable and durable consumption. Particularly, the main conclusion from these results is that a value of θ which matches the relative volatility of nondurable consumption to income significantly improves the model predictions along the other three dimensions of durable consumption as well. Specifically, using the 1955 – 2007 data, when $\theta = 62\%$, the SID model can improve the model's prediction on the relative volatility of durable to nondurable consumption, the first-order autocorrelation of durable consumption, and the contemporaneous correlation between durable and nondurable consumption by 44%, 95%, and 80%, respectively.⁴² As a reminder, we are not directly choosing θ to match any moments on durable consumption (although this can easily be done). Thus, these results suggest that the SID mechanism is important for explaining the behavior of durable consumption.

6. Extensions and Discussion

6.1. Bernanke's Adjustment Costs (AC) Model

The main difference between the model present in Section 3.1 and the model in Bernanke (1985) is that the latter assumes changing durables stocks involves quadratic adjustment costs because purchases of durables require leisure expenditure. Specifically, the utility function of a representative consumer during a given period t is assumed to be

$$u(c_t, k_t, k_{t-1}) = -\frac{1}{2}(\bar{c} - c_t)^2 - \frac{\varrho}{2}(\bar{k} - k_t)^2 - \frac{\vartheta}{2}(k_t - k_{t-1})^2, \quad (52)$$

where ϑ measures the importance of adjustment costs in utility.⁴³ Solving this model yields the following dynamics for nondurable and durable consumption:

$$\Delta c_t = G_c G_y \varepsilon_t, \quad (53)$$

$$\Delta k_t = \frac{x_1 (1 - \beta (1 - \delta))}{\vartheta (1 - x_2^{-1})} \frac{G_c G_y \varepsilon_t}{1 - x_1 \cdot L}, \quad (54)$$

where L is the lag operator. (See Online Appendix for the derivation.)

⁴²Using the 1955 – 2012 data and $\theta = 60\%$, we find that SID can improve these three dimensions by 62%, 43%, and 80%, respectively.

⁴³Bernanke (1985) assumes that utility is a non-separable function of nondurables and durables consumption; that is, there is an additional term $-m(\bar{c} - c_t)(\bar{k} - k_t)$ in the utility function. However, the estimated m , the parameter measuring the degree of non-separability, is not significantly different from 0. Hence, for simplicity here we assume that $m = 0$ and focus on the effect of adjustment costs.

Clearly, (54) is an MA(∞) process with decreasing MA coefficients, which means that durables consumption reacts to the income shock gradually in the presence of adjustment costs. Figure 1 illustrates the impulse responses of durables consumption growth Δk_t to the income shock when the parameters are the same as that estimated in Bernanke (1985) ($R = 1.01$, $\vartheta = 0.706$, $\delta = 0.025$, $\varrho = 0.0286$, $x_1 = 0.828$). Expression (54) also shows that the presence of adjustment costs can improve the model's predictions in the following aspects: (1) it increases excess smoothness of durables consumption and (2) it increases the autocorrelation of durables consumption by introducing a slow adjustment mechanism.

However, although the presence of adjustment costs reduces the initial response of nondurables consumption to the income shock because $G_c G_y < 1$ when $\vartheta > 0$ as is clear in Equation (53), it does not affect the dynamic responses of nondurables consumption, which is still the random walk result of Hall (1978).⁴⁴ Specifically, the introduction of adjustment costs reduces the relative volatility of aggregate nondurables consumption to income, defined as

$$\mu = \frac{\text{sd}(\Delta C_t)}{\text{sd}(\Delta Y_t)} = \left(1 + (1 - \beta(1 - \delta))^2 \frac{R}{R - x_1} \frac{x_1}{\vartheta(1 - x_2^{-1})} \right)^{-1} < 1.$$

Using the same parameter values as in the last section, we find that $\mu = 0.98$, which is the same as that obtained in the FI-RE model and is well above its empirical counterpart (0.66). In other words, costs of adjusting durable stocks do not improve the model's predictions for the joint behavior of aggregate nondurables consumption and income sufficiently; in US data nondurables consumption is much smoother than income. In addition, it is clear from (53) that the impulse response of aggregate nondurable consumption to aggregate income is flat with an immediate upward jump in the initial period that persists indefinitely, which is not consistent with the VAR evidence documented in the literature that the impulse response of aggregate nondurable consumption to income takes a hump-shaped form.

To compare the AC model with the SID model, we set $x_1 = (1 - \theta)R$ such that the two models have the same propagation mechanism in the dynamics of durable consumption. We report the results in Tables 4-5. It is clear from the tables that SID did a better job in explaining the relative volatility of nondurable consumption to income and the contemporaneous correlation between nondurable and durable consumption: The AC model's predictions on these two moments are

⁴⁴In other words, nondurable consumption is not sensitive to past information, as predicted by the standard permanent income model.

the same as that obtained in the FI-RE model. It is worth noting that the AC model's prediction on the relative volatility of durable to nondurable consumption matches the data better than the SID model at the cost of worsening the model's prediction on the relative volatility of nondurable consumption to income.⁴⁵

6.2. Incomplete Information about Current Income (IC)

In this subsection, we consider an extended incomplete information (IC) model in which consumers cannot distinguish the two components in income specified in (12)-(14). Specifically, following Muth (1960) and Pischke (1991), given that the change in income is

$$\Delta y_{t+1} = \varepsilon_{t+1} + \epsilon_{t+1} - \epsilon_t, \quad (55)$$

the best forecast is to recognize that Δy_{t+1} is a moving-average process of order one:

$$\Delta y_{t+1} = v_{t+1} - \alpha v_t, \quad (56)$$

where the innovation, v_t , with mean 0 and variance ω_v^2 , is not a fundamental driving process – it contains information on current and lagged permanent and transitory income shocks. Equating the variances and autocorrelation coefficients of the original and derived processes (55) and (56), we have

$$\omega_v^2 = \frac{\omega_\epsilon^2}{\alpha} \text{ and } \alpha = -\frac{1 - \sqrt{1 - 4\rho^2}}{2\rho},$$

where $\rho = -\omega_\epsilon^2 / (\omega^2 + 2\omega_\epsilon^2)$ and $\alpha \in [0, 1]$ will be large if the variance of the transitory shock ω_ϵ^2 is large relative to the variance of the permanent shock ω^2 and will converge to 0 as ω_ϵ^2 approaches to 0.

Following the same procedure in Section 4.1, the new state variable and the original budget constraint can be written as:

$$s_t = a_t + \frac{1 - \delta}{R} k_{t-1} + \frac{1}{R - 1} y_t^p - \frac{\alpha}{R - 1} v_t,$$

$$s_{t+1} = R s_t - c_t - (1 - \beta(1 - \delta)) k_t + \zeta_{t+1},$$

respectively, where $\zeta_{t+1} = [(R - \alpha) / (R - 1)] v_{t+1}$. The expressions for changes in aggregate non-

⁴⁵Note that $\frac{sd(\Delta E)}{sd(\Delta C)} = \frac{sd(\Delta E)}{sd(\Delta Y)} / \frac{sd(\Delta C)}{sd(\Delta Y)}$.

durables and durables can then be rewritten as:

$$\Delta C_t = H_c \frac{R - \alpha}{R(R - 1)} \left[\frac{\theta \varepsilon_{t+1}}{(1 - (1 - \theta)R \cdot L)(1 - \alpha \cdot L)} \right], \quad (57)$$

$$\Delta K_t = \frac{1 - \beta(1 - \delta)}{\beta \varrho} \Delta C_t, \quad (58)$$

respectively, where the iid idiosyncratic noises in the expressions for individual consumption dynamics are canceled out. These equations bring out two salient points in our extended model. First, both SID and incomplete information provide endogenous propagation mechanisms of the model – they are characterized by the two factors, $\frac{1}{1 - (1 - \theta)R \cdot L}$ and $\frac{1}{1 - \alpha \cdot L}$, respectively, and thus contribute to the stickiness of aggregate nondurable and durable consumption. Second, under incomplete information, the presence of the transitory shock plays a role in strengthening the inertial responses to the aggregate income shock because α is a function of the variance of the transitory shock. If α is a large value, the effect will be initially small but highly persistent. However, given that $\omega^2 = 125.7^2$ and $\omega_\varepsilon^2 = 2.39^2$ in our estimation using the U.S. data from 1955 – 2007, we can easily calculate that $\alpha = 3.61 \times 10^{-4}$ which is close to 0.⁴⁶ In other words, given the estimated income process, the propagation mechanism in the IC model is extremely weak and the expressions for the changes in nondurable and durable consumption are almost identical to that we obtained in our benchmark model. Note that when $\theta = 1$, these two expressions reduce to $\Delta C_t = H_c \frac{R - \alpha}{R(R - 1)} \frac{\varepsilon_{t+1}}{1 - \alpha \cdot L}$ and $\Delta K_t = \frac{1 - \beta(1 - \delta)}{\beta \varrho} \Delta C_t$, which are almost identical to that obtain in the FI-RE case (see the results reported in Tables 4-5).

6.3. Observational Equivalence between the Benchmark Model and the CARA Model

In this subsection, we consider an SID model with CARA utility and durable consumption, and show that the CARA model and the benchmark SID model presented in Section 4.1 are observationally equivalent in the sense that they lead to the same dynamics of aggregate consumption and savings. Following Caballero (1990b), the typical consumer has the following utility function with constant coefficient of absolute risk aversion (CARA): $u(c_t, k_t) = -\exp(-\alpha_c c_t) / \alpha_c - \varrho \exp(-\alpha_k k_t) / \alpha_k$, where $\alpha_c > 0$ and $\alpha_k > 0$. Following the same procedure adopted in Caballero (1990b) and incorporating the SID assumption into the CARA model, we can solve for the follow-

⁴⁶Using the 1955 – 2012 data set, we obtain the similar result.

ing consumption and durable accumulation functions:⁴⁷

$$\begin{aligned} c_t &= H_c \widehat{s}_t + \Omega_c + \Pi_c, \\ k_t &= \Omega_k + \frac{\alpha_c}{\alpha_k} c_t, \end{aligned}$$

where H_c is the same as that obtained in the benchmark model, $\Omega_c = -\frac{R+\delta-1}{R} \left[1 + \frac{(R+\delta-1)^2}{\varrho R}\right]^{-1} \Omega_k$, $\Omega_k = -\frac{1}{\alpha_k} \ln\left(\frac{R+\delta-1}{\varrho}\right)$, $\Pi_c = -\frac{1}{R-1} \Phi$, and

$$\Phi = \frac{1}{\alpha_c} \ln(\beta R) + \frac{1}{2} \alpha_c \left[1 + \frac{(R+\delta-1)^2}{\varrho R}\right]^{-2} \omega^2.$$

In addition, given the expression for individual saving, $d_t (\equiv (R-1)a_t + y_t - c_t - (k_t - (1-\delta)k_{t-1}))$, following the same aggregation procedure presented in the last section, aggregating across all consumers yields the following expression for aggregate saving:

$$\bar{d}_t = \left[1 - \frac{1 + (R+\delta-1)/\varrho}{1 + (R+\delta-1)^2/(\varrho R)}\right] \frac{\theta \zeta_t}{1 - (1-\theta)R \cdot L} + \frac{(R-1)(1-\theta)\zeta_t}{1 - (1-\theta)R \cdot L} + \left[\frac{R}{H_c} - \left(1 + \frac{R+\delta-1}{\varrho}\right)\right] \Phi. \quad (59)$$

Based on (59) and the corresponding expression for aggregate saving in the benchmark model in which $\bar{d}_t = \left[1 - \frac{1+(R+\delta-1)/\varrho}{1+(R+\delta-1)^2/(\varrho R)}\right] \frac{\theta \zeta_t}{1-(1-\theta)R \cdot L} + \frac{(R-1)(1-\theta)\zeta_t}{1-(1-\theta)R \cdot L}$, we have the following proposition:

Proposition 7. *When*

$$\beta^* = \frac{1}{R} \exp\left(-\frac{1}{2} \alpha_c^2 \left[1 + \frac{(R+\delta-1)^2}{\varrho R}\right]^{-2} \omega_\eta^2\right) < \frac{1}{R}, \quad (60)$$

the benchmark model and the CARA model with SID and durables are observationally equivalent in the sense that they lead to the same expressions for the dynamics of aggregate consumption and saving.

Proof. Given that $\frac{R}{H_c} - \left(1 + \frac{R+\delta-1}{\varrho}\right) \neq 0$, setting $\Phi = 0$ yields (60). ■

The intuition behind this result is simple. The constant precautionary saving demand due to the interaction of exponential utility and fundamental uncertainty and constant dissavings due to impatience (smaller β) cancel out. This result is also emphasized in Wang (2003) in a general equilibrium setting in which the consumers only face idiosyncratic income shocks. Luo and Young

⁴⁷See Online Appendix for the derivation.

(2010) derive a similar result in a model with agents that are averse to model misspecification.

6.4. Discussion on Fixed Costs and Infrequent Adjustment

The previous section shows that introducing SID in an otherwise standard PIH model can better explain the joint dynamics of durable and nondurable goods at the aggregate level. However, at the individual level, the benchmark SID model cannot capture the observed inertial behavior, i.e., infrequent and lumpy purchases on durables in the micro-level data. Bar-Ilan and Blinder (1992) show in an FI-RE model that a consumer with full information about the state chooses to adjust when the welfare improvements from adjusting are greater than fixed costs induced by adjusting, and it is always optimal for the consumer to adjust every period if the fixed cost is zero.

The key feature of our benchmark SID model is that all individuals facing noisy signals will adjust their durable consumption gradually in every period. In this case, the typical consumer under SID suffers from welfare losses due to his or her incomplete consumption adjustments. Following the sticky expectations literature (e.g., Carroll 2003, Carroll and Slacalek 2006), we can also calculate the welfare losses due to deviation from the first-best instantaneously adjusted path when the consumer under SID updates his information set and adjusts optimal plans with an exogenously given probability. That is, the consumer featuring both learning the state slowly and adjusting infrequently suffers from two types of welfare losses: one from incomplete adjustment and the other from infrequent adjustment. In this case, introducing fixed adjustment costs can endogenize the probability of re-adjusting. Specifically, with fixed costs, consumers will optimally choose not to adjust durable goods stocks in every period, while when they make adjustments, they still adjust in a gradual way. In Online Appendix, we show in a SID model in which $u(k_t) = -(\bar{k} - k_t)^2 / 2$ that if the fixed cost (F) is small enough ($F < F^* \equiv H_k^2 \omega_\zeta^2 / 2$, where H_k is the MPC out of perceived permanent income), it is optimal for the consumer to adjust in each period. This conclusion is similar to that obtained in the FI-RE model in Bar-Ilan and Blinder (1992). That is, regardless of full-information or imperfect information about the state, it is always optimal for the consumer to adjust every period if the fixed cost is sufficiently small.

Finally, if the value of fixed cost, F , is above F^* , the consumers in the economy would always choose to adjust optimal consumption infrequently. In this case, the dynamics of aggregate durable consumption under infrequent adjustment (K_t^{ia}) can be written as

$$\Delta K_t^{ia} = \frac{\pi \theta H_k \zeta_t}{(1 - (1 - \pi) \cdot L) (1 - (1 - \theta) R \cdot L)}, \quad (61)$$

where $\pi^* = H_k \omega_{\zeta} / (\beta \sqrt{2F}) + (\beta - 1) / \beta \in (0, 1]$ is the optimal frequency of adjustment.⁴⁸ This expression clearly shows that infrequent adjustment at the microeconomic level due to fixed costs lead to additional stickiness in durable consumption at the aggregate level.⁴⁹

7. Conclusion

This paper has examined the implications of slow information diffusion (SID) for the joint dynamics of aggregate nondurable and durable consumption. In particular, we have shown that the models with slow information better explain the following aspects in the aggregate data: (i) the relative volatility of aggregate nondurable to durable consumption expenditures, (ii) the first-order serial correlation of aggregate expenditures on durables, and (iii) the contemporaneous correlation between nondurable and durable expenditures. In addition, we show that incorporating fixed cost into the benchmark SID model can better characterize the observed behavior of durable consumption at both the micro- and macro-levels.

More work clearly needs to be done. The restriction to quadratic utility may limit the generality of our results, since it rules out the precautionary behavior that seems important at the micro-level (see Carroll and Samwick 1998). However, solving information-constrained consumer problems in their full nonlinear generality has proven difficult (see Sims 2006 or Tutino 2013); whether our results continue to hold when such precautionary considerations are incorporated is an open question. As noted earlier, it is likely our results survive the introduction of aversion to model misspecification (as in Luo and Young 2010) given the range of observationally equivalent results found in that paper.

8. Appendix

8.1. AppendixA: Deriving the Optimal Decisions under FI-RE

We can formulate the Lagrange function as follows:

$$L = E_0 \left[\sum_{t=0}^{\infty} \beta^t \left(-\frac{1}{2} (\bar{c} - c_t)^2 - \frac{\theta}{2} (\bar{k} - k_t)^2 - \lambda_t \{a_{t+1} - [Ra_t + (1 - \delta) k_t - k_{t+1} + y_t - c_t]\} \right) \right],$$

⁴⁸See Online Appendix for the derivation of ΔK_{t+1}^{ia} .

⁴⁹It is straightforward to show that if $F < F^*$, (61) reduces to $\Delta K_t^{ia} = \frac{\theta H_k \zeta_t}{1 - (1 - \theta) R \cdot L}$.

where λ_t is the Lagrange multiplier. The first-order conditions are

$$\lambda_t = \bar{c} - c_t, \quad (62)$$

$$\lambda_t = \beta R E_t [\lambda_{t+1}], \quad (63)$$

$$\lambda_t = \beta E_t \left[\varrho \left(\bar{k} - k_{t+1} \right) + (1 - \delta) \lambda_{t+1} \right], \forall t. \quad (64)$$

Assuming $\beta R = 1$, the above first-order conditions mean

$$\begin{aligned} c_t &= E_t [c_{t+1}], \\ E_t \left[\bar{k} - k_{t+1} \right] &= \frac{R + \delta - 1}{\varrho} (\bar{c} - c_t) \end{aligned}$$

Substituting them into (5) and taking the conditional expectation on both sides gives the optimal decisions for nondurables and durables, (7) and (8), in the text.

8.2. Appendix B: Solving for the Value Function and the Consumption Function under SID

First, conjecture that the value function under SID takes the following quadratic form:

$$v(\hat{s}_t) = A_0 + A_1 \hat{s}_t + A_2 \hat{s}_t^2, \quad (65)$$

where A_0, A_1 , and A_2 are constants to be determined. Substituting Equations (65) into the Bellman equation, (34), in the text yields

$$A_0 + A_1 \hat{s}_t + A_2 \hat{s}_t^2 = \max_{c_t, k_t} \left\{ -\frac{1}{2} (\bar{c} - c_t)^2 - \frac{\varrho}{2} (\bar{k} - k_t)^2 + \beta E_t [A_0 + A_1 \hat{s}_{t+1} + A_2 \hat{s}_{t+1}^2] \right\}, \quad (66)$$

subject to

$$\hat{s}_{t+1} = R \hat{s}_t - c_t - [1 - \beta(1 - \delta)] k_t + \eta_{t+1}. \quad (67)$$

Second, performing the indicated optimization yields the following efficiency condition:

$$\bar{k} - k_t = \frac{R + \delta - 1}{\varrho R} (\bar{c} - c_t).$$

Substituting this condition into (66) and (67) yields:

$$A_0 + A_1 \hat{s}_t + A_2 \hat{s}_t^2 = \max_{c_t} \left\{ -\frac{1}{2} \Omega (\bar{c} - c_t)^2 + \beta E_t [A_0 + A_1 \hat{s}_{t+1} + A_2 \hat{s}_{t+1}^2] \right\}, \quad (68)$$

where $\Omega = 1 + (R + \delta - 1)^2 / (\rho R^2)$, subject to:

$$\widehat{s}_{t+1} = R\widehat{s}_t - \Omega c_t - [1 - \beta(1 - \delta)] \left(\bar{k} - \frac{R + \delta - 1}{\rho R} \bar{c} \right) + \eta_{t+1}. \quad (69)$$

The FOC with respect to c_t yields:

$$c_t = \frac{-2RA_2}{R - 2A_2\Omega} \widehat{s}_t + \frac{R\bar{c} - A_1}{R - 2A_2\Omega}. \quad (70)$$

Substituting it into (69) yields the following state transition equation:

$$\widehat{s}_{t+1} = \frac{R^2}{R - 2A_2\Omega} \widehat{s}_t - \Omega \frac{R\bar{c} - A_1}{R - 2A_2\Omega} + \eta_{t+1}$$

Substituting these expressions back into Equation (68) to arrive at the following equation:

$$\begin{aligned} A_0 + A_1\widehat{s}_t + A_2\widehat{s}_t^2 &= -\frac{1}{2}\Omega \left(\frac{-2RA_2}{R - 2A_2\Omega} \widehat{s}_t + \frac{R\bar{c} - A_1}{R - 2A_2\Omega} - \bar{c} \right)^2 + \beta E_t [A_0 + A_1\widehat{s}_{t+1} + A_2\widehat{s}_{t+1}^2] \\ &= -\frac{1}{2}\Omega \left[\left(\frac{-2RA_2}{R - 2A_2\Omega} \right)^2 \widehat{s}_t^2 + 2 \left(\frac{-2RA_2}{R - 2A_2\Omega} \right) \left(\frac{2A_2\Omega\bar{c} - A_1}{R - 2A_2\Omega} \right) \widehat{s}_t + \left(\frac{2A_2\Omega\bar{c} - A_1}{R - 2A_2\Omega} \right)^2 \right] \\ &\quad + \beta A_0 + \beta A_1 E_t \left[\frac{R^2}{R - 2A_2\Omega} \widehat{s}_t - \Omega \frac{R\bar{c} - A_1}{R - 2A_2\Omega} \right] \\ &\quad + \beta A_2 \left\{ \left(\frac{R^2}{R - 2A_2\Omega} \right)^2 \widehat{s}_t^2 + \frac{2R^2}{R - 2A_2\Omega} \left(-\Omega \frac{R\bar{c} - A_1}{R - 2A_2\Omega} \right) \widehat{s}_t + \left(\Omega \frac{R\bar{c} - A_1}{R - 2A_2\Omega} \right)^2 + \text{var}(\eta_{t+1}) \right\}, \end{aligned}$$

where $\eta_{t+1} = \theta \left[\left(\frac{\zeta_{t+1}}{1 - (1 - \theta)R \cdot L} \right) + \left(\zeta_{t+1} - \frac{\theta R \zeta_t}{1 - (1 - \theta)R \cdot L} \right) \right]$. Finally, after collecting and matching the \widehat{s}_t^2 , \widehat{s}_t , and constant terms, respectively, the undetermined coefficients turn out to be:

$$A_2 = -\frac{R(R - 1)}{2\Omega}, A_1 = R\bar{c}, A_0 = -\frac{R\Omega}{2(R - 1)}\bar{c}^2 - \frac{R}{2\Omega} \text{var}(\eta_{t+1}).$$

Substituting these back into Equations (65) and (70) yields the value function, (37), and the consumption function, (35), of the text.

8.3. Appendix C: Aggregation Mechanism

Consider the expression for individual consumption dynamics:

$$\Delta c_t = \theta H_c \left[\frac{\zeta_t}{1 - (1 - \theta)R \cdot L} + \left(\zeta_t - \frac{\theta R \zeta_{t-1}}{1 - (1 - \theta)R \cdot L} \right) \right]$$

where ζ_t is the iid RI-induced noise shock. Given this expression, the cross-sectional sample average given the sample size is n can be written as:

$$E_{T_n,t} = \frac{1}{n} \sum_{i=1}^n \left(\zeta_{i,t} - \frac{\theta R \zeta_{i,t-1}}{1 - (1 - \theta)R \cdot L} \right),$$

and the corresponding variance is

$$V_n = \frac{\omega_{\zeta}^2}{n^2} \sum_{i=1}^n \left[1 + \frac{(\theta R)^2}{1 - ((1 - \theta)R)^2} \right] = \frac{\omega_{\zeta}^2}{n} \left[1 + \frac{(\theta R)^2}{1 - ((1 - \theta)R)^2} \right],$$

which clearly shows that as long as $1 - ((1 - \theta)R)^2$ is finite, the idiosyncratic components, $\zeta_{i,t}$ and $\frac{\theta R \zeta_{i,t-1}}{1 - (1 - \theta)R \cdot L}$, converge to zero in mean square as n goes to infinity. It is worth noting that this result is consistent with that obtained in Uhlig (1996) in which he shows that the measurability problem pointed out by Judd (1985) can be avoided by requiring convergence in mean square rather than convergence almost everywhere.

References

- [1] Adda, Jerome and Russell Cooper (2006). "The Dynamics of Car Sales: A Discrete Choice Approach." NBER Working Paper #7785.
- [2] Angeletos, George-Marios and Jennifer La'O (2010). "Noisy Business Cycles." *NBER Macroeconomics Annual 2009*, Vol. 24, 319-378.
- [3] Angeletos, George-Marios and Jennifer La'O (2013). "Sentiments." *Econometrica*, 81, 739-779
- [4] Angeletos, George-Marios and Jennifer La'O (2012). "Optimal Monetary Policy with Informational Frictions." mimeo, MIT.
- [5] Attanasio, Orazio (2000). "Consumer Durables and Inertial Behavior: Estimation and Aggregation of (S,s) Rules for Automobile Purchases." *Review of Economic Studies*, 67, 659-77.
- [6] Bar-Ilan, Avner and Alan Blinder (1992). "Consumer Durables: Evidence on the Optimality of Doing Nothing." *Journal of Money, Credit and Banking*, 24, 253-72.
- [7] Bernanke, Ben (1985). "Adjustment Costs, Durables, and Aggregate Consumption." *Journal of Monetary Economics*, 15, 41-68.
- [8] Bertola, Giuseppe and Caballero, Ricardo (1990). "Kinked Adjustment Costs and Aggregate Dynamics." *NBER Macroeconomics Annual 1990*, Vol. 5, 237-288.

- [9] Caballero, Ricardo (1990a). "Expenditure on Durable Goods: A Case for Slow Adjustment." *Quarterly Journal of Economics*, 105, 727-743.
- [10] Caballero, Ricardo (1990b). "Consumption Puzzles and Precautionary Savings." *Journal of Monetary Economics*, 25, 113-136.
- [11] Caballero, Ricardo (1993). "Durable Goods: An Explanation for Their Slow Adjustment." *Journal of Political Economy*, 101, 351-384.
- [12] Caballero, Ricardo (1994). "Notes on the Theory and Evidence on Aggregate Purchases of Durable Goods." *Oxford Review of Economic Policy*, 10, 107-117.
- [13] Carroll, Christopher (2003). "Macroeconomic Expectations of Households and Professional Forecasts." *Quarterly Journal of Economics* 118, 269-298.
- [14] Carroll, Christopher and Jiri Slacalek (2006). "Sticky Expectations and Consumption Dynamics." mimeo, Johns Hopkins University.
- [15] Deaton, Angus (1992). *Understanding Consumption*. Oxford University Press.
- [16] Dynan, Karen (2000). "Habit Formation in Consumer Preferences: Evidence from Panel Data." *American Economic Review*, 90, 391-406.
- [17] Eberly, Janice (1994). "Adjustment of Consumers' Durables Stocks: Evidence from Automobile Purchases." *Journal of Political Economy*, 102, 403-436.
- [18] Feldman, Mark and Christian Gilles (1985). "An Expository Note on Individual Risk without Aggregate Uncertainty." *Journal of Economic Theory*, 35, 26-32.
- [19] Galí, Jordi (1993). "Variability of Durable and Nondurable Consumption: Evidence for Six O.E.C.D. Countries." *Review of Economics and Statistics*, 75, 418-428.
- [20] Hall, Robert E (1978). "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence." *Journal of Political Economy*, 86, 971-987.
- [21] Hansen, Lars P. (1987). "Calculating Asset Prices in Three Example Economies." in Bewley, Truman F. ed., *Advances in Econometrics, Fifth World Congress*. Cambridge University Press.
- [22] Judd, Kenneth L. (1985). "A Law of Large Numbers with a Continuum of IID Random Variables." *Journal of Economic Theory*, 35, 19-25.
- [23] Lam, Pok-Sang (1991). "Permanent Income, Liquidity, and Adjustments of Automobile Stocks: Evidence From Panel Data." *Quarterly Journal of Economics*, 106, 203-230.

- [24] Landauer, Thomas K. (1986). "How Much Do People Remember? Some Estimates of the Amount of Learned Information in Long-Term Memory." *Cognitive Science*, 10, 477-493.
- [25] Lorenzoni, Guido (2009). "A Theory of Demand Shocks." *American Economic Review*, 99, 2050-2084.
- [26] Lucas, Robert (1972). "Expectations and the Neutrality of Money." *Journal of Economic Theory*, 4, 103-124.
- [27] Luo, Yulei (2008). "Consumption Dynamics under Information Processing Constraints." *Review of Economic Dynamics*, 11, 366-385.
- [28] Luo, Yulei and Eric R. Young (2010). "Risk-sensitive Consumption and Savings under Rational Inattention." *American Economic Journal: Macroeconomics*, 2, 281-325.
- [29] Luo, Yulei and Eric R. Young (2014). "Signal Extraction and Rational Inattention." *Economic Inquiry*, 52, 811-829.
- [30] Maćkowiak, Bartosz and Mirko Wiederholt (2009). "Optimal Sticky Prices under Rational Inattention." *American Economic Review*, 99, 769-803.
- [31] Mankiw, Gregory (1982). "Hall's Consumption Hypothesis and Durable Goods." *Journal of Monetary Economics*, 10, 417-425.
- [32] Mankiw, Gregory and Ricardo Reis (2002). "Sticky Information Versus Sticky Prices: A Proposal To Replace The New Keynesian Phillips Curve." *Quarterly Journal of Economics*, 117, 1295-1328.
- [33] Monacelli, Tommaso (2009). "New Keynesian Models, Durable Goods, and Collateral Constraints." *Journal of Monetary Economics*, 56, 242-254.
- [34] Muth, John (1960). "Optimal Properties of Exponentially Weighted Forecasts." *Journal of the American Statistical Association*, 55, 299-306.
- [35] Pischke, Jörn-Steffen (1995). "Individual Income, Incomplete Information, and Aggregate Consumption." *Econometrica*, 63, 805-840.
- [36] Quah, Danny (1990). "Permanent and Transitory Movements in Labor Income: An Explanation for "Excess Smoothness" in Consumption." *Journal of Political Economy*, 98, 449-475.
- [37] Reis, Ricardo (2006). "Inattentive Consumers." *Journal of Monetary Economics*, 53, 1761-1800.
- [38] Sims, Christopher A. (2003). "Implications of Rational Inattention." *Journal of Monetary Economics*, 50, 665-690.

- [39] Sims, Christopher A. (2006). "Rational Inattention: Beyond the Linear-Quadratic Case." *American Economic Review*, 96, 158-163.
- [40] Sims, Christopher A. (2010). "Rational Inattention and Monetary Economics." in Friedman, B.M. and M. Woodford (Eds.), *Handbook of Monetary Economics*, Vol. 3, 155-181.
- [41] Sun, Yeneng (2006). "The Exact Law of Large Numbers via Fubini Extension and Characterization of Insurable Risks." *Journal of Economic Theory*, 126, 31-69.
- [42] Tutino, Antonella (2013). "Rationally Inattentive Consumption Choices." *Review of Economic Dynamics*, 16, 421-439.
- [43] Uhlig, Harald (1996). "A Law of Large Numbers for Large Economies." *Economic Theory*, 8, 41-50.
- [44] Wang, Neng (2003). "Caballero Meets Bewley: The Permanent-Income Hypothesis in General Equilibrium." *American Economic Review*, 93, 927-936.
- [45] Zaffaroni, Paolo (2004). "Contemporaneous Aggregation of Linear Dynamic Models in Large Economies." *Journal of Econometrics*, 120, 75-102.

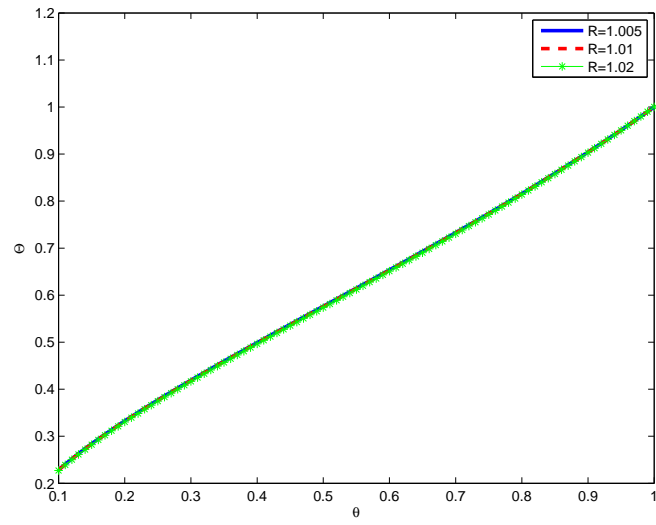


Figure 1. Relative Volatility of ΔC to ΔY

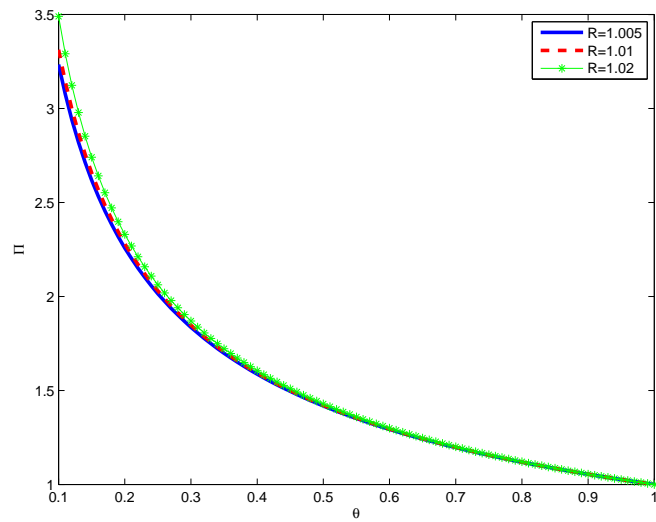


Figure 2. Relative Volatility of ΔE to ΔC

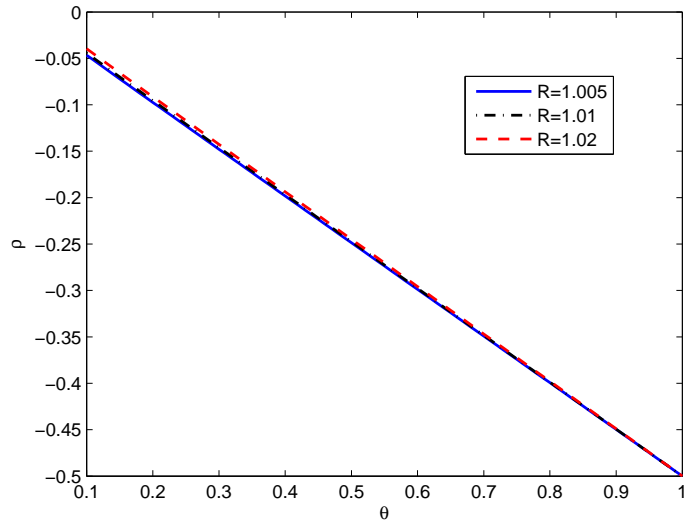


Figure 3. First-order Autocorrelation of ΔE

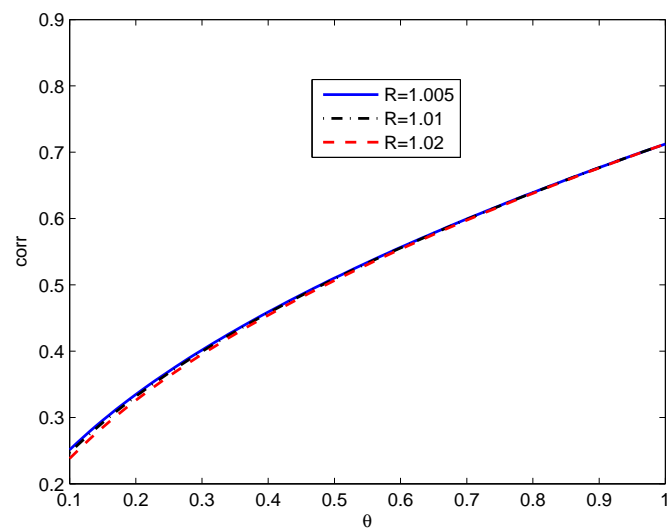


Figure 4. Correlation between ΔE and ΔC

Table 1. Summary of Key Moments in the Data (1955 – 2007)

Name of the Statistic	Value (s.e.)
Relative volatility of ΔC to ΔY $\left(\frac{sd(\Delta C)}{sd(\Delta Y)}\right)$	0.66 (0.02)
Relative volatility of ΔE to ΔC $\left(\frac{sd(\Delta E)}{sd(\Delta C)}\right)$	0.62 (0.08)
Autocorrelation of ΔE ($\rho_{\Delta E}$)	-0.30 (0.07)
Correlation between ΔE and ΔC ($corr(\Delta E, \Delta C)$)	0.46 (0.08)

Table 2. Summary of Key Moments in the Data (1955 – 2012)

Name of the Statistic	Value (s.e.)
Relative volatility of ΔC to ΔY $\left(\frac{sd(\Delta C)}{sd(\Delta Y)}\right)$	0.64 (0.02)
Relative volatility of ΔE to ΔC $\left(\frac{sd(\Delta E)}{sd(\Delta C)}\right)$	0.76 (0.08)
Autocorrelation of ΔE ($\rho_{\Delta E}$)	-0.03 (0.12)
Correlation between ΔE and ΔC ($corr(\Delta E, \Delta C)$)	0.46 (0.07)

Table 3. Parameter Values

Parameter	Value	Targets
R	1.01	annual interest rate of 4%
β	0.99	$\beta R = 1$
ϱ	0.029	Bernanke (1985)
δ	0.015	Bernanke (1985), Monacelli (2009)

Table 4. Model Comparison (1955 – 2007)

	Data	RE ($\theta = 1$)	RI ($\theta = 0.62$)	AC	IC	Improvement
$\frac{sd(\Delta C)}{sd(\Delta Y)}$	0.66	0.98	0.66	0.98	0.98	100%
$\frac{sd(\Delta E)}{sd(\Delta C)}$	0.62	1.21	0.95	0.60	1.21	44%
$\rho_{\Delta E}$	-0.30	-0.50	-0.31	-0.31	-0.50	95%
$corr(\Delta E, \Delta C)$	0.46	0.71	0.51	0.70	0.71	80%

Table 5. Model Comparison (1955 – 2012)

	Data	RE ($\theta = 1$)	RI ($\theta = 0.6$)	AC	IC	Improvement
$\frac{sd(\Delta C)}{sd(\Delta Y)}$	0.64	0.98	0.64	0.98	0.98	100%
$\frac{sd(\Delta E)}{sd(\Delta C)}$	0.76	1.21	0.93	0.62	1.21	62%
$\rho_{\Delta E}$	-0.03	-0.50	-0.30	-0.30	-0.50	43%
$corr(\Delta E, \Delta C)$	0.46	0.71	0.51	0.71	0.71	80%

Note: The values of “Improvement” in Tables 4 and 5 are calculated using $(1 - \frac{|M_{RI} - M_{data}|}{|M_{RE} - M_{data}|}) \times 100\%$.