Sticky Information Diffusion and the Inertial Behavior of Durables Consumption*

Yulei Luo†
UC Davis and University of Hong Kong

Eric R. Young‡
University of Virginia

October 28, 2009

Abstract

This paper studies the dynamics of durable and nondurable consumption under two alternative assumptions about information updating by households – rational inattention and sticky expectations. We first show that the two types of sticky information diffusion can help generate strong excess smoothness in durables consumption. We then find that sticky expectations due to a fixed cost does a better job of reproducing the infrequent adjustments of durables consumption at the individual level and the slow adjustments at the aggregate level. Finally, we show that finite information-processing capacity can by itself lead to infrequent adjustments even if there is no adjustment cost.

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

Keywords: Durability, Rational Inattention, Infrequent Adjustments, Consumption Stickiness.

*We thank Taeyoung Doh, Edward Knotek II, Jun Nie, Jordan Rappaport, Tom Sargent, Jonathan Willis and seminar participants at Federal Reserve Bank of Kansas City and Washington and Lee University for helpful comments and Wei Li for valuable research assistance. Luo thanks the General Research Fund (GRF) in Hong Kong for financial support. Young thanks the Bankard Fund for Political Economy at Virginia for financial support. All remaining errors are our responsibility.

†Corresponding author. School of Economics and Finance, University of Hong Kong, Hong Kong. Email address: `yluo@econ.hku.hk`.

‡ Department of Economics, University of Virginia, Charlottesville, VA 22904. Email address: `ey2d@virginia.edu`.

1. Introduction

Mankiw (1982) argues that the permanent income hypothesis (PIH) model with durable goods is inconsistent with empirical evidence. In particular, he shows that in Hall (1978)'s PIH model in which utility is a quadratic function of the stock of durable goods, the optimal expenditure on durables should follow an ARMA$(1,1)$ process, where the MA component is determined by the depreciation rate of durable goods. He then estimates an ARMA$(1,1)$ model of expenditures on durable goods using quarterly U.S. data and finds that expenditures on durable goods follows an AR(1) process. As we discuss here, to make the model fit the data the depreciation rate should be about 100 percent; this finding is called the Mankiw puzzle in the literature. In addition, as documented in Caballero (1994), in the presence of durable goods the PIH model predicts that the change in the stock of durable goods should also be unpredictable white noise.[1] However, the change in the stock of durables has strong positive serial correlation in the post-war U.S. quarterly data. He then argued that this rejection is an order of magnitude larger than that for nondurables (the rejection of the martingale property of aggregate consumption) and is robust across of categories of durable goods. Over the past two decades there have been a number of papers examining Mankiw's results. Caballero (1990, 1994) and Adda and Cooper (2006) find that the Mankiw puzzle is robust across different time periods, different frequencies, and different countries.

Bernanke (1985) studies the joint behavior of nondurables and durables consumption in the presence of adjustment costs of changing durables stocks within a simple representative agent framework. He finds that the costs of adjusting durables stocks are substantial and can help improve the model's prediction for the joint behavior of aggregate consumption and income.[2] 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.[3] 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

[1] Hall (1978) shows that under the PIH, the change in nondurable consumption is unpredictable.
[2] We will revisit Bernanke's model in Section 2.
[3] Lam (1991) reports that households only occasionally adjust their stock of durables.

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.

We take an alternative approach to the Mankiw puzzle, one based on information frictions at the micro-level. Specifically, we study a permanent income model with durable goods and examine implications of two types of imperfect information diffusion – rational inattention (henceforth RI) and sticky expectations (SE) – for the joint dynamics of nondurables and durables consumption at both micro- and macro-levels. Rational inattention (RI), a consequence of information-processing constraints, 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.[4] 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. The idea of SE is to relax the assumption that all consumers' expectations are completely updated at every period; specifically, only a fraction of the population update their expectations on permanent income and re-optimize in any given period.[5]

We find that although both RI and SE can improve the model's predictions at the aggregate level, SE is a better candidate to characterize the inertial behavior at the individual level (infrequent and lumpy purchases on durables). Specifically, both hypotheses improve the model in the following aspects: (1) they reduce the relative volatility of aggregate nondurables to durables consumption, (2) they increase the first-order serial correlation of expenditures on aggregate durables consumption, and (3) they also increase the sum of MA coefficients in the process of durables consumption. In particular, we show that both hypotheses have the potential to explain the Mankiw puzzle. The reason that RI cannot capture the behavior of individual consumers is that consumers under information-processing constraints adjust their durable stock gradually in response to income shocks, which is inconsistent with the observed intermittent large adjustments in the micro-level data. We then show that heterogeneity in channel capacity can be used to endogenize SE and provide a micro-foundation for infrequent adjustments. Therefore, an aggregate model composed by a continuum of consumers with heterogenous degrees of attention can capture both infrequent adjustments at the individual level and gradual adjustments at the

[4] Luo (2008), Luo and Young (2009b), and Tutino (2009) use the RI framework to examine the dynamics of nondurable consumption. There are a number of other papers as well that study business cycle dynamics, including Luo and Young (2009a), Maćkowiak and Weiderholdt (2009), and Martins and Sinigaglia (2009).

[5] Reis (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."

aggregate level, and this model is founded on information theory rather than difficult to measure costs of adjustment.

Specifically, the 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 rational expectations. As households gradually learn about a change in the state, their stock of durables will slowly adjust. Indeed, Caballero (1990) explicitly suggests that slow diffusion of information could account for the particular adjustment process he posits. Thus, our rational inattention model provides a simple microfoundation for the slow adjustment used in that paper.

The rest of the paper is organized as follows. Section 2 proposes a stylized permanent income model with durable goods and discuss the model's predictions on the dynamics of durables consumption. Section 3 solves permanent income models with durable goods and two types of sticky information diffusion, and examines stochastic implications of RI and SE for the dynamics of durables consumption at both individual and aggregate levels. Section 4 shows how heterogeneity in channel capacity can endogenize SE and can potentially better explain both micro- and macro-level data. Section 5 concludes.

2. A Stylized Permanent Income Model with Durable Goods

In this section we present a standard 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 two types of sticky information diffusion, rational inattention (RI) and sticky expectations (SE), into this stylized model affects the joint behavior of nondurables and durables 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.

2.1. The Model

Following Mankiw (1982), Bernanke (1985), and Galí (1993), we consider an RE version of the PIH model which integrates both durables and nondurables consumption, where the latter includes both nondurable goods and services.[6] The optimizing decisions of a representative

[6] Although the original Mankiw model only considers durables consumption, including nondurables consumption in preferences does not change his main conclusion provided they enter in a separable manner.

consumer in the RE-PIH model with durable goods can be formulated as follows:

$$\max_{\{c_t, k_{t+1}\}} E_0 \left[\sum_{t=0}^{\infty} \beta^t u(c_t, k_t) \right], \tag{2.1}$$

subject to the budget constraint[7]

$$a_{t+1} = Ra_t + y_t - c_t - e_t, \tag{2.2}$$

and the accumulation equation for durables

$$k_{t+1} = (1 - \delta) k_t + e_t, \tag{2.3}$$

where $u(c_t, k_t) = -\frac{1}{2} (\overline{c} - c_t)^2 - \frac{\varrho}{2} \left(\overline{k} - k_t\right)^2$ is the utility function, $\overline{c}$ and $\overline{k}$ 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). Combining (2.2) and (2.3) gives the period-to-period finance constraint of the consumer:

$$a_{t+1} = Ra_t + (1 - \delta) k_t - k_{t+1} + y_t - c_t. \tag{2.4}$$

Solving this optimization problem gives optimal decisions for nondurable and durable consumption:

$$c_t = H_c s_t \tag{2.5}$$

$$k_{t+1} = E_t [k_{t+1}] = \frac{R + \delta - 1}{\varrho} c_t, \tag{2.6}$$

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

$$H_c = (R - 1) \left(1 + \frac{(1 - \beta (1 - \delta))^2}{\beta \varrho} \right)^{-1} \tag{2.7}$$

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

(see Appendix 6.1 for derivations), and

$$s_t = a_t + \frac{1-\delta}{R}k_t + \frac{1}{R}\sum_{j=0}^{\infty} R^{-j} E_t\left[y_{t+j}\right] \tag{2.8}$$

is the expected present value of lifetime resources, consisting of financial wealth (the risk free foreign bond) plus human wealth.

As shown in Luo (2008), to facilitate the introduction of 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 after introducing rational inattention..[8] Specifically, if s_t is defined as a new state variable, the original finance constraint can be rewritten as

$$s_{t+1} = Rs_t - c_t - (1 - \beta(1-\delta))\, k_{t+1} + \zeta_{t+1}, \tag{2.9}$$

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

$$\zeta_{t+1} = \frac{1}{R}\sum_{j=0}^{\infty}\left(\frac{1}{R}\right)^{j}(E_{t+1} - E_t)\left[y_{t+1+j}\right], \tag{2.10}$$

We close the model by assuming that the income process follows a random walk,

$$y_{t+1} = y_t + \varepsilon_{t+1}, \tag{2.11}$$

where ε_{t+1} is a white noise with mean 0 and variance ω^2. In this case, permanent income s_t, can be written as

$$s_t = a_t + \frac{1-\delta}{R}k_t + \frac{1}{R-1}y_t; \tag{2.12}$$

that is, s_t is a linear combination of three state variables, financial wealth, the stock of durable goods, and labor income. The innovation to permanent income $\zeta_{t+1} = \frac{1}{R-1}\varepsilon_{t+1}$. Combining (2.5), (2.6), with (2.9) gives the expressions for the changes in nondurable goods and the stock

[8] Reduction of the state space to univariate is convenient because we will introduce RI into this model and 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 (we thank Chris Sims for correcting our previous beliefs about the tractability of a multivariate version of the RI-LQ problem).

of durable goods:

$$\Delta c_t = H_c \zeta_t$$
$$\Delta k_{t+1} = \frac{R+\delta-1}{\varrho} H_c \zeta_t.$$

Given the specification of the income process, (2.11), we have

$$\Delta c_t = \left(1 + \frac{(1-\beta(1-\delta))^2}{\beta\varrho}\right)^{-1} \varepsilon_t,$$

which is the random walk result of Hall (1978), and the expenditure on durable goods follows the following ARMA$(1,1)$ process:

$$e_t = e_{t-1} + \varsigma_t - (1-\delta)\varsigma_{t-1}, \tag{2.13}$$

where $\varsigma_t = \frac{R+\delta-1}{\varrho}\left(1 + \frac{(1-\beta(1-\delta))^2}{\beta\varrho}\right)^{-1} \varepsilon_t$ is a shock 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, (2.13) implies that the first-order autocorrelation of Δe_t is simply

$$\rho_{\Delta e_t}(1) = \frac{\delta-1}{1+(1-\delta)^2} < 0$$

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), $\rho_{\Delta e_t}(1) = -0.48$. However, the estimated value of $\rho_{\Delta e_t}(1)$ 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 implying $\delta = 0.99527$).[9] Obviously, a model with this property is going to be difficult to calibrate to observed aggregate data.

[9]The data used for the new estimates is seasonally-adjusted nominal consumer durable expenditures from 1947-2009, deflated using the PCE deflator and first-differenced. Alternative choices for the deflator do not materially affect the results.

2.2. Bernanke's Adjustment Costs Model Revisited

The main difference between the model present in Section 2.1 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

$$u(c_t, k_t, k_{t+1}) = -\frac{1}{2}(\bar{c} - c_t)^2 - \frac{\varrho}{2}\left(\bar{k} - k_t\right)^2 - \frac{\vartheta}{2}(k_{t+1} - k_t)^2, \tag{2.14}$$

where ϑ measures the importance of adjustment costs in utility.[10] Given the utility function (2.14) and the budget constraint (2.4), solving the optimization problem gives decision rules of nondurables and durables:

$$c_t = H_c(a_t + H_k k_t + H_y y_t) + g_0, \tag{2.15}$$

$$k_{t+1} = x_1 k_t + \frac{x_1(1 - \beta(1-\delta))}{d\left(1 - x_2^{-1}\right)} c_t + h_0. \tag{2.16}$$

x_1 and x_2 (which satisfy $x_1 + x_2 = \frac{\beta\varrho + (1+\beta)\vartheta}{\beta\vartheta}$ and $x_1 x_2 = \frac{1}{\beta}$) are two real eigenvalues (suppose $x_1 < x_2$ without loss of generality) for the second-order stochastic difference equation $k_t + hk_{t+1} + \beta E_t[k_{t+2}] = 0$,

$$H_y = \frac{1}{R-1}, \tag{2.17}$$

$$H_k = \beta(1-\delta) + (\beta(1-\delta) - 1)\frac{x_1}{R - x_1}, \tag{2.18}$$

$$H_c = (R-1)\left(1 + (1 - \beta(1-\delta))^2 \frac{R}{R - x_1} \frac{x_1}{\vartheta\left(1 - x_2^{-1}\right)}\right)^{-1}, \tag{2.19}$$

and g_0 and h_0 are irrelevant constant terms.[11]

To obtain the explicit dynamics of nondurables and durables consumption, we define a new state variable, permanent income s_t, as

$$s_t = a_t + H_k k_t + H_y y_t, \tag{2.20}$$

[10] Bernanke (1985) assumes that utility is a non-separable function of nondurables and durables consumption; that is, there is an additional term $-m(\bar{c} - c_t)\left(\bar{k} - k_t\right)$ in the utility function. However, the estimated m, the parameter measuring the degree of non-separability, is insignificantly different from 0. Hence, for simplicity here we assume that $m = 0$ and focus on the effect of adjustment costs.

[11] Note that as ϑ goes to 0, (2.15) and (2.16) reduces to (2.5) and (2.6) because $\lim_{\vartheta\to 0} x_1 = 0$ and $\lim_{\vartheta\to 0} \frac{x_1}{\vartheta\left(1 - x_2^{-1}\right)} = \frac{1}{\beta\varrho}$.

and reformulate the original budget constraint (2.4) as

$$s_{t+1} = s_t + H_y \varepsilon_{t+1}. \tag{2.21}$$

after using (2.15) and (2.16). We can then obtain the dynamics of (c_t, k_{t+1}) by combining (2.15), (2.16), with (2.21):[12]

$$\Delta c_t = H_c H_y \varepsilon_t, \tag{2.22}$$

$$\Delta k_{t+1} = \frac{x_1 (1 - \beta (1 - \delta))}{\vartheta \left(1 - x_2^{-1}\right)} \frac{H_c H_y \varepsilon_t}{1 - x_1 \cdot L}, \tag{2.23}$$

where L is the lag operator.

Equation (2.22) shows that the presence of adjustment costs reduces the initial response of nondurables consumption to the income shock because $H_c H_y < 1$ when $\vartheta > 0$, while it does not affect the dynamic responses of nondurables consumption because it is just the random walk result of Hall (1978).[13] Specifically, the introduction of adjustment costs reduces the relative volatility of nondurables consumption to income, defined as

$$\mu = \frac{\mathrm{sd}\,[\Delta c_t]}{\mathrm{sd}\,[\Delta y_t]} = \left(1 + (1 - \beta (1 - \delta))^2 \frac{R}{R - x_1} \frac{x_1}{\vartheta \left(1 - x_2^{-1}\right)}\right)^{-1} < 1.$$

Using the parameter values estimated in Bernanke (1985) ($R = 1.01$, $\vartheta = 0.706$, $x_1 = 0.828$), we find that $\mu = 0.96$, which is still well above its empirical counterpart. (In US data, μ is around 0.5.) In other words, the introduction of costs of adjusting durable stocks does not improve the model's predictions for the joint behavior of aggregate nondurables consumption and income; in US data nondurables consumption is much smoother than income and is sensitive to past information.[14]

Clearly, (2.23) 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+1} to the income shock when $\vartheta = 0.706$, $\varrho = 0.0286$ and $x_1 = 0.828$.[15] Expression (2.23) also shows that the presence of adjustment costs can improve the model's predictions in the following aspects: (1) it

[12] Note that in this representative-agent setting, individual and aggregate dynamics are identical.

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

[14] These two anomalies have been termed the excess smoothness and excess sensitivity puzzles in the literature. See Deaton (1992) for a recent review.

[15] These parameter values are taken from Table 2 in Bernanke (1985).

increases excess smoothness of durables consumption and (2) it increases the autocorrelation of durables consumption. However, due to the assumption of convex adjustment costs, the model predicts partial and continuous reactions of durables to income shocks, which is inconsistent with the empirical evidence that individual consumers make adjustments of durables stocks only *intermittently* and *lumpily*.

3. Permanent Income Models with Durable Goods and Sticky Information Diffusion

In this section, we propose permanent income models with durable goods and sticky information and explore how two types of information imperfections (RI and SE) affect the dynamic effects of income shocks on the joint behavior of nondurables and durables consumption.

3.1. Consumption Dynamics under RI

In this subsection, we introduce RI into the model proposed in Section 2.1. 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) and Luo (2008), the consumer's information-processing constraint can be characterized by the inequality

$$\mathcal{H}\left(s_{t+1}|\mathcal{I}_t\right) - \mathcal{H}\left(s_{t+1}|\mathcal{I}_{t+1}\right) \leq \kappa, \tag{3.1}$$

where $\mathcal{I}_t$ is the consumer's currently processed information, κ is the consumer's channel capacity, $\mathcal{H}\left(s_{t+1}|\mathcal{I}_t\right)$ denotes the entropy of the state prior to observing the new signal at $t+1$, and $\mathcal{H}\left(s_{t+1}|\mathcal{I}_{t+1}\right)$ is the entropy after observing the new signal.[16] (3.1) 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), $\mathcal{D}_t$ is a normal distribution $N\left(\widehat{s}_t, \sigma_t^2\right)$; as a result, (3.1) can be reduced to

$$\log|\psi_t^2| - \log|\sigma_{t+1}^2| \leq 2\kappa \tag{3.2}$$

where $\sigma_{t+1}^2 = \operatorname{var}\left[s_{t+1}|\ \mathcal{I}_{t+1}\right]$ and $\psi_t^2 = \operatorname{var}\left[s_{t+1}|\ \mathcal{I}_t\right]$ are the posterior and prior variances of the state variable, respectively.[17]

[16] In this section we assume that consumers have the identical degree of inattention. Later we will assume that consumers have heterogenous degrees of inattention and argue that this assumption may be used to rationalize sticky expectations proposed by Carroll (2003) and Carroll and Sommer (2003).

[17] Since observing signals is costless, the constraint is binding in all time periods.

In the univariate case (3.2) has a steady state $\sigma^2 = \frac{\omega_\zeta^2}{\exp(2k)-R^2}$, in which 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 variance $\lambda_t^2 = \operatorname{var}[\xi_{t+1}|\,\mathcal{I}_t]$; in the steady state $\lambda^2 = \left(\sigma^{-2} - \psi^{-2}\right)^{-1}$. As shown in Sims (2003) and Luo (2008), given the LQG specification, the consumption and durable accumulation functions under RI are

$$c_t = H_c \widehat{s}_t \tag{3.3}$$

$$k_{t+1} = \frac{1-\beta(1-\delta)}{\beta\varrho} H_c \widehat{s}_t, \tag{3.4}$$

where H_c is defined in (2.7). The conditional mean $\widehat{s}_t$ evolves according to the Kalman filter equation

$$\widehat{s}_{t+1} = (1-\theta)\widehat{s}_t + \theta(s_{t+1} + \xi_{t+1}), \tag{3.5}$$

where $\theta = 1 - 1/\exp(2\kappa) \in [0,1]$ is the optimal weight on any new observation. Straightforward calculations imply that

$$\Delta c_t = \theta H_c \left[\frac{\zeta_t}{1-(1-\theta)R\cdot L} + \left(\xi_t - \frac{\theta R \xi_{t-1}}{1-(1-\theta)R\cdot L}\right)\right] \tag{3.6}$$

$$\Delta k_{t+1} = \frac{1-\beta(1-\delta)}{\beta\varrho} \Delta c_t, \tag{3.7}$$

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.[18] Expenditure on durable goods follows the process

$$\Delta e_t = \frac{1-\beta(1-\delta)}{\beta\varrho}\theta H_c \left[\begin{array}{c} \left(\zeta_t + \frac{((1-\theta)R-(1-\delta))\zeta_{t-1}}{1-(1-\theta)R\cdot L}\right) + \\ \left(\xi_t - (1-\delta+\theta R)\,\xi_{t-1} + \frac{\theta R((1-\delta)-(1-\theta)R)\xi_{t-2}}{1-(1-\theta)RL}\right) \end{array}\right], \tag{3.8}$$

which reduces to $\Delta e_t = \frac{1-\beta(1-\delta)}{\beta\varrho} H_c \zeta_t$ when $\theta = 1$.

After aggregating over all households under an assumption of identical κ, we obtain the

[18] This assumption only has bite when θ is very close to 0. As shown in Batchuluun, Luo, and Young (2008), households with very low channel capacity display markedly different behavior; for these households the assumption of quadratic utility will not be innocuous.

expressions for changes in aggregate nondurables and durables:

$$\Delta C_t = \theta H_c \left[\left(\frac{\zeta_t}{1-(1-\theta)R \cdot L} \right) + \left(\overline{\xi}_t - \frac{\theta R \overline{\xi}_{t-1}}{1-(1-\theta)R \cdot L} \right) \right], \tag{3.9}$$

$$\Delta K_{t+1} = \frac{1-\beta(1-\delta)}{\beta \varrho} \Delta C_t \tag{3.10}$$

and

$$\Delta E_t = \frac{1-\beta(1-\delta)}{\beta\varrho} \theta H_c \left[\begin{array}{c} \left(\zeta_t + \frac{((1-\theta)R-(1-\delta))\zeta_{t-1}}{1-(1-\theta)R \cdot L} \right) + \\ \left(\overline{\xi}_t - (1-\delta+\theta R)\, \overline{\xi}_{t-1} + \frac{\theta R[(1-\delta)-(1-\theta)R]\overline{\xi}_{t-2}}{1-(1-\theta)R \cdot L} \right) \end{array} \right], \tag{3.11}$$

respectively, where $E^i[\cdot]$ is the population average and $\overline{\xi}_t = E^i[\xi_t]$ is the common noise. As argued in Sims (2003), although the randomness in an individual's response to the aggregate shock should also be *idiosyncratic* because it arises from his *own* internal information-processing constraint, a considerable part of the idiosyncratic error may be common across individuals. The intuition is that people's needs for coding macroeconomic information efficiently are similar, so they rely on common sources of coded information. Therefore, the common term of the idiosyncratic error, $\overline{\xi}_t$, is between 0 and the part of the idiosyncratic error, ξ_t, caused by the common shock to permanent income, ζ_t. Formally, assume that ξ_t consists of two independent noises: $\xi_t = \overline{\xi}_t + \xi_t^i$, where $\overline{\xi}_t = E^i[\xi_t]$ and ξ_t^i are the common and idiosyncratic components of the error generated by ζ_t, respectively. A single parameter,

$$\lambda = \frac{\operatorname{var}\left[\overline{\xi}_t\right]}{\operatorname{var}[\xi_t]} \in [0,1],$$

can be used to measure the common source of coded information on the aggregate component (or the relative importance of $\overline{\xi}_t$ vs. ξ_t).

We can now examine how RI affects some key stochastic properties of aggregate consumption processes derived above. Specifically, we discuss how RI affects (1) the relative volatility of expenditures on durables and nondurables, (2) the first-order autocorrelation of changes in durables expenditures, and (3) the sum of the MA coefficients for durable expenditures. Our interest lies in determining whether the Mankiw puzzle can be resolved by breaking the tight link between the MA coefficient and the depreciation rate implied by the RE assumption.

3.1.1. The Relative Volatility of ΔE_t to ΔC_t

Given (3.9) and (3.11), the volatility of changes in aggregate non-durables and durables consumption can be written as

$$\mathrm{sd}\left[\Delta C_t\right] = \left(\theta H_c\right)\sqrt{\frac{1}{1-\left(\left(1-\theta\right)R\right)^2}+\lambda^2\left(1+\frac{\theta^2R^2}{1-\left(\left(1-\theta\right)R\right)^2}\right)\frac{1-\theta}{\theta\left(1-\left(1-\theta\right)R^2\right)}}\omega_\zeta \tag{3.12}$$

and

$$\mathrm{sd}\left[\Delta E_t\right] = \left(\frac{R+\delta-1}{\varrho}\theta H_c\right)\sqrt{\begin{array}{c}\frac{1+(1-\delta)^2-2(1-\delta)(1-\theta)R}{1-((1-\theta)R)^2}+\\ \lambda^2\left(1+\left(1-\delta+\theta R\right)^2+\frac{\theta^2R^2((1-\delta)-(1-\theta)R)^2}{1-((1-\theta)R)^2}\right)\frac{1-\theta}{\theta(1-(1-\theta)R^2)}\end{array}}\omega_\zeta, \tag{3.13}$$

where $\omega_\zeta^2 = \mathrm{var}\left[\zeta_t\right]$, $\lambda = \frac{\mathrm{var}\left[\overline{\xi}_t\right]}{\mathrm{var}\left[\xi_t\right]} \in [0,1]$ is the aggregation factor, and $\mathrm{var}\left[\xi_t\right] = \frac{1-\theta}{\theta\left[1-(1-\theta)R^2\right]}\omega_\zeta^2$. Therefore, the relative volatility of the changes in durable to nondurable consumption can be defined as follows

$$\mathrm{rv}\left(\Delta E_t, \Delta C_t\right) = \frac{\mathrm{sd}\left[\Delta E_t\right]}{\mathrm{sd}\left[\Delta C_t\right]}, \tag{3.14}$$

where $\mathrm{sd}\left[\Delta C_t\right]$ and $\mathrm{sd}\left[\Delta E_t\right]$ are defined in (3.12) and (3.13), respectively. We first consider a special case in which all noises are idiosyncratic (so that individuals live on isolated islands and do not interact with each other directly or indirectly via conversation, imitation, newspapers, or other media); in this special case ($\lambda = 0$) the variability ratio can be written as

$$\mathrm{rv}\left(\Delta E_t, \Delta C_t\right) = \left(\frac{R+\delta-1}{\varrho}\right)\sqrt{1+\left(1-\delta\right)^2-2\left(1-\delta\right)\left(1-\theta\right)R}, \tag{3.15}$$

which means that the relative volatility of changes in durables and non-durables consumption is increasing with the degree of attention, θ, as

$$\frac{\partial\,\mathrm{rv}\left(\Delta E_t, \Delta C_t\right)}{\partial\theta} > 0.$$

That is, RI increases the excess smoothness of durables goods relative to nondurables consumption. Using data from the US, Canada, the UK, Japan, France, and Italy, Galí (1993) finds that the rejection of the PIH in the durables case is much stronger, in the sense that durables consumption exhibits much stronger excess smoothness. In particular, he shows that the actual variability ratio $\mathrm{rv}\left(\Delta E_t, \Delta C_t\right)$ in the US data is 0.25, whereas the stylized model predicts that

$\text{rv}(\Delta E_t, \Delta C_t) = 2.89$ if $\delta = 0.05$, 1.71 if $\delta = 0.025$, and 0.98 if $\delta = 0.01$.[19] To examine how RI affects the relative variability, define

$$\mu = \frac{\text{rv}(\Delta E_t, \Delta C_t; \theta = 1)}{\text{rv}(\Delta E_t, \Delta C_t; \theta < 1)} = \sqrt{\frac{1 + (1-\delta)^2}{1 + (1-\delta)^2 - 2(1-\delta)(1-\theta)R}}. \tag{3.16}$$

Figure 2 clearly shows that the presence of RI can improve the model's prediction for the observed variability ratio for different values of δ in the data. For example, when $\delta = 0.05$ and $\theta = 10\%$, $\mu = 3.7$. That is, if 10 percent of any new information is transmitted each period, (that is, 10% of the uncertainty is removed upon the receipt of a new signal) the predicted relative variability can be reduced by about 4 times. However, it is worth noting that $\theta = 10\%$ is only a small fraction of individuals' capacity and the model's predicted $\text{rv}(\Delta E_t, \Delta C_t)$ $(= 0.78)$ is still well above the empirical counterpart. In Section 3.4, we will examine how the combination of adjustment costs and rational inattention better fits the data.

We now consider another special case in which $\lambda = 1$.[20] In this case, it is straightforward to show that

$$\text{rv}(\Delta E_t, \Delta C_t) = \left(\frac{R + \delta - 1}{\varrho}\right)\sqrt{1 + (1-\delta)^2}, \tag{3.17}$$

which is independent of the degree of attention. Figure 3 illustrates how RI affects the relative variability, μ for $\lambda = 0.1$, 0.2, and 0.3 when $R = 1.02$ and $\delta = 0.05$. The figures clearly show that the smaller the value of θ, the larger the relative volatility of ΔC_t and ΔE_t, which again improves the model's prediction for the asymmetry in the size of variability across goods.

3.1.2. The First-order Autocorrelation of ΔE_t

By construction, (3.11) can be rewritten as

$$\begin{aligned}\Delta E_t = {} & \varsigma_t + \frac{((1-\theta)R - (1-\delta))\varsigma_{t-1}}{1 - (1-\theta)R \cdot L} \\ & + \left(\frac{R+\delta-1}{\varrho}\theta H_c\right)\left(\overline{\xi}_t - (1-\delta+\theta R)\overline{\xi}_{t-1} + \frac{\theta R((1-\delta)-(1-\theta)R)\overline{\xi}_{t-2}}{1-(1-\theta)R\cdot L}\right),\end{aligned} \tag{3.18}$$

[19] As reported in Galí (1993), the difference between the estimates of predicted and actual $\text{rv}(\Delta E_t, \Delta C_t)$ is also very large for all other countries.

[20] The case of $\lambda = 1$ means that consumers' needs for information coding are exactly the same and they completely rely the common source of coded information.

where $\varsigma_t = \frac{R+\delta-1}{\varrho}\theta H_c \zeta_t$ with $\mathrm{var}[\varsigma_t] = \left(\frac{R+\delta-1}{\varrho}\theta H_c\right)^2 \omega_\zeta^2$. Given (3.18), the first-order autocorrelation of ΔE_t can be written as

$$\rho_{\Delta E_t}(1) = \frac{\mathrm{cov}[\Delta E_{t+1}, \Delta E_t]}{\mathrm{var}[\Delta E_t]},$$

where $\mathrm{var}[\Delta E_t]$ is defined in (3.13) and

$$\mathrm{cov}[\Delta E_{t+1}, \Delta E_t] \tag{3.19}$$
$$= \left(\frac{R+\delta-1}{\varrho}\theta H_c\right)^2 \left\{ \begin{array}{c} \left((1-\theta)R-(1-\delta)+\frac{(1-\theta)R((1-\theta)R-(1-\delta))^2}{1-((1-\theta)R)^2}\right)+ \\ \lambda^2\Upsilon\left[\begin{array}{c} -(1-\delta+\theta R)-\theta R(1-\delta+\theta R)((1-\delta)-(1-\theta)R) \\ +\theta^2(1-\theta)R^3((1-\delta)-(1-\theta)R)^2\frac{1}{1-((1-\theta)R)^2} \end{array}\right] \end{array} \right\} \omega_\zeta^2,$$

where $\Upsilon = \frac{1-\theta}{\theta[1-(1-\theta)R^2]}$.[21] We may also consider the special case in which $\lambda = 0$; in this case

$$\rho_{\Delta E_t}(1) = (1-\theta)R-(1-\delta)\frac{1-((1-\theta)R)^2}{1+(1-\delta)^2-2(1-\delta)(1-\theta)R}, \tag{3.20}$$

which shows how the combination of (θ, δ) affects the first-order autocorrelation of the expenditure on durables.[22] It clearly shows that RI increases the first-order autocorrelation of ΔE_t, i.e., the less the value of θ, the larger $\rho_{\Delta E_t}(1)$. In addition, the higher the depreciation rate, the larger $\rho_{\Delta E_t}(1)$. These results are included in the following proposition.

Proposition 1.

$$\frac{\partial\rho}{\partial\theta} < 0, \tag{3.21}$$
$$\frac{\partial\rho}{\partial\delta} > 0. \tag{3.22}$$

Proof. *Given (3.20), it is straightforward to show that* $\frac{\partial\rho}{\partial\theta} < 0$ *because* θ, $(1-\theta)R \in (0,1)$, *and*

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

■

[21] See Appendix 6.2 for derivations.

[22] Note that when $\theta = 1$, $\rho_{\Delta E_t}(1) = \frac{-(1-\delta)}{1+(1-\delta)^2}$.

Given that $R = 1.01$ and $\delta = 0.05$, $\rho_{\Delta E_t}(1) = 0.5$ when $\theta = 100\%$, $\rho_{\Delta E_t}(1) = -0.25$ when $\theta = 50\%$, and $\rho_{\Delta E_t}(1) = -0.03$ when $\theta = 10\%$. Therefore, when $\theta = 10\%$, the model's prediction on the first-order autocorrelation is close to the empirical counterpart. (0.06 in Mankiw's dataset and -0.04 in more recent dataset $1947-2009$.) For more general cases in which $\lambda > 0$, we cannot obtain the explicit expression for the effect of RI on the first-order autocorrelation, so we present some numerical examples. Figure 4 illustrates how RI affects the autocorrelations when $\lambda = 0$, 0.2, and 0.4, respectively. Clearly, RI does not have monotonic impacts on the first-order correlation due to the presence of the common noise term. However, it does tend to move the correlation closer to zero relative to the standard model. Therefore, RI provides an alternative explanation for the slow adjustment of expenditures on durable goods. For example, when $\theta = 0.1$, $\delta = 0.2$, and $R = 1.03$, the model predicts $\rho_{\Delta E_t}(1) = 0.02$, a value much closer to its empirical counterpart in US data ($\rho_{\Delta E_t}(1) = -0.04$) than that predicted by Mankiw (1982). While the depreciation rate may still seem excessive, it is no longer completely unreasonable (particularly for goods like automobiles).

3.1.3. The Sum of the MA Coefficients for ΔE_t

We also study the dynamic effect of RI on ΔE_t by looking at the autocorrelation over substantially longer periods. Caballero (1990,1994) confirms Mankiw's puzzling finding but also shows that the negative correlation in ΔE_t predicted by the model arises with a delay of several quarters; by looking only at the first-order autocorrelation, Mankiw (1982) missed the reversion observed at longer lags. For example, Table 1 in Caballero (1990) reports that the first MA coefficient, -0.092 (with standard error 0.081), is very far from -0.95 (a value obtained if the quarterly depreciation rate is 0.05), and the sum of all eight MA coefficients is significant and negative, -0.541 (with standard error, 0.234). This feature cannot be captured by the standard PIH model of durable goods, but is predicted by our RI model. Specifically, (3.18) implies that the sum of the MA coefficients for expenditures on durable goods can be written as

$$\begin{aligned}\operatorname{sum}\left(\mathrm{MA}_i\left(\Delta E_t\right)\right) &= \left(\left(1-\theta\right)R-\left(1-\delta\right)\right)\sum_{j=0}^{\infty}\left(\left(1-\theta\right)R\right)^j \\ &= -\frac{\left(1-\delta\right)-\left(1-\theta\right)R}{1-\left(1-\theta\right)R}.\end{aligned}$$

It is clear that if $(1-\delta)-(1-\theta)R > 0$, i.e., $\delta < 1-(1-\theta)R \triangleq \delta_{\max}$, $\operatorname{sum}(MA_i) < 0$.

Proposition 2.

$$\frac{\partial\,\mathrm{sum}\,(\mathrm{MA}_{\,i})}{\partial\theta} = -\frac{\delta R}{(1-(1-\theta)\,R)^2} < 0,$$
$$\frac{\partial\,\mathrm{sum}\,(\mathrm{MA}_{\,i})}{\partial\delta} = \frac{1}{1-(1-\theta)\,R} > 0.$$

Proof. *Straightforward calculation.* ■

Note that under RE ($\theta = 1$), $\mathrm{sum}\,(\mathrm{MA}_{\,i}) = \delta - 1 = -0.95$ if the quarterly depreciation rate is 0.05. Figure 5 illustrates the effects of RI on the sum of the MA coefficients in (3.18). It clearly shows that the degree of inattention increases the sum. For example, if $\theta = 20\%$, $\mathrm{sum}\,(MA_i) = 0.7$, which is closer to its empirical counterpart, -0.541. In addition, in our RI model the sum of the MA coefficients for nondurable goods can be written as,

$$\mathrm{sum}\,(\mathrm{MA}_{\,i}\,(C_t)) = (1-\theta)\,R\sum_{j=0}^{\infty}((1-\theta)\,R)^j = \frac{(1-\theta)\,R}{1-(1-\theta)\,R} > 0,$$

which is also consistent with the evidence reported in Table II in Caballero (1990): the sum of MA coefficients should be positive for nondurables.

3.2. Consumption Dynamics under SE

In this subsection, we present an alternative model of imperfect information diffusion (sticky expectations or "SE") proposed by Carroll (2003) and Carroll and Slacalek (2006), and examine how it affects the dynamics of nondurables and durables consumption. The key assumption of SE is that consumers are inattentive in the sense that every period they update the information about their permanent income with some probability; in other words, only a fraction of them update their information and make optimal adjustments in any period.[23] In the next section, we will present a model in which RI can be used to rationalize SE and the updating probability is determined by the distribution of channel capacity across all consumers in the economy.

As defined in Section 2.1, permanent income that determines optimal consumption decisions, s_t, is

$$s_t = E_t\left[a_t + \frac{1-\delta}{R}k_t + \sum_{j=1}^{\infty}R^{-j}y_{t+j}\right]. \tag{3.23}$$

Denote $\widetilde{k}_{i,t+1}$ the optimal durables stock chosen by a household i who updated expectations

[23] See also Reis (2006).

about permanent income in (the current) period t. Hence, this consumer's actual consumption equals the optimal levels of consumption chosen:

$$c_{i,t} = \widetilde{c}_{i,t} \tag{3.24}$$

$$k_{i,t+1} = \widetilde{k}_{i,t+1}. \tag{3.25}$$

Households who do not update their expectations in period $t+1$ consume

$$c_{i,t+1} = c_{i,t} = \widetilde{c}_{i,t} \tag{3.26}$$

$$k_{i,t+2} = k_{i,t+1} = \widetilde{k}_{i,t+1} \tag{3.27}$$

until they update their expectations; we assume updating happens with the probability π.[24] Nondurables and durables consumption per capita in period t that would prevail if *all* consumers updated their expectations are

$$\Delta\widetilde{C}_t = \int_0^1 \Delta\widetilde{c}_{i,t} di. \tag{3.28}$$

$$\Delta\widetilde{K}_{t+1} = \int_0^1 \Delta\widetilde{k}_{i,t+1} di. \tag{3.29}$$

Because the set of consumers who choose to update is randomly selected from the continuum of agents, the mean consumption of those consumers who choose to update can be written as

$$\Delta C_t^{\pi} = \int_0^1 \pi_{i,t}\Delta\widetilde{c}_{i,t} di = \pi\Delta\widetilde{C}_t \text{ and } \Delta K_{t+1}^{\pi} = \int_0^1 \pi_{i,t}\Delta\widetilde{k}_{i,t+1} di = \pi\Delta\widetilde{K}_{t+1}.$$

If this result holds in every past period, it leads to the following expressions for per capita nondurables and durables consumption[25]

$$\Delta C_t = \pi \sum_{j=0}^{\infty} (1-\pi)^j \, \Delta\widetilde{C}_{t-j}, \tag{3.30}$$

$$\Delta K_{t+1} = \pi \sum_{j=0}^{\infty} (1-\pi)^j \, \Delta\widetilde{K}_{t+1-j}. \tag{3.31}$$

[24] This specification differs from Reis (2006); in his model, consumption can evolve deterministically over the "planning period." Given any distaste for intertemporal variance of consumption combined with perfect capital markets, such deterministic variation would be suboptimal anyway.

[25] See Carroll and Sommer (2003) for the derivation.

Aggregating the change in individual consumption across all consumers,

$$\Delta \widetilde{c}_{i,t} = H_c H_y \varepsilon_t \tag{3.32}$$

$$\Delta \widetilde{k}_{i,t+1} = \frac{1-\beta(1-\delta)}{\beta\varrho} \Delta \widetilde{c}_{i,t} \tag{3.33}$$

we obtain

$$\Delta \widetilde{C}_t = H_c H_y \varepsilon_t \text{ and } \Delta \widetilde{K}_{t+1} = \frac{1-\beta(1-\delta)}{\beta\varrho} \Delta \widetilde{C}_t,$$

which means that

$$\Delta C_t = \pi \sum_{j=0}^{\infty} (1-\pi)^j H_c H_y \varepsilon_{t-j} = \frac{\pi H_c H_y \varepsilon_t}{1-(1-\pi)\cdot L} \tag{3.34}$$

$$\Delta K_{t+1} = \frac{\pi \varsigma_t}{1-(1-\pi)\cdot L} \tag{3.35}$$

$$\Delta E_t = \frac{\pi\left[\varsigma_t - (1-\delta)\varsigma_{t-1}\right]}{1-(1-\pi)\cdot L}, \tag{3.36}$$

where $\varsigma_t = \frac{1-\beta(1-\delta)}{\beta\varrho} H_c H_y \varepsilon_t$.

3.3. Comparison of RI and SE

Comparing (3.34) and (3.35) with (3.9) and (3.10), it is clear that the implications of RI and SE for aggregate nondurables and durables consumption could be similar because aggregating across all individuals would weaken or even eliminate the impacts of the endogenous noises on aggregate consumption growth. Indeed, as discussed in Section 2.2, if all consumers are completely isolated (that is, there is no communication among them and $\overline{\xi}_t = 0$), $\theta = \pi$, and $R = 1$, the two models generate the same dynamics of aggregate nondurables and durables consumption. In other words, they are observationally equivalent in terms of the joint behavior of aggregate consumption and income under some restrictions.[26] Therefore, SE can have similar implications for (1) the relative volatility of aggregate non-durables to durables consumption, (2) the first-order serial correlation of expenditures on aggregate durables consumption, and (3) the sum of the MA coefficients for durable expenditures as RI does, and improve the model in these dimensions.

However, the two hypotheses have different implications for consumption decisions at in-

[26] Luo and Young (2009b) study a range of observational equivalence results in consumption-savings models with RI. Obviously, $R = 1$ is a special case; however, most households in the US save using instruments that pay very low real rates of return (often negative), so it may be a good approximation.

dividual level. Specifically, under RI, consumers adjust optimal consumption plans *frequently but incompletely*, whereas under SE consumers adjust consumption *infrequently but completely* once they choose to adjust.[27] As noted above, in the data consumers do not alter their optimal purchases on durables frequently and the purchases are lumpy. Therefore, SE better captures this feature because consumers are *assumed* to update their expectations on permanent income intermittently (more precisely, with a constant probability). However, the "microfoundation" for SE is more problematic – it relies critically on costs of adjusting expectations that are difficult to measure. RI, in contrast, relies on a well-specified theory of information flow that is used in many contexts and is potentially subject to empirical assessment.[28]

3.4. Sticky Information Diffusion and Adjustment Costs

In the preceding subsection we show that both RI and SE can improve the Mankiw model's predictions on the joint dynamics of nondurables and durables consumption. However, the introduction of sticky information diffusion still cannot explain the observed extreme smoothness of durables consumption. In this subsection we propose a way to introduce sticky information diffusion into Bernanke's adjustment costs model and then show that the interaction of the two assumptions can further improve the model's predictions on aggregate dynamics.

In the Bernanke model proposed in Section 2.2, it is clear that s_t is the unique state variable for nondurables consumption, whereas both k_t and s_t are state variables for the stock of durable goods. As we have discussed in Section 2.1, with more than one state variable, agents with finite capacity can allocate their attention differently across these variables and thus reduce their uncertainty at different rates; consequently, multivariate versions of the RI model are not any more tractable and cannot be solved explicitly. To mitigate this problem, here we assume that consumers can observe their stock of durable goods (k_t), but cannot observe their permanent income (s_t) as determined by the present expected value of financial and human wealth. This assumption is plausible because in reality some components in labor income cannot be observed perfectly and in some situations consumers cannot distinguish different components in labor income, while the information regarding the stock of durable goods can be better observed and perceived.[29] Consequently, permanent income determined by current and expected future labor

[27]Consider an analogy to costs of adjusting capital – RI is analogous to convex costs of adjustment while SE is analogous to fixed costs.

[28]For example, Landauer (1986) estimates the information processing flow in the laboratory. One could adapt this experimental approach to get some information about the flow of signals about economic variables in principle.

[29]See Muth (1960), Pischke (1995), and Hansen and Sargent (2007).

income may easily be more difficult to observe perfectly than the stock of durables.[30]

Therefore, given decision rules of nondurables and durables under RE, (2.15) and (2.16), our hidden state model leads to the following decision rules:

$$c_t = H_c \widehat{s}_t + g_0, \tag{3.37}$$

$$k_{t+1} = x_1 k_t + \frac{x_1 (1 - \beta (1 - \delta))}{\vartheta \left(1 - x_2^{-1}\right)} H_c \widehat{s}_t + h_0, \tag{3.38}$$

where x_1, x_2, and H_c are the same as that defined in Section 2.2, k_0 is given, and $s_0 \sim N\left(\widehat{s}_0, \sigma^2\right)$. Note that given that s_t evolves according to (2.21), the corresponding Kalman filter equation is also governed by (3.5) in Section 3.1. Therefore, the aggregate dynamics is governed by the following equations:[31]

$$\Delta C_t = \frac{\theta H_c \zeta_t}{1 - (1 - \theta) R \cdot L}, \tag{3.39}$$

$$\Delta K_{t+1} = \frac{1}{1 - x_1 \cdot L} \frac{x_1 (1 - \beta (1 - \delta))}{\vartheta \left(1 - x_2^{-1}\right)} \Delta C_t \tag{3.40}$$

and

$$\Delta E_t = \frac{x_1 (1 - \beta (1 - \delta)) \theta H_c}{\vartheta \left(1 - x_2^{-1}\right)} \left(\frac{\zeta_t}{(1 - x_1 \cdot L)(1 - (1 - \theta) R \cdot L)} - \frac{(1 - \delta) \zeta_{t-1}}{(1 - x_1 \cdot L)(1 - (1 - \theta) R \cdot L)} \right) \tag{3.41}$$

$$= \frac{x_1 (1 - \beta (1 - \delta)) \theta H_c}{\vartheta \left(1 - x_2^{-1}\right)} \left(\zeta_t + \sum_{j=1}^{\infty} \Gamma_j \zeta_{t-j} \right),$$

where $\Gamma_j = \sum_{k=0}^{j} \left(x_1^k ((1 - \theta) R)^{j-k} \right) - (1 - \delta) \sum_{k=0}^{j-1} \left(x_1^k ((1 - \theta) R)^{j-1-k} \right)$. The above equations clearly show that both adjustment costs and RI contribute to the propagation mechanism of the model: one is determined by $1 - x_1 \cdot L$ and the other is determined by $1 - (1 - \theta) R \cdot L$. When $\vartheta = 0$ (no adjustment costs), (3.41) reduces to (3.13).

Using expression (3.41), the variability ratio of the change in purchases on durables and

[30]It may be hard for agents to know the *value* of their durables, however. Provided service flow is produced by the stock, not the market value, our assumption about observability seems reasonable.

[31]In this subsection we only consider the case that $\lambda = 0$, as this case is closest to the SE model and can be regarded as the upper bound for the smoothness of durables consumption.

nondurables can be written as

$$\mathrm{rv}\left(\Delta E_t, \Delta C_t\right) = \frac{x_1\left(1-\beta\left(1-\delta\right)\right)}{\vartheta\left(1-x_2^{-1}\right)}\sqrt{\left(1+\sum_{j=1}^{\infty}\Gamma_j^2\right)\left(1-\left((1-\theta)R\right)^2\right)}, \tag{3.42}$$

where we use the fact that $\mathrm{sd}\left[\Delta E_t\right] = \frac{x_1(1-\beta(1-\delta))\theta H_c}{\vartheta\left(1-x_2^{-1}\right)}\sqrt{1+\sum_{j=1}^{\infty}\Gamma_j^2}\omega_\zeta$. Figure 6 illustrates how the combination of adjustment costs and RI affects the variability ratio when $R = 1.01$, $\delta = 0.025$, and $\varrho = 0.0286$. It is clear that incorporating RI into the Bernanke model can further reduce the variability ratio and thus bring the model and the data closer along this dimension. For example, when $\theta = 25\%$ and $\vartheta = 0.3$, the variability ratio is about 0.25, equal to its empirical counterpart in US data.

Similarly, by introducing SE into the Bernanke model, we can obtain similar results on the excess smoothness of durables consumption. Note that aggregate consumption under SE is exactly the same as that under RI when $\lambda = 0$. In this model, both adjustment costs and SE contribute to the propagation mechanism of the model: one is determined by $1 - x_1 \cdot L$ and the other is determined by $1 - \pi \cdot L$.

4. RI vs. SE

In this section, we calculate the welfare losses due to deviating from the first-best instantaneously-adjusted path in both SE and RI economies, and show that RI has a potential to endogenize the probability of re-optimizing in each period.

4.1. Welfare Implications of SE

Consider the SE model proposed in Section 3.2. Here for simplicity we only consider the original Mankiw model (no nondurable goods) in which the utility function $u(k_t) = -\frac{1}{2}\left(\overline{k} - k_t\right)^2$. We further assume that the exogenous probability at which the typical consumer updates his expectations and re-optimizes in any given period is π, independent of the length of time since the optimal plan was set. The consumer sets his optimal consumption plan at t to minimize a quadratic loss function that depends on the difference between the consumer's *actual* consumption plan at period t, k_t, and his first-best instantaneously-adjusted plan k_t^*. If the consumer chooses to adjust at period t, he sets optimal consumption to minimize

$$\frac{1}{2}E_t\left[\sum_{j=t}^{\infty}\beta^{j-t}(k_j - k_j^*)^2\right], \tag{4.1}$$

where $E_t[\cdot]$ is formed using all available information. The following diagram illustrates the evolution of infrequent adjustments over time.[32]

$$\text{If adjust at } t \left\{ \begin{array}{l} \pi\text{: adjust at } t+1 \\ \\ 1-\pi\text{: not adjust at } t+1 \left\{ \begin{array}{l} \pi\text{: adjust at } t+2 \\ \\ 1-\pi\text{: not adjust at } t+2 \left\{ \begin{array}{l} \pi\text{: adjust at } t+3 \\ 1-\pi\text{: not adjust at } t+3 \end{array} \right. \end{array} \right. \end{array} \right. \quad \cdots$$

Therefore, the present discounted welfare losses if the agent *adjusts* at time t (and re-optimizes with the same probability π at $t+1$), v^{SE}, can be written as

$$v^{SE} = E_t \left[\frac{1}{2} \sum_{j=t}^{\infty} ((1-\pi)\beta)^{j-t} (k_t - k_j^*)^2 \right] + \left[\sum_{j=t+1}^{\infty} \left((1-\pi)^{j-t-1} \beta^{j-t} \right) \right] \pi \left(v^{SE} + F \right), \quad (4.2)$$

where F is a fixed cost of adjustment. The first term in the expression measures the welfare losses due to deviations of actual plans from desired (first-best) plans, and the losses are discounted by the discount factor $\left(\beta^{j-t}\right)$ and the probability that k_t will still be set in period j $\left((1-\pi)^{j-t}\right)$; the second term represents the value of adjusting consumption plans at period j ($j > t$) and continuing the procedure. Solving it gives

$$(1-\beta) v^{SE} = \frac{1}{2} [1 - \beta(1-\pi)] E_t \left[\sum_{j=t}^{\infty} ((1-\pi)\beta)^{j-t} (k_t - k_j^*)^2 \right] + \pi\beta F; \quad (4.3)$$

the first order condition with respect to k_t means that

$$\begin{aligned} k_t &= [1 - \beta(1-\pi)] \sum_{j=t}^{\infty} (\beta(1-\pi))^{j-t} E_t \left[k_j^*\right], j \geq t \\ &= \frac{1-\beta(1-\pi)}{1-\beta(1-\pi)} H_k s_t \\ &= H_k s_t = k_t^*, \text{ for any } t \geq 0, \end{aligned}$$

where we use the facts that $k_j^* = H_k s_j$ and $s_{j+1} = s_j + \zeta_{j+1}$. We can therefore calculate that

$$(1-\beta) v^{SE} = \frac{1}{2} \frac{\beta(1-\pi) H_k^2}{1-\beta(1-\pi)} \omega_\zeta^2 + \pi\beta F.$$

[32]Kiley (2000) applied a similar idea in firm's optimizing problem.

which means that

$$v^{SE} = \frac{1}{2}\frac{\beta\left(1-\pi\right)H_k^2\omega_\zeta^2}{\left(1-\beta\right)\left(1-\beta\left(1-\pi\right)\right)} + \frac{\pi\beta}{1-\beta}F. \tag{4.4}$$

In this case, if the agent adjusts in every period there are no welfare losses due to infrequent adjustments, and the total welfare loss the agent suffers is just the present value of fixed costs of adjustment, $\frac{1}{1-\beta}F$.[33] Therefore, we have following proposition:

Proposition 1. *If the fixed cost is small enough, i.e.,*

$$F < \frac{1}{2}\frac{H_k^2\omega_\zeta^2}{1-\beta\left(1-\pi\right)}, \tag{4.5}$$

it is optimal for the consumer to adjust in each period.

Proof. *It is optimal for the consumer to adjust in each period if and only if $v^{SE} + F > \frac{1}{1-\beta}F$. Substituting (4.4) into this inequality yields (4.5).* ■

This conclusion is similar to that obtained in Bar-Ilan and Blinder (1992); they show in an RE case that consumers with full information choose to adjust when the welfare improvements from adjusting are greater than fixed costs induced by adjusting. It is clear that it is always optimal for the consumer to adjust every period if the fixed cost $F = 0$. Furthermore, if we allow for endogenous choice of the probability π, the first-order condition for (4.4) implies that the optimal probability is

$$\pi^* = \frac{H_k\omega_\zeta}{\beta\sqrt{2F}} + \frac{\beta-1}{\beta}, \tag{4.6}$$

which means that the optimal frequency of adjustment is increasing in the volatility of the innovation to permanent income (ω_ζ) and decreasing in the fixed cost F.

4.2. Welfare Implications of RI

Under RI, the agent adjusts optimal consumption plans in every period but the adjustments are incomplete due to finite information-processing capacity. In this case, the welfare losses due to incomplete adjustments are

$$v^{RI} = \min E_t\left[\frac{1}{2}\sum_{j=t}^{\infty}\beta^{j-t}(k_j - k_j^*)^2\right], \tag{4.7}$$

[33] Note that we have imposed the restriction that $\beta R = 1$.

where $E_t[\cdot]$ is formed using processed information and is subject to the information-processing constraint described in Section 3.1, and $k_j^* = H_k s_j$ is the first-best RE (full information) plan. As shown in Sims (2003) and Luo (2008), in this permanent income model optimal consumption plan under RI can be written as

$$k_t = H_k E_t[s_t],$$

where $E_t[s_t]$ is the perceived state variable. Substituting the optimal rule under RI into the objective function, the welfare loss due to incomplete adjustments can be rewritten as

$$\begin{aligned} v^{RI} &= \frac{1}{2}\left[\sum_{j=t}^{\infty}\beta^{s-t}E_t(k_j - H_k E_j[s_j])^2\right] \\ &= \frac{1}{2}\left[\sum_{s=t}^{\infty}\beta^{s-t}E_t\left[E_j(k_j - H_k E_j[s_j])^2\right]\right] \\ &= \frac{1}{2}\frac{H_k^2\sigma^2}{1-\beta}, \end{aligned} \tag{4.8}$$

where $j \geq t$, and $\sigma^2 = \mathrm{var}_j(s_j) = E_j(k_j - H_k E_j[s_j])^2$ is the steady state conditional variance from the Kalman filter. Note that in the RI case the consumer with finite capacity adjusts his optimal plans and suffers from fixed costs in each period; the present value of fixed costs is then $\frac{F}{1-\beta}$.

4.3. How does RI rationalize SE?

To link the SE and RI hypotheses, we consider the situation of a consumer who is given the opportunity to adjust his consumption plan with a probability π, and adjusts in the initial period t. In this situation, the consumer has two options: either (1) increase his capacity to infinity and fully adjust his consumption plans infrequently from $t+1$ on (on average $1/\pi$ period after the last adjustment), or (2) adjust the plans incompletely due to limited capacity in each period. Given expressions (4.4) and (4.8), we have following proposition:

Proposition 2. *If the fixed cost is small enough, i.e.,*

$$F < \frac{1}{2}H_k^2\omega_\zeta^2\Gamma, \tag{4.9}$$

where $\Gamma = \frac{1}{1-\beta(1-\pi)} - \frac{1}{\beta(1-\pi)(1/(1-\theta)-R^2)}$, *it is optimal for the consumer to adjust in each period even if he only has incomplete information about the true state.*

Proof. *It is optimal for the consumer to adjust in each period if and only if*

$$v^{SE} + F > v^{RI} + \frac{F}{1-\beta}. \tag{4.10}$$

Substituting (4.4) and (4.8) into this inequality and using the facts that $\sigma^2 = \frac{\omega_\zeta^2}{\exp(2\kappa)-R^2}$ *and* $\theta = 1 - 1/\exp(2\kappa)$ *.yields (4.9).* ■

Comparing (4.5) with (4.9), we can see that RI introduces additional welfare losses due to incomplete observations and thus reduces the threshold value of the fixed costs above which the consumer chooses not to adjust in each period. It is worth noting that in the RE case it is always optimal to adjust in each period if the fixed cost is zero, whereas it might not be true in the RI case as v^{RI} could be greater than v^{SE} $\left(\text{i.e., } \frac{1}{1/(1-\theta)-1/\beta^2} - \frac{\beta(1-\pi)}{1-\beta(1-\pi)} > 0\right)$. In other words, under RI it might be optimal for the consumer to adjust infrequently even when $F = 0$ as he still suffers from the welfare losses due to incomplete information. It is straightforward to prove the following proposition, which is a sufficient condition for infrequent adjustments in the RI model.

Proposition 3. *If the consumer's channel capacity is small enough, i.e.,*

$$\theta < \widetilde{\theta} = \frac{(1-\pi)\left(1-2\beta^2\right)+\beta}{(1-\pi)\left(1-\beta^2\right)+\beta} \in (0,1), \tag{4.11}$$

it is optimal for the consumer to adjust infrequently even if the fixed cost F is zero.

Furthermore, when $F = 0$, the consumer could be indifferent with the two options, as the welfare losses are the same for the two strategies. Specifically, $v^{SE} = v^{RI}$ $\left(\text{i.e., } \frac{1}{1/(1-\theta)-1/\beta^2} = \frac{\beta(1-\pi)}{1-\beta(1-\pi)}\right)$ implies that the exogenous probability of adjusting in each period (π) can be written as an increasing function of the degree of attention (θ):

$$\pi = 1 - \left(\beta + \frac{\beta}{1-\widetilde{\theta}} - \frac{1}{\beta}\right)^{-1}. \tag{4.12}$$

Note that

$$\frac{d\pi}{d\widetilde{\theta}}|_{v^{SE}=v^{RI}} = \frac{\beta^3}{\left(\beta^2(\theta-2)+1-\theta\right)^2} > 0.$$

Figure 7 clearly shows that the monotonic relationship between π and $\widetilde{\theta}$ for different values of $R\,(=1/\beta)$.[34] The mechanism of the connection of the two hypotheses is as follows. At the

[34]Note that π increases to 1 as θ converges to 100%.

beginning, after the nature chooses a π, (or an optimal π is determined by (4.6),) (4.12) can help determine the corresponding critical capacity $\widetilde{\theta}$ such that the consumers with $\theta < \widetilde{\theta}$ choose adjust infrequently, while those with $\theta > \widetilde{\theta}$ choose to adjust in each period. In the next subsection, we examine the effects of heterogeneity in capacity on the dynamics of durables consumption.

When $F > 0$, the relationship between π and the critical capacity $\widetilde{\theta}$ satisfies the following equality:

$$J\left(\pi, \widetilde{\theta}, F\right) = \left(\frac{1}{1-\beta(1-\pi)} - \frac{2\beta(1-\pi)F}{H_k^2\omega_\zeta^2}\right) - \left(1 + \frac{1}{1/\left(1-\widetilde{\theta}\right) - 1/\beta^2}\right) = 0. \quad (4.13)$$

Applying the implicit function theorem, given π, we have

$$\frac{d\widetilde{\theta}}{dF} = -\frac{\partial G/\partial F}{\partial G/\partial \widetilde{\theta}} > 0,$$

which means that given the probability π, $\widetilde{\theta}$ is increasing in F. Intuitively, with positive fixed costs, the critical capacity would be greater as adjustments in each period are more costly.

4.4. Heterogeneity in Capacity, Infrequent Adjustments, and Aggregate Consumption

Given the relationship between $\widetilde{\theta}$ and π, we can now consider how allowing for heterogeneous channel capacity can generate the smoothness of aggregate consumption while allowing for infrequent adjustments in the microeconomic level. Consider an economy with a continuum of consumers with a measure of 1. We model heterogeneity in capacity by assuming that channel capacity θ is drawn from the distribution $G(\cdot)$. According to Proposition 3 in Section 4.3, given $\widetilde{\theta}$, the fraction of consumers who choose option 1 (the SE option) is $\lambda = G\left(\widetilde{\theta}\right)$, and the remaining consumers, $1-\lambda$, choose option 2 (the RI option). To examine the aggregate implications of heterogeneity in capacity, for simplicity we consider two special distributions of θ: Bernoulli and uniform distributions. We choose these distributions for analytical tractability only.

4.4.1. Case 1: Bernoulli Distribution of θ

Consider the case in which individual capacity follows a Bernoulli distribution. Specifically, we assume that there are only two types of consumers in the economy: type-I consumers have the identical value of attention $\underline{\theta} < \widetilde{\theta}$ and type-II consumers have the identical value of attention $\overline{\theta} > \widetilde{\theta}$. Therefore, in the initial period t, type-I consumers choose the SE option and fully adjust their optimal plan with the probability π, while type-II consumers choose the RI option and do optimization under information-processing constraints every period. We assume that drawing consumers from the population follows a Bernoulli distribution: the probabilities of being type-I and type-II are λ and $1-\lambda$, respectively.

Given the expressions for optimal consumption under SE and RI, the expression for the growth of aggregate consumption is

$$\Delta K_{t+1} = \lambda \Delta K_{t+1}^{SE} + (1-\lambda)\,\Delta K_{t+1}^{RI}, \tag{4.14}$$

where ΔK_{t+1}^{SE} and ΔK_{t+1}^{RI} are consumption per capita if *all* consumers are type-I and type-II, respectively. Specifically,

$$\Delta K_{t+1}^{SE} = \pi \sum_{j=0}^{\infty} (1-\pi)^j \, H_c H_y \varepsilon_{t+1-j},$$
$$\Delta K_{t+1}^{RI} = \overline{\theta} \sum_{j=0}^{\infty} \left(1-(1-\overline{\theta})R\right)^j H_c H_y \varepsilon_{t+1-j}.$$

Therefore,

$$\Delta K_{t+1} = \left[\lambda\pi \sum_{j=0}^{\infty} (1-\pi)^j + (1-\lambda)\,\overline{\theta} \sum_{j=0}^{\infty} \left(1-(1-\overline{\theta})R\right)^j\right] H_c H_y \varepsilon_{t-j}, \tag{4.15}$$

where $\pi = 1-\left(\beta + \frac{\beta}{1-\widetilde{\theta}} - \frac{1}{\beta}\right)^{-1}$. Expression (4.15) shows that the growth of durables consumption is an MA(∞) process with decreasing coefficients, which means that durables consumption adjusts slowly and gradually to income shocks, with reactions that build up over time. Note that these coefficients are functions of the distribution of the two types of consumers (λ), the critical level of capacity $\left(\widetilde{\theta}\right)$, and the upper bound on capacity $\left(\overline{\theta}\right)$. It is clear that the pattern of the response of K_{t+1} in this case is similar to those under the pure RI and SE hypotheses, (3.10) and (3.35).

4.4.2. Case 2: Uniform Distribution of θ

Consider the second case that channel capacity θ follows a uniform distribution along $[0,1]$. At the beginning of every period, every consumer is assigned randomly a value of $\theta \sim U(0,1)$ (and these draws are iid over time and across agents). In this uniform distribution case the fraction of the population whose capacity is below $\widetilde{\theta}$ (we call them type-I consumers) is just $\widetilde{\theta}$. Therefore, the remaining fraction of the population, $1-\widetilde{\theta}$, (we call them type-II consumers) will choose option 2 and do optimization under finite capacity.

If *all* consumers are type-I and type-II, respectively, ΔK_{t+1}^{SE} and ΔK_{t+1}^{RI} can be written as

$$\Delta K_{t+1}^{SE} = \pi \sum_{j=0}^{\infty} (1-\pi)^j H_c H_y \varepsilon_{t+1-j},$$
$$\Delta K_{t+1}^{RI} = \sum_{j=0}^{\infty} \left[\int_{\widetilde{\theta}}^{1} \theta_i \left((1-\theta_i) R \right)^j d\theta_i \right] H_c H_y \varepsilon_{t+1-j}.$$

Therefore, the expression for the growth of aggregate consumption is

$$\begin{aligned} \Delta K_{t+1} &= \widetilde{\theta} \Delta K_{t+1}^{SE} + \left(1-\widetilde{\theta}\right) \Delta K_{t+1}^{RI} \\ &= \left\{ \widetilde{\theta}\pi \sum_{j=0}^{\infty} (1-\pi)^j + \left(1-\widetilde{\theta}\right) \sum_{j=0}^{\infty} \left[\int_{\widetilde{\theta}}^{1} \theta_i \left((1-\theta_i) R \right)^j d\theta_i \right] \right\} H_c H_y \varepsilon_{t+1-j}, \end{aligned} \tag{4.16}$$

where $\pi = 1 - \left(\beta + \frac{\beta}{1-\widetilde{\theta}} - \frac{1}{\beta} \right)^{-1}$. Expression (4.16) also shows that the growth of durables consumption is an MA(∞) process with decreasing coefficients that are functions of the distribution of θ over $[0,1]$ and the critical level of capacity $\left(\widetilde{\theta}\right)$.[35]

5. Conclusion

This paper has examined the implications of two types of sticky information diffusion (finite information-processing capacity or rational inattention and sticky expectations) for the joint dynamics of nondurables and durables consumption at both individual and aggregate levels. In particular, we have shown that the models with sticky information better explain the following aspects in the aggregate data: (1) the relative volatility of aggregate non-durables to durables consumption, (2) the first-order serial correlation of expenditures on aggregate durables consumption, and (3) the sum of the MA coefficients for durable expenditures. In addition, we show that sticky expectations better characterize the observed inertial behavior of durables consumption at the micro level than rational inattention: infrequent and lumpy purchases on durables. Finally, we show that rational inattention endogenizes the exogenous probability as-

[35] We can obtain closed form solutions for the integral terms:

$$\int_{\widetilde{\theta}}^{1} \theta_i \left((1-\theta_i) R \right)^j d\theta_i = \frac{R^j \left(1-\widetilde{\theta}\right)^j}{3j+2+j^2} \left(j\widetilde{\theta} - j\widetilde{\theta}^2 - \widetilde{\theta}^2 + 1 \right).$$

We assume that $\widetilde{\theta}$ is sufficiently large that the sum converges.

sumed in sticky expectations, and the model with heterogeneous capacity can better explain the observed behavior of durables consumption at both micro- and micro-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 microlevel (see Carroll and Samwick 1997). However, solving information-constrained consumer problems in their full nonlinear generality has proven difficult (see Sims 2006, Tutino 2009, or Batchuluun, Luo, and Young 2009 for attempts); whether our results continue to hold when such precautionary considerations are incorporated is an open question. One tractable extension that permits a form of precautionary savings – the risk-sensitive rational inattention model of Luo and Young (2009b) – seems unlikely to alter any conclusions we reached here given the observational equivalence results that hold for that class of models.

6. Appendix

6.1. Deriving the Optimal Decisions

We can formulate the Lagrange function as follows:

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

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

$$\lambda_t = \overline{c}-c_t \tag{6.1}$$

$$\lambda_t = \beta R E_t\left[\lambda_{t+1}\right] \tag{6.2}$$

$$\lambda_t = \beta E_t\left[\varrho\left(\overline{k}-k_{t+1}\right)+\left(1-\delta\right)\lambda_{t+1}\right],\forall t \tag{6.3}$$

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

$$c_t = E_t\left[c_{t+1}\right] \tag{6.4}$$

$$E_t\left[\overline{k}-k_{t+1}\right] = \frac{R+\delta-1}{\varrho}\left(\overline{c}-c_t\right) \tag{6.5}$$

Substituting them into (2.4) and taking the conditional expectation on both sides gives the optimal decisions for nondurables and durables, (2.5) and (2.6), in the text.

6.2. Deriving the Covariance between ΔE_t and ΔE_{t-1}

Using the expression for the change in ΔE_t, (3.11), we have

$$\begin{aligned}
\operatorname{cov}\left[\Delta E_t, \Delta E_{t-1}\right] &= \operatorname{cov}\left[\begin{array}{c} \left(\zeta_t + \frac{((1-\theta)R-(1-\delta))\zeta_{t-1}}{1-(1-\theta)R\cdot L}\right) + \left(\xi_t - (1-\delta+\theta R)\,\xi_{t-1} + \frac{\theta R[(1-\delta)-(1-\theta)R]\xi_{t-2}}{1-(1-\theta)RL}\right), \\ \left(\zeta_{t-1} + \frac{((1-\theta)R-(1-\delta))\zeta_{t-2}}{1-(1-\theta)R\cdot L}\right) + \left(\xi_{t-1} - (1-\delta+\theta R)\,\xi_{t-2} + \frac{\theta R[(1-\delta)-(1-\theta)R]\xi_{t-3}}{1-(1-\theta)RL}\right) \end{array}\right] \\
&= \operatorname{cov}\left[\zeta_t + \frac{((1-\theta)\,R-(1-\delta))\,\zeta_{t-1}}{1-(1-\theta)\,R\cdot L}, \zeta_{t-1} + \frac{((1-\theta)\,R-(1-\delta))\,\zeta_{t-2}}{1-(1-\theta)\,R\cdot L}\right] \\
&+ \operatorname{cov}\left[\begin{array}{c} \xi_t - (1-\delta+\theta R)\,\xi_{t-1} + \frac{\theta R((1-\delta)-(1-\theta)R)\xi_{t-2}}{1-(1-\theta)RL}, \\ \xi_{t-1} - (1-\delta+\theta R)\,\xi_{t-2} + \frac{\theta R((1-\delta)-(1-\theta)R)\xi_{t-3}}{1-(1-\theta)RL} \end{array}\right] \\
&= \left([(1-\theta)\,R-(1-\delta)] + \frac{(1-\theta)\,R\,((1-\theta)\,R-(1-\delta))^2}{1-((1-\theta)R)^2}\right)\omega_\zeta^2 \\
&+ \left[\begin{array}{c} -(1-\delta+\theta R) - \theta R\,(1-\delta+\theta R)\,((1-\delta)-(1-\theta)R) \\ +\theta^2(1-\theta)R^3\,((1-\delta)-(1-\theta)R)^2\,\frac{1}{1-((1-\theta)R)^2} \end{array}\right]\operatorname{var}\left[\xi_t\right],
\end{aligned}$$

which implies (3.19) in the text because $\operatorname{var}[\xi_t] = \frac{1-\theta}{\theta(1-(1-\theta)R^2)}\omega_\zeta^2$.

References

[1] Adda, J. and R. Cooper (2006), "The Dynamics of Car Sales: A Discrete Choice Approach," NBER Working Paper #7785.

[2] Attanasio, O. (2000), "Consumer Durables and Inertial Behaviour: Estimation and Aggregation of (S,s) Rules for Automobile Purchases," *Review of Economic Studies* 67, 659-77.

[3] Bar-Ilan, A. and A.S. Blinder (1992), "Consumer Durables: Evidence on the Optimality of Doing Nothing," *Journal of Money, Credit and Banking* 24, 253-72.

[4] Bernanke, B. (1985), "Adjustment Costs, Durables, and Aggregate Consumption," *Journal of Monetary Economics* 15, 41-68.

[5] Bertola, G. and Caballero, R.J. (1990), "Kinked Adjustment Costs and Aggregate Dynamics," NBER Macroeconomic Annual Vol. 5, pp. 237-288.

[6] Caballero, R.J. (1990), "Expenditure on Durable Goods: A Case for Slow Adjustment," *Quarterly Journal of Economics* 105, 727-743.

[7] Caballero, R.J. (1993), “Durable Goods: An Explanation for Their Slow Adjustment,” *Journal of Political Economy* 101, 351-384.

[8] Caballero, R.J. (1994), “Notes on the Theory and Evidence on Aggregate Purchases of Durable Goods,” *Oxford Review of Economic Policy* 10, 107-117.

[9] Caballero, R.J. and E.M.R.A. Engel (1993), “Microeconomic Rigidities and Aggregate Price Dynamics,” *European Economic Review* 37, 697-711.

[10] Calvo, G.A. (1983), “Staggered Prices in a Utility Maximizing Framework,” *Journal of Monetary Economics* 12, 383-398.

[11] Carroll, C.D. (2003), “Macroeconomic Expectations of Households and Professional Forecasts,” *Quarterly Journal of Economics* 118, 269-298.

[12] Carroll, C.D. and A.A. Samwick (1997), “The Nature of Precautionary Wealth,” *Journal of Monetary Economics* 40, 41-71.

[13] Carroll, C.D. and J. Slacalek (2006), “Sticky Expectations and Consumption Dynamics,” mimeo, Johns Hopkins University.

[14] Carroll, C.D. and M. Sommer (2003), “Epidemiological Expectations and Consumption Dynamics,” mimeo, Johns Hopkins University.

[15] Carroll, C.D., J. Slacalek, and M. Sommer (2008), “International Evidence on Sticky Consumption Growth,” mimeo, Johns Hopkins University.

[16] Coibion, O. and Y. Gorodnichenko (2009), “What Can Survey Forecasts Tell Us About Informational Rigidities?” mimeo, College of William and Mary and UC Berkeley.

[17] Deaton, A. (1992), *Understanding Consumption*, Oxford University Press.

[18] Eberly, J. (1994), “Adjustment of Consumers’ Durables Stocks: Evidence from Automobile Purchases,” *Journal of Political Economy* 102, 403-436.

[19] Galí, J. (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, R.E. (1978), “Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence,” *Journal of Political Economy* 91, 249-265.

[21] Kiley, M. (2000), "Endogenous Price Stickiness and Business Cycle Persistence," *Journal of Money, Credit and Banking* 32, 28-53.

[22] Lam, P. (1991), "Permanent Income, Liquidity, and Adjustments of Automobile Stocks: Evidence From Panel Data," *Quarterly Journal of Economics* 106, 203-230.

[23] Landauer, T.K. (1986), "How Much Do People Remember? Some Estimates of the Amount of Learned Information in Long-Term Memory," *Cognitive Science* 10, 477-493.

[24] Luo, Y. (2008), "Consumption Dynamics under Information Processing Constraints," *Review of Economic Dynamics* 11, 366-385.

[25] Luo, Y. and E.R. Young (2009a), "Rational Inattention and Aggregate Fluctuations," *Berkeley Electronic Journal of Macroeconomics (Contributions)* 9(1), Article 14.

[26] Luo, Y. and E.R. Young (2009b), "Risk-sensitive Consumption and Savings under Rational Inattention," forthcoming, *American Economic Journal: Macroeconomics.*

[27] Maćkowiak, B. and M. Wiederholt (2009), "Business Cycle Dynamics under Rational Inattention," mimeo, European Central Bank and Northwestern University.

[28] Mankiw, G.N. (1982), "Hall's Consumption Hypothesis and Durable Goods," *Journal of Monetary Economics* 10, 417-425.

[29] Martins, G. and D. Sinigaglia (2009), "Real Business Cycle Dynamics under Rational Inattention," mimeo, Columbia University and PUC-Rio.

[30] Reis, R. (2006), "Inattentive Consumers," *Journal of Monetary Economics* 53, 1761-800.

[31] Sims, C.A. (2003), "Implications of Rational Inattention," *Journal of Monetary Economics* 50, 665-690.

[32] Sims, C.A. (2006), "Rational Inattention: Beyond the Linear-Quadratic Case," *American Economic Review* 96, 158-163.

[33] Tutino, A. (2009), "The Rigidity of Choice: Life Cycle Savings with Information-Processing Limits," manuscript, Princeton University.

[34] Woodford, M. (2009), "Information-Constrained State-Dependent Pricing," forthcoming, *Journal of Monetary Economics.*

Figure 6.1: Impulse Response of Δk_{t+1} When $\vartheta = 0.706$ and $\vartheta = 0.0$

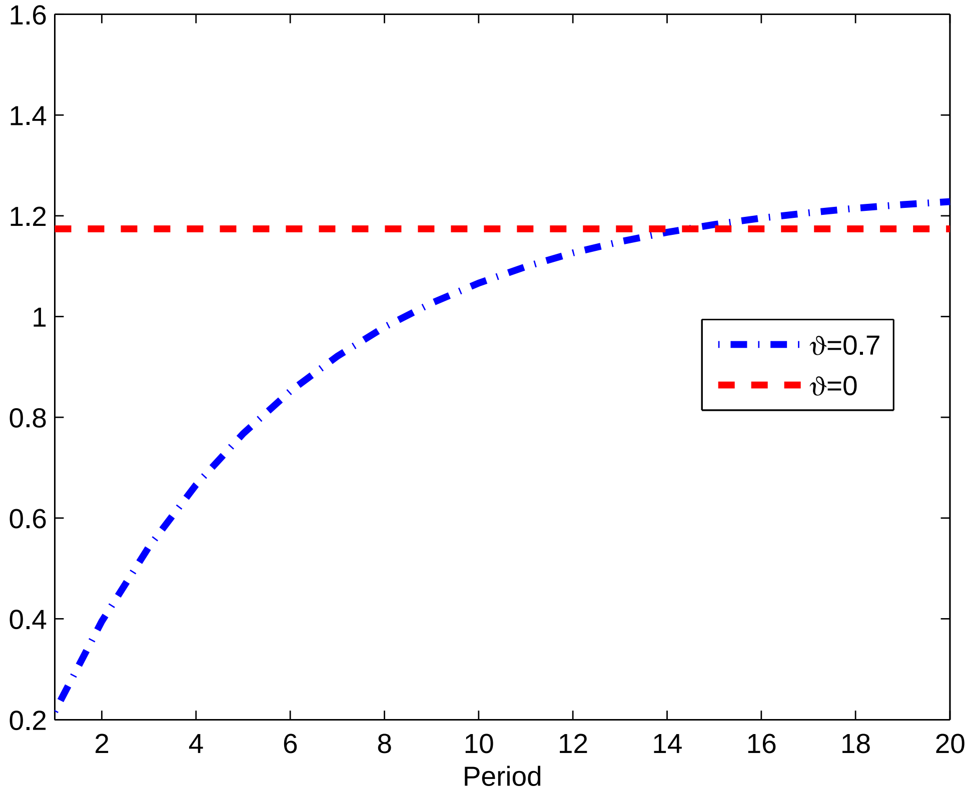

Figure 6.2: The Effects of RI on the Relative Volatility of Durables to Non-Durables

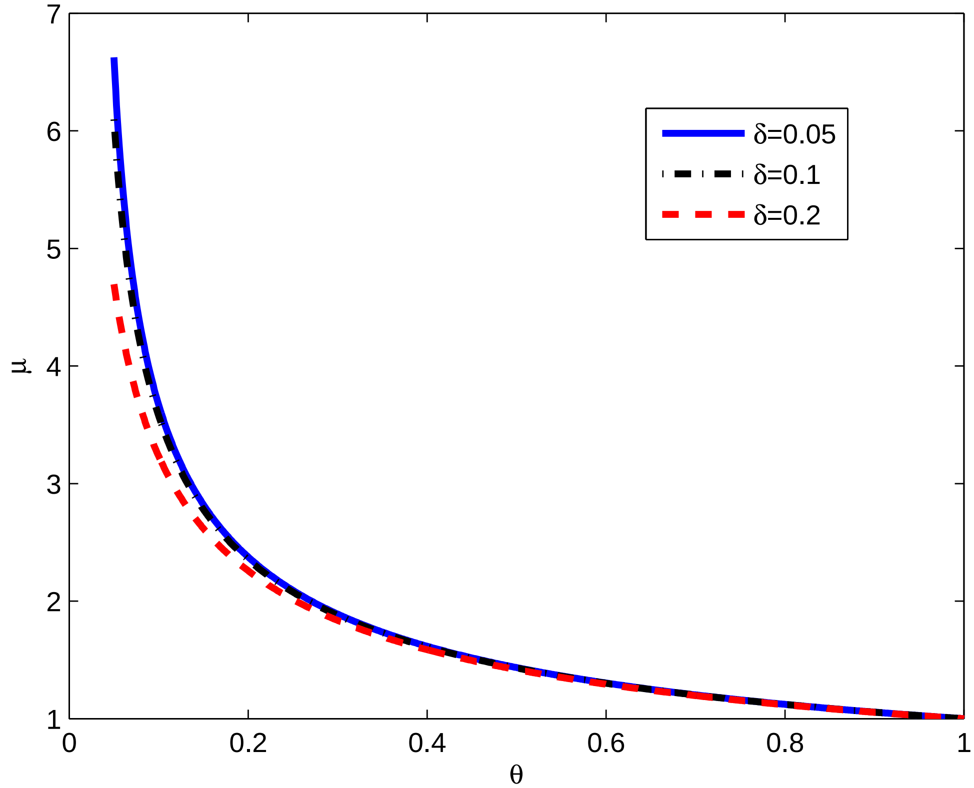

Figure 6.3: The Effects of RI on the Relative Volatility of Durables to Non-Durables

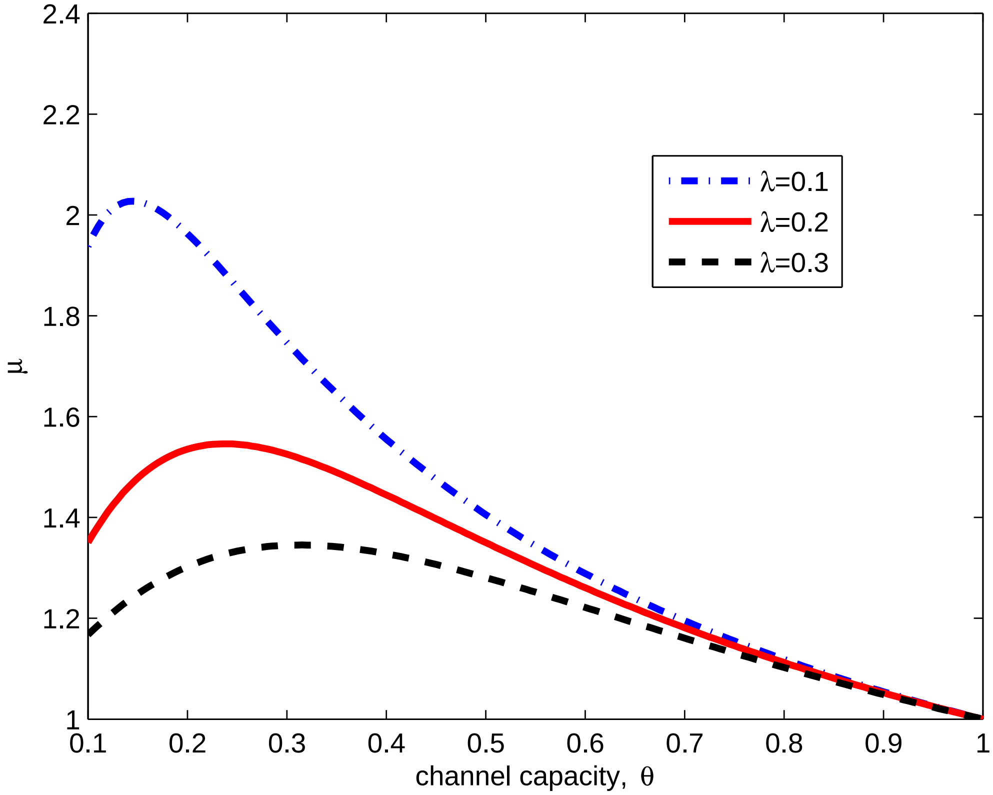

Figure 6.4: The First-order Autocorrelation in Durables

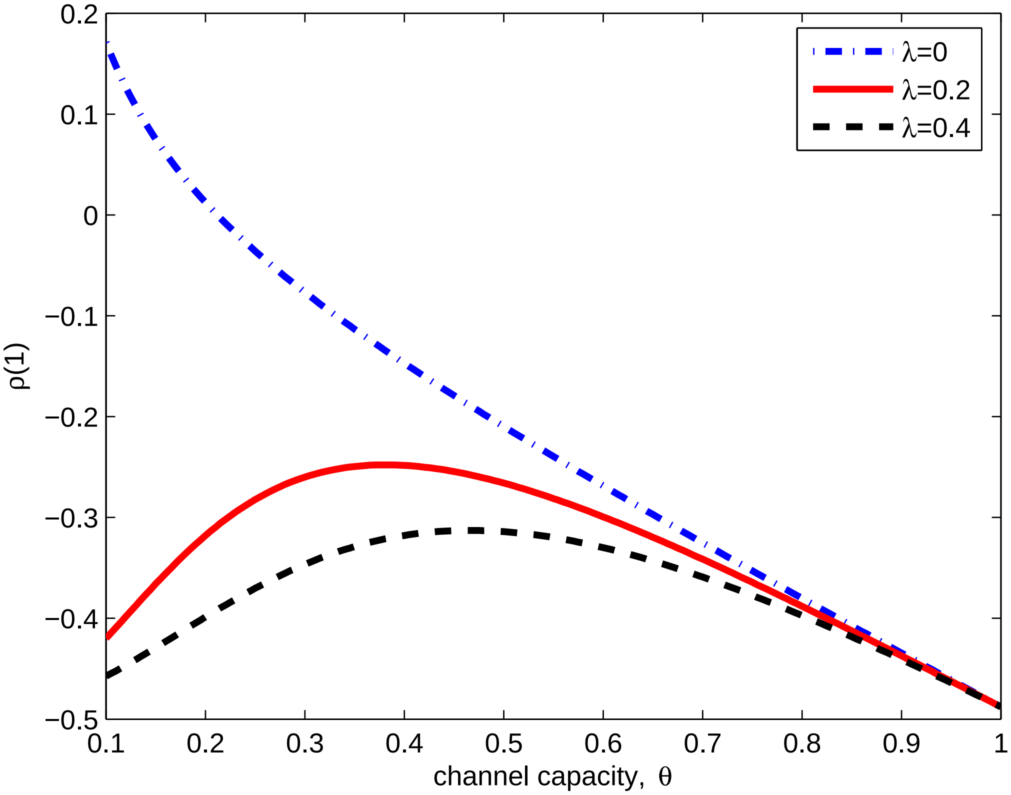

Figure 6.5: The sum of the MA coefficients

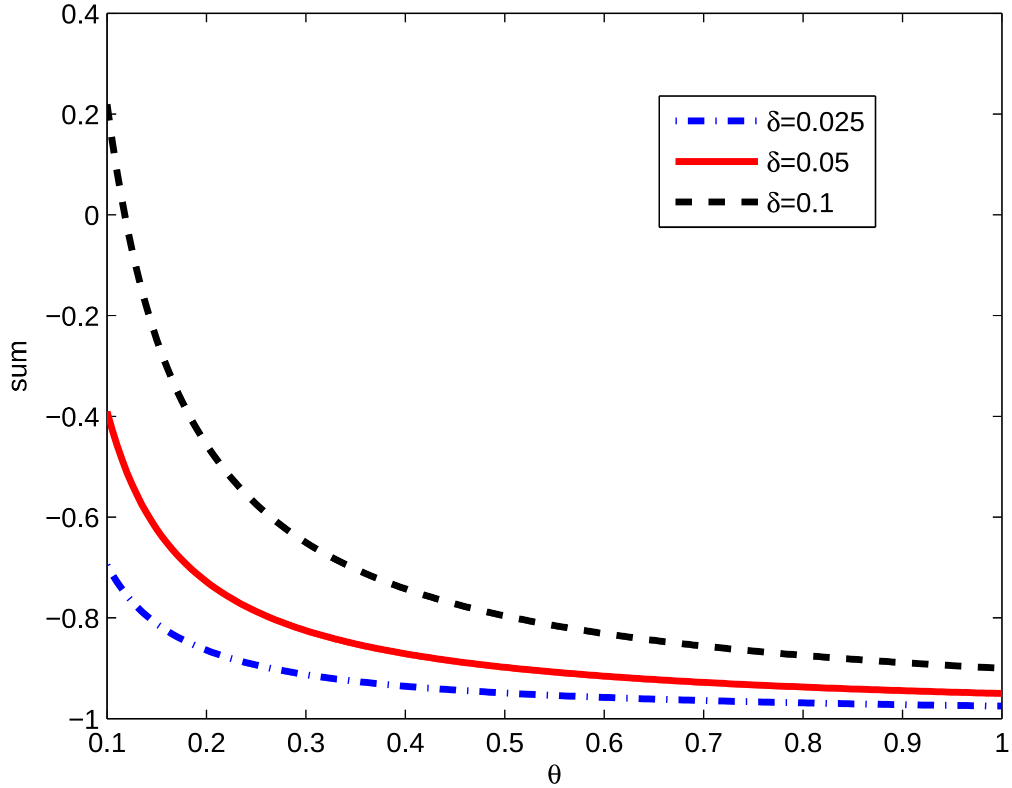

Figure 6.6: The Effect of AC and RI on rv

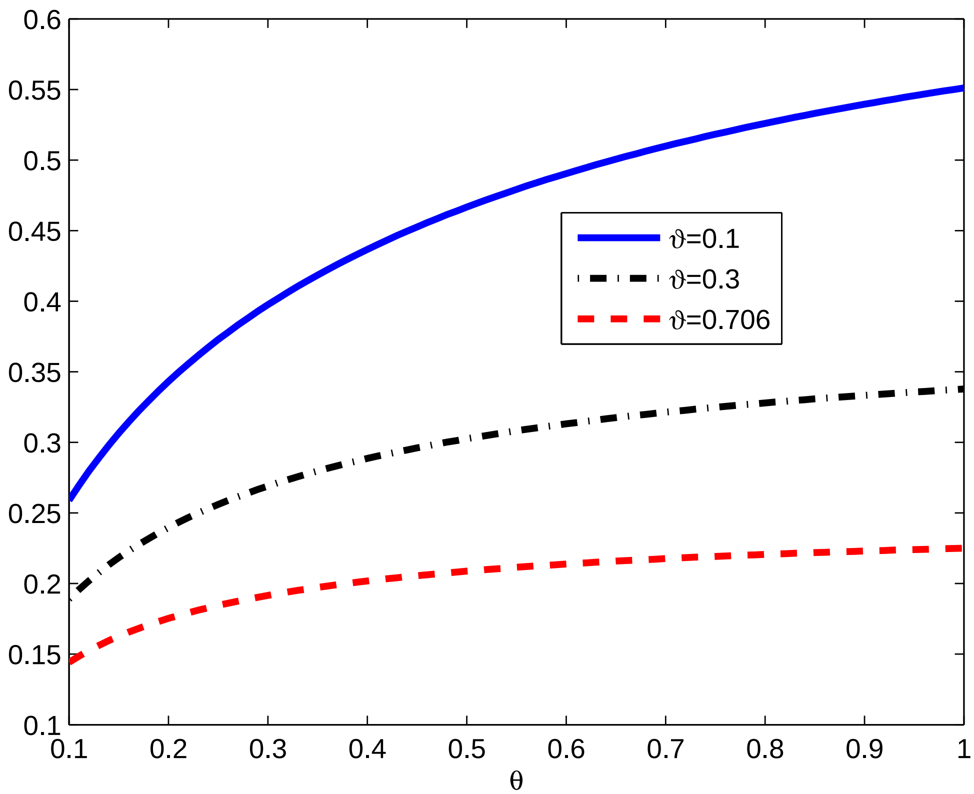

Figure 6.7: The Relationship between θ and π