Elastic Attention, Risk Sharing, and International Comovements*

Wei Li†
Shanghai University of Finance and Economics

Yulei Luo‡
University of Hong Kong

Jun Nie§
Federal Reserve Bank of Kansas City

March 21, 2017

Abstract

In this paper we examine the effects of elastic information-processing capacity (or elastic attention) proposed in Sims (2010) on international consumption and income correlations in a tractable small open economy (SOE) model with exogenous income processes. We find that in the presence of capital mobility in financial markets, elastic attention due to a fixed information-processing cost lowers the international consumption correlations by generating heterogeneous consumption adjustments to income shocks across countries facing different macroeconomic uncertainty. In addition, we show that elastic attention can improve the model's predictions for the other key moments of the joint dynamics of consumption and income.

Keywords: Rational Inattention, Elastic Capacity, Risk Sharing, International Consumption Correlations.

JEL Classification Numbers: D83, E21, F41, G15.

---

*We are grateful to Thomas Lubik (editor) and two anonymous referees for many constructive suggestions and comments. We also would like to thank Anton Braun, Hongyi Chen, Kaiji Chen, Martin Ellison, Charles Engel, Anastasios Karantounias, Federico Mandelman, Chris Otrok, Kang Shi, Chris Sims, Eric R. Young, Vivian Yue, Chi-Wa Yuen, Tao Zha, and seminar participants at University of Hong Kong, Hong Kong Institute of Monetary Research, Federal Reserve Banks of Atlanta and Kansas City, Shanghai Jiao Tong University and Shanghai University of Finance and Economics for helpful suggestions and discussions. Luo thanks the Hong Kong General Research Fund (#HKU791913) for financial support. Part of this work was conducted while Li was visiting the Hong Kong Institute for Monetary Research, whose financial support and hospitality are greatly appreciated. All errors are the responsibility of the authors. 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.

†School of International Business Administration, Shanghai University of Finance and Economics. E-mail: vera.liwei@gmail.com

‡Faculty of Business and Economics, The University of Hong Kong, Hong Kong. E-mail: `yulei.luo@gmail.com`.

§Economic Research Department, Federal Reserve Bank of Kansas City. E-mail: `jun.nie@kc.frb.org`.

1 Introduction

Standard international real business cycle models have difficulty explaining some stylized facts in open economies. One of the major inconsistencies between the models' predictions and the empirical evidence concerns cross-county consumption and income correlations. Specifically, in a canonical open economy model under the complete market assumption, risk averse consumers will insure country-specific risk using international financial markets, which leads to highly, even fully correlated consumption regardless of output (or income) correlations.[1] However, the empirical evidence suggests that cross-country consumption is far from perfectly correlated; in fact they are lower than output (or income) correlation in most cases.[2] Backus, Kehoe and Kydland (1992) call this inconsistency the "most striking discrepancy" between data and theory (Henceforth, the BKK puzzle).

In this paper, we propose a novel explanation for the BKK puzzle by incorporating elastic attention, which is optimally chosen by agents in response to fundamental shocks, into an otherwise standard small open economy (SOE) model. Specifically, we follow Sims (2010) and assume that individual consumers face fixed information-processing costs and thus have only limited and elastic information-processing capacity when making major economic decisions. Consequently, consumers cannot perfectly observe the state of the economy and learn the true state using noisy observations. Furthermore, they optimally choose information-processing capacity and make decisions based on perceived information.[3] Many empirical studies found that incomplete information about the state plays an important role in affecting individual agents' optimal decisions. For example, Coibion and Gorodnichenko (2012, 2015) find pervasive evidence consistent with Sims' rational inattention model using the U.S. surveys of professional forecasters, ordinary consumers and investors.[4] In particular, Coibion and Gorodnichenko (2015) find that information rigidities were falling from the late 1960s to the early 1980s as the volatility of macroeconomic variables was rising, while had been consistently increasing since the start of the Great Moderation ($1983-1984$). They then argue that one should be careful when treating information rigidities at the macro level as a structural parameter because these rigidities vary over time in response to changes in macroeconomic conditions.[5]

[1]See Chapter 6 in Obstfeld and Rogoff (1996) for a textbook treatment on this topic.

[2]Table 1 of this paper reports the cross-country consumption and income correlations using the G-7 data.

[3]The assumption is also consistent with a psychology theory on elastic attention proposed in Kahneman (1973). It is worth noting that Sims (2003) first proposed the rational inattention hypothesis by assuming that economic agents only have limited and fixed capacity/attention, and did not study the elastic attention case.

[4]Hong, Torous, and Valkanov (2007) find supportive evidence for rational inattention in the financial markets. Specifically, they find that investors in the stock market react gradually to information contained in industry returns about their fundamentals and that information diffuses only gradually across markets.

[5]See Maćkowiak and Wiederholt (2015), Matejka and McKay (2015), Cheremukhin and Tutino (2016), and Luo and Young (2016) for applications of elastic attention in other macroeconomic settings.

After incorporating this elastic attention idea, we first solve the model in closed-form solution and explicitly show that the elastic attention mechanism can endogenously generate heterogenous and gradual consumption adjustments to income shocks across countries, and thus make the model better explain international consumption and income correlations as well as some other key stochastic properties of the joint dynamics of consumption and income in individual countries.

Our closed-form solution can help analytically inspect the key mechanism behind the results. We show there are three competing channels that interact and determine the consumption correlation in our benchmark model. The first channel is *the slow adjustment channel*. Specifically, if the home country and the rest of the world in our model economy have the same degree of slow adjustment, imperfect state observations generate a same pattern of gradual responses of consumption growth to income shocks, and the channel has no impact on the cross-country correlation because its impacts on consumption variance and cross-country consumption covariance are just cancelled out.[6] The second channel is *the common noise channel*. The common noise arising from imperfect observations is partially common within countries and is independent across countries. After aggregating over individual consumers, the common noise reduces consumption correlations across countries because it increases consumption volatility while having no effect on the covariance of consumption across countries. This channel can help distinguish RI from other models such as the habit formation model and the model with incomplete information about income which cannot reduce the consumption correlation. The third channel is *the elastic attention channel*. Individual consumers facing fixed information-processing costs optimally choose their information-processing capacity (i.e., the degree of elastic attention) that affects the speed of consumption adjustment via interacting with the fundamental uncertainty. The gap between heterogenous responses of consumption to income due to the elastic attention further reduces the cross-country consumption correlation.

It is worth noting that the elastic attention channel is different from that obtained in the rational inattention model with fixed capacity (e.g., Sims 2003 and Luo 2008). Specifically, when the marginal cost of processing information is fixed while the optimal information-processing capacity can be adjusted in response to fundamental shocks, both the variance of noise and the speed of adjustment depend on the amount of fundamental uncertainty, which differs across countries by nature. This endogenous variation in the optimal information-processing capacity is the key underlying mechanism that generates *greater* heterogeneity and thus lowers cross-country consumption correlations. In addition, to separate the elastic attention channel and the common noise channel, we calibrate the key parameter on the common noise channel using three different national media

[6] As will be discussed in Section 3.3, some other hypotheses such as habit formation can generate the same effect on the consumption correlation. Note that here we just use the basic formula for the correlation between two variables X and Y: $\text{cov}(X,Y)/(\text{sd}(X)\,\text{sd}(Y))$, where $\text{cov}(X,Y)$ is the covariance between X and Y, and $\text{sd}(X)$ and $\text{sd}(Y)$ are standard deviations of X and Y, respectively.

concentration measures. By fixing the common noise parameter at the calibrated values, our quantitative results show that the model with elastic attention does a reasonably good job in matching the empirical counterparts of consumption correlations across countries.

To the best of our knowledge, most of the previous efforts to solve the BKK puzzle assume that consumers have infinite information-processing ability. In contrast, as shown in Sims (2003, 2010), the rational inattention hypothesis can provide a micro-foundation for modeling stickiness, randomness and delays observed in economic behavior.[7] This paper considers an important application of RI and shows that small deviations from the standard full-information rational expectations due to limited attention can significantly improve the model's predictions on international comovements as well as the joint dynamics of consumption and income. Our paper is closely related to Boz, Daude, and Durdu (2011) and Luo, Nie, and Young (2014). Boz, Daude, and Durdu (2011) incorporate incomplete information about the trend and cyclical shocks and learning into a SOE-RBC model and examine how it affects business cycle dynamics in emerging markets. Luo, Nie, and Young (2014) examine the effects of model uncertainty due to a concern about model specification (robustness) on international consumption correlations in a SOE real business cycles (RBC) model. By contrast, this paper tackles the problem from a different angle where agents trust their models, but are unable to process all required information due to limited information capacity when making decisions. This paper is also related to the SOE model with habit formation proposed in Fuhrer and Klein (2006). They find that with habit formation, a common interest rate shock can generate a positive consumption correlation even when no risk sharing exists. Therefore, risk diversification is even less prevalent than standard empirical tests suggest, worsening the puzzle. This mechanism is similar to the slow adjustment channel discussed above.

Literature review. There are other theories proposed in the literature to tackle the BKK puzzle. For example, Devereux, Gregory, and Smith (1992) apply nonseparable utility to generate more realistic cross-country consumption correlations but fail to make them consistently lower than output correlations. Stockman and Tesar (1995) show that the presence of nontraded goods in the complete-market model can improve, though not resolve, the problem. Engel and Wang (2011) show in their two-sector two-country model that introducing durable goods can better explain the observed behavior of trade and the consumption correlation when the technology innovations in both durable and nondurable goods sectors are highly correlated. Colacito, Croce, Ho, and Howard (2014) document a new anomaly that a canonical international RBC model cannot explain, i.e., capital outflow in response to a positive long-run productivity shock. Although they find that introducing the Epstein-Zin recursive preference can help resolve this puzzle and improves the model's predictions of the high equity premium and the volatility of the exchange rate, the consumption

[7] See Luo (2008) and Maćkowiak and Wiederholt (2009) for the applications of RI in the consumption and firm decisions within the linear-quadratic-Gaussian setting.

correlation in their model is slightly higher than the output correlation. Chen and Crucini (2014) suggest that the presence of temporary productivity spillover from the foreign economy can account for the low consumption correlation. Another important theory in the literature is the presence of demand shocks. (See, for example, Wen 2007 and Bai and Ríos-Rull 2015.) It is worth noting that the presence of demand shocks has the potential to reduce the consumption correlation to a realistic level at the cost of excessive volatility of consumption relative to output. This mechanism is similar to the common noise channel of our elastic attention model. In addition, some efforts have also been devoted to examining how financial market imperfections affect international comovements. (See, for examples, Kollman 1996, Lewis 1996, Kehoe and Perri 2002, and Bai and Zhang 2012.)

The remainder of the paper is organized as follows. Section 2 presents the standard full-information rational expectations SOE model and discusses the model's puzzling implications for international consumption correlations. Section 3 introduces RI into this SOE model and examines the theoretical implications of elastic attention. Section 4 presents the quantitative findings about how elastic attention improves the model's performance on the cross-country consumption correlation and other key stochastic properties of the joint dynamics of consumption and income. Section 5 concludes the discussion.

2 Benchmark: Full-information Rational Expectations Small Open Economy Model

2.1 Model Setup

In this section we present a full-information rational expectations (FI-RE) version of a small open economy (SOE) model and will discuss how to incorporate rational inattention (RI) into this stylized model in the next section. Following the incomplete financial market literature, we consider an economy with a continuum of ex ante identical consumers and assume that the only asset that is traded internationally is a risk-free bond. Following Glick and Rogoff (1995) and Obstfeld and Rogoff (1996), we formulate the FI-RE SOE model as

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

subject to the flow budget constraint

$$b_{t+1} = Rb_t - c_t + y_t, \tag{2}$$

where $u(c_t) = -\left(\bar{c} - c_t\right)^2/2$ is the utility function,[8] c_t is consumption, $\bar{c}$ is the bliss point, $R \geq 1$ is the exogenous and constant gross world interest rate, b_t is the amount of the risk-free world bond

[8] Our main results do not change if we adopt the constant-absolute-risk-aversion (CARA) utility function, $u\left(c_t\right) = -\exp\left(-\alpha c_t\right)/\alpha$, where $\alpha > 0$ is the coefficient of absolute risk aversion. The main reason for this result is that the

held at the beginning of period t, y_t is real income in period t, and $E_0[\cdot]$ is the typical consumer's expectation operator conditional on his processed information at time 0.[9] Here we assume that the household sector takes y_t as given to keep our model tractable.[10] The model assumes perfect capital mobility in that domestic consumers have access to the bond offered by the rest of the world and that the real return on this bond is the same across countries. In other words, the world risk-free bond provides a mechanism for domestic households to smooth consumption using the international capital market. Finally we assume that the no-Ponzi-scheme condition is satisfied.

A similar problem can be formulated for the rest of the world (ROW). We use an asterisk ("$*$") to represent the rest of the world variables. For example, we assume that y_t^* is the aggregate income of the rest of the world (G7, OECD, or EU). Furthermore, we assume that the domestic endowment and the ROW endowment are correlated.

Let $\beta R = 1$; optimal consumption is then determined by permanent income:

$$c_t = (R-1)\, s_t, \tag{3}$$

where

$$s_t = b_t + \frac{1}{R}\sum_{j=0}^{\infty} R^{-j} E_t\left[y_{t+j}\right]$$

is the expected discounted present value of lifetime total wealth, consisting of financial wealth (the risk free foreign bond) plus human wealth. From (3), we can see that uncertainty does not explicitly appear in the consumption function and thus the certainty equivalence principle holds.

In order to facilitate the introduction of the rational inattention hypothesis, we follow Luo (2008) and Luo, Nie, and Young (2015), and reduce the multivariate model with a general income process to a univariate model with iid innovations to total wealth s_t that can be solved analytically.[11] Letting s_t be defined as a new state variable, we can reformulate the SOE model as

$$v(s_0) = \max_{\{c_t\}_{t=0}^{\infty}} \left\{ E_0 \left[\sum_{t=0}^{\infty} \beta^t u(c_t) \right] \right\} \tag{4}$$

subject to

$$s_{t+1} = R s_t - c_t + \zeta_{t+1}, \tag{5}$$

stochastic property of the joint dynamics of consumption and income is mainly determined by the marginal propensity to consume (MPC) out of expected total wealth, and the CARA and LQ specifications have the same MPC. Of course, under imperfect state observations we need to assume that the loss function due to imperfect observations is still approximately quadratic even if the utility function is CARA.

[9] Here we ignore the investment and government spending components.

[10] Incorporating the firm sector and modelling investment decisions explicitly do not change the main results in this paper. (See the online appendix for an extension to an RI-SOE model with endogeneous capital accumulation.)

[11] See Luo (2008) for a formal proof of this reduction. Multivariate versions of the RI model are numerically, but not analytically, tractable, as the variance-covariance matrix of the states cannot generally be obtained in closed form.

where the time $(t+1)$ innovation to total wealth can be written as

$$\zeta_{t+1} = \frac{1}{R} \sum_{j=t+1}^{\infty} \left(\frac{1}{R} \right)^{j-(t+1)} \left(E_{t+1} - E_t \right) [y_j] \, ; \tag{6}$$

$v(s_0)$ is the consumer's value function under FI-RE. Under the FI-RE hypothesis, this model with quadratic utility leads to the well-known random walk result of Hall (1978):

$$\begin{aligned} \Delta c_t &= \frac{R-1}{R} \left(E_t - E_{t-1} \right) \left[\sum_{j=0}^{\infty} \left(\frac{1}{R} \right)^j \left(y_{t+j} \right) \right] \qquad (7) \\ &= (R-1) \, \zeta_t, \end{aligned}$$

which relates changes in consumption to income innovations.[12] In this case, the change in consumption depends neither on the past history of labor income nor on anticipated changes in labor income.

As argued in Hansen (1987) and Cochrane (Chapter 2, 2005), the above endowment economy model can be regarded as a general equilibrium model with a linear production technology. Specifically, in our model setting, R can be regarded as the return on technology and is not yet the interest rate (the equilibrium rate of return on one-period claims to consumption). As proposed in Cochrane (2005), we first find optimal consumption in the model and then price one-period claims using the equilibrium consumption stream. Denoting the risk free rate by R^f, we have the following Euler equation for the home country:

$$\frac{1}{R^f} = E_t \left[\beta \frac{u'(c_{t+1})}{u'(c_t)} \right] = \beta E_t \left[\frac{\overline{c} - c_{t+1}}{\overline{c} - c_t} \right] = \beta = \frac{1}{R},$$

where $E_t[\cdot]$ is the consumer's expectation operator conditional on his processed information at time t. That is, the equilibrium interest rate, R^f, is the same as the *exogenously* specified return on the technology, R. Since the rest of the world has the same value of the discount factor β and faces the similar Euler equation, it is straightforward to show that both the home country and the rest of world face the same equilibrium interest rate.

As mentioned before, we adopt the small-open economy model with a constant interest rate and quadratic utility rather than a two-country general equilibrium model with the constant-relative-risk-averse (CRRA) utility (e.g., Kollman 1996) for two reasons. First, most existing RI models assume that the objective functions are quadratic and the state transition equations are linear; consequently, the Gaussian ex post distribution of the state is optimal, which greatly simplifies the model. As shown in Sims (2006), fully solving non-LQG models is extremely difficult; the models

[12] Under FI-RE the expression for the change in individual consumption is the same as that for the change in aggregate consumption.

solved in those papers have either very short horizons or extremely simple setups due to numerical obstacles – the state of the world is the distribution of the state and this distribution is not well-behaved.[13] Second, there is no two-country general equilibrium in which the general equilibrium interest rate is constant.[14] (See Section 3.1 in Luo, Nie, and Young 2014 for a detailed discussion on this issue.)

2.2 Estimating Income Processes

We follow the intertemporal consumption literature (Quah, 1990; Pischke, 1995; Luo, Nie and Young, 2015) and assume that the income process in the home country, y_t, consists of two components, a random walk and a white noise:

$$y_t = y_t^p + y_t^i, \tag{8}$$

$$y_t^p = y_{t-1}^p + \varepsilon_t, \tag{9}$$

$$y_t^i = \overline{y} + \epsilon_t, \tag{10}$$

where y_t^p and y_t^i are the permanent and transitory income components, respectively, and ε_t and ϵ_t are orthogonal iid shocks with mean 0 and variance ω^2 and ω_ϵ^2, respectively.

For the rest of the world (ROW), income is assumed to have similar processes:

$$y_t^* = y_t^{p*} + y_t^{i*}, \tag{11}$$

$$y_t^{p*} = y_{t-1}^{p*} + \varepsilon_t^*, \tag{12}$$

$$y_t^{i*} = \overline{y}^* + \epsilon_t^*, \tag{13}$$

where ε_t^* and ϵ_t^* are orthogonal iid shocks with with mean 0 and variance ω^{*2} and ω_ϵ^{*2}, respectively. In addition, we allow contemporaneous correlations between the SOE and the ROW,

$$\text{corr}(\varepsilon, \varepsilon^*) = \eta > 0 \text{ and } \text{corr}(\epsilon, \epsilon^*) = \rho > 0.$$

Since $\Delta y_t = \varepsilon_t + \epsilon_t - \epsilon_{t-1}$ and $\Delta y_t^* = \varepsilon_t^* + \epsilon_t^* - \epsilon_{t-1}^*$, the international correlation of income growth can be written as:

$$\text{corr}(\Delta y_t, \Delta y_t^*) = \frac{E[\varepsilon_t \varepsilon_t^*] + 2E[\epsilon_t \epsilon_t^*]}{\sqrt{\omega^2 + 2\omega_\epsilon^2}\sqrt{\omega^{*2} + 2\omega_\epsilon^{*2}}}. \tag{14}$$

Using annual GDP data from 1950 – 2010 from Penn World Tables (version 7.1), we find that income volatility is mainly determined by the permanent shocks for G-7 countries. For example,

[13] Note that within the LQ setting, the first two moments are sufficient to characterize the distribution of the state.

[14] Note that in our model setting, the resulting stochastic interest rate may make our RI model intractable because it is no longer a LQG specification and thus the first two moments are not sufficient to characterize the entire distribution of the true state.

for the U.S., ω/ω_ϵ is about 75000, suggesting $\omega \gg \omega_\epsilon$. This empirical result is consistent with the existing studies on estimating the two-component income process. For example, Quah (1990) proposes this specification and argues that it has the potential to solve the excess smoothness puzzle in the consumption literature. He estimates that the transitory income component accounts for 1% to 2% of total variance of consumption. Luo, Nie, and Young (2015) find that the ω_ϵ/ω ratio is 0.0061 using quarterly US data over the period of $1955-2012$. The relative variance of the transitory shock to the permanent shock remains small for other G-7 countries and different versions of PWT real GDP data.[15] Therefore, we will focus on the permanent income component for the rest of our discussion. The cross country income correlation can be approximated as follows:

$$\operatorname{corr}(\Delta y_t, \Delta y_t^*) \approx \frac{E[\varepsilon_t \varepsilon_t^*]}{\omega\omega^*} = \operatorname{corr}(\varepsilon_{t+1}, \varepsilon_{t+1}^*).$$

Under the income specification, the innovation to life-time wealth can be reduced to $\zeta_t = \varepsilon_t/\left(R-1\right) + \epsilon_t/R \sim N(0, \omega_\zeta^2)$ which is approximately equal to $\varepsilon_t/\left(R-1\right)$ with variance $\omega^2/(R-1)^2$.

2.3 Implications for Cross-Country Consumption Correlations

In the FI-RE model proposed in Section 2.1, consumption growth can be written as

$$\Delta c_t = \left(R-1\right)\zeta_t,$$

which means that consumption growth is white noise and the impulse response of consumption to the income shock is *flat* with an immediate upward jump in the initial period that persists indefinitely. It is worth noting that this consumption behavior does not fit the data well. As has been well documented in the consumption literature (e.g., Reis 2006), the impulse response of aggregate consumption to aggregate income takes a hump-shaped form, which means that aggregate consumption growth reacts to income shocks gradually. In Sections 3 and 4, we will show that introducing elastic attention can help generate more realistic impulse responses of consumption to income.

Given that consumption dynamics in the ROW is

$$\Delta c_t^* = \left(R-1\right)\zeta_t^*,$$

the international consumption correlation can thus be written as

$$\operatorname{corr}(\Delta c_t, \Delta c_t^*) \approx \operatorname{corr}(\varepsilon_t, \varepsilon_t^*) \approx \operatorname{corr}(\Delta y_t, \Delta y_t^*). \tag{15}$$

[15] For the other G7 countries, the transitory-permanent variance ratios are 0.00009, 0.0148, 0.00013, 0.0019, 0.00051, and 0.0340 for Canada, France, UK, Italy, Japan and Germany, respectively. The ratios remain very low when we adopt different specifications for real income: (1) real GDP minus government spending, (2) real GDP minus investment, and (3) real GDP minus investment and government spending.

Note that this prediction would not be consistent with the empirical evidence, as international consumption correlations are lower than output correlations for most pairs of countries. (See Table 1 for the consumption-income correlations in the four relatively small countries in the G7.)

3 Theoretical Implications of RI for Consumption-Income Comovements

3.1 Introducing RI with Elastic Attention

Following Sims (2003, 2010), we incorporate rational inattention (RI) due to finite information-processing capacity into the FI-RE SOE model specified above. Under RI, agents have only finite Shannon channel capacity to observe the state of the world. Specifically, we use the concept of entropy from information theory to characterize the uncertainty about a random variable; the reduction in entropy is thus a natural measure of information flow.[16] With finite capacity $\kappa \in (0,\infty)$, the state variable s following a continuous distribution cannot be observed without error and thus the information set at time $t+1$, $\mathcal{I}_{t+1}$, is generated by the entire history of noisy signals $\left\{s_j^*\right\}_{j=0}^{t+1}$. Agents with finite capacity will choose a new signal $s_{t+1}^* \in \mathcal{I}_{t+1} = \left\{s_1^*, s_2^*, \cdots, s_{t+1}^*\right\}$ that reduces their uncertainty about the state variable s_{t+1} as much as possible. Formally, this idea can be described by the information constraint

$$\mathcal{H}\left(s_{t+1}|\mathcal{I}_t\right) - \mathcal{H}\left(s_{t+1}|\mathcal{I}_{t+1}\right) \leq \kappa, \tag{16}$$

where κ is the typical consumer's information 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. Finally, following the literature, we suppose that the prior distribution of s_{t+1} is Gaussian.

Under the linear-quadratic-Gaussian (LQG) setting, as has been shown in Sims (2003, 2010), the ex post Gaussian distribution, $s_t|\mathcal{I}_t \sim N\left(E\left[s_t|\mathcal{I}_t\right], \Sigma_t\right)$, where $\Sigma_t = E_t\left[\left(s_t - \widehat{s}_t\right)^2\right]$, is optimal.[17] In addition, Maćkowiak and Wiederholt (2009) also show in a tracking problem that when the variables being tracked follow a stationary Gaussian process, signals which take the form of "true state plus white noise error" (i.e., $s_{t+1}^* = s_{t+1} + \xi_{t+1}$, where ξ_{t+1} is the iid endogenous noise due to RI) are optimal. Although the SOE model proposed in this paper is not a pure tracking problem,

[16]Formally, entropy is defined as the expectation of the negative of the (natural) log of the density function, $-E\left[\ln\left(f\left(s\right)\right)\right]$. The entropy of a discrete distribution with equal weight on two points is simply $E\left[\ln\left(f\left(s\right)\right)\right] = -0.5\ln\left(0.5\right) - 0.5\ln\left(0.5\right) = 0.69$, and the unit of information contained in this distribution is 0.69 "nats". In this case, an agent can remove all uncertainty about s if the capacity devoted to monitoring s is $\kappa = 0.69$ nats.

[17]This result is often assumed as a matter of convenience in signal extraction models with exogenous noises, and RI can rationalize this assumption.

the proof for the optimal form of the noisy signal in Maćkowiak and Wiederholt ((2009) can be extended to our model in which the state being tracked is a random walk if the channel capacity, κ, is greater than a lower bound so that all conditional moments are finite and well-defined.[18] Specifically, within the LQG setting, the information-processing constraint, (16), can be reduced to

$$\ln\left(R^2\Sigma_t+\omega_\zeta^2\right)-\ln\left(\Sigma_{t+1}\right)\leq 2\kappa; \tag{17}$$

Since this constraint is always binding, we can compute the value of the steady state conditional variance Σ: $\Sigma=\omega_\zeta^2/\left(\exp\left(2\kappa\right)-R^2\right)$. Given this expression for Σ and that the noisy signal takes the following additive form:

$$s_t^*=s_t+\xi_t, \tag{18}$$

where $\xi_t \sim N\left(0,\Lambda\right)$, we can use the usual formula for updating the conditional variance of a Gaussian distribution Σ to recover the variance of the endogenous noise (Λ): $\Lambda=\left(\Sigma^{-1}-\Psi^{-1}\right)^{-1}$, where $\Psi=R^2\Sigma+\omega_\zeta^2$ is the posterior variance of the state. Note that the specification in (18) is standard in the signal extraction literature and captures the situation where agents happen or choose to have imperfect knowledge of the underlying shocks.[19] Since imperfect observations of the state lead to welfare losses, agents use the processed information to estimate the true state.[20] Specifically, we assume that households use the Kalman filter to update the perceived state $\widehat{s}_t=E_t\left[s_t\right]$ after observing new signals in the steady state in which the conditional variance of s_t, $\Sigma_t=\text{var}_t\left(s_t\right)$, has converged to a constant Σ:

$$\widehat{s}_{t+1}=\left(1-\theta\right)\left(R\widehat{s}_t-c_t\right)+\theta\left(s_{t+1}+\xi_{t+1}\right), \tag{19}$$

where $\theta=1-\exp\left(-2\kappa\right)$ is the Kalman gain.[21]

Combining (5) with (19), we obtain the following equation governing the perceived state $\widehat{s}_t$:

$$\widehat{s}_{t+1}=R\widehat{s}_t-c_t+\eta_{t+1}, \tag{20}$$

where

$$\eta_{t+1}=\theta R\left(s_t-\widehat{s}_t\right)+\theta\left(\zeta_{t+1}+\xi_{t+1}\right) \tag{21}$$

is the innovation to the mean of the distribution of perceived permanent income,

$$s_t-\widehat{s}_t=\frac{\left(1-\theta\right)\zeta_t}{1-(1-\theta)R\cdot L}-\frac{\theta\xi_t}{1-(1-\theta)R\cdot L} \tag{22}$$

[18] Here we assume that the agents know that their consumption function is $c_t=rs_t$ if they have perfect information on the state.

[19] Note that this noisy signal specification is consistent with that adopted in traditional signal extraction models with exogenous noises. See Angeletos and La'O (2010) and Luo and Young (2014) for its recent applications.

[20] See Luo (2008) and Luo, Nie, and Young (2015) for details about the welfare losses due to information imperfections within the partial equilibrium permanent income hypothesis framework.

[21] θ measures how much uncertainty about the state can be removed upon receiving the new signals about the state.

is the estimation error where L is the lag operator, and $E_t[\eta_{t+1}] = 0$. Note that η_{t+1} can be rewritten as

$$\eta_{t+1} = \theta\left[\left(\frac{\zeta_{t+1}}{1-(1-\theta)R\cdot L}\right) + \left(\xi_{t+1} - \frac{\theta R\xi_t}{1-(1-\theta)R\cdot L}\right)\right], \tag{23}$$

where $\omega_\xi^2 = \text{var}(\xi_{t+1}) = \frac{1}{\theta}\frac{1}{1/(1-\theta)-R^2}\omega_\zeta^2$. Note that here we need to impose that $1-\theta < (1/R)^2 < 1/R$ to ensure that the η_{t+1} sequence converges and ω_ξ^2 is finite. Note that using the definition of θ we can write this restriction as $\kappa > \ln R \cong r$. Expression (23) clearly 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 decreases with θ. More specifically, as θ increases, the first term in (23) becomes less important because $(1-\theta)\zeta_t$ in the numerator decreases, and the second term also becomes less important because the importance of ξ_t decreases as θ increases.[22]

Following Sims (2010) and Luo and Young (2014), we assume that consumers minimize the mean square error (MSE) due to imperfect observations under finite information-processing capacity. Assuming a constant information cost λ, the filtering problem can be written as:

$$\min_{\{\Sigma_t\}} \sum_{t=0}^{\infty} \beta^t \left[\Sigma_t + \lambda \ln\left(\frac{R^2\Sigma_{t-1} + \omega_\zeta^2}{\Sigma_t}\right)\right],$$

where Σ_t is variance of s_t after collecting time t information, while $R^2\Sigma_{t-1} + \omega_\zeta^2$ is the variance before information collection. This minimization problem demonstrates consumer's trade-off between the uncertainty of the perceived state and the cost attached to the reduction in uncertainty. In an extreme case when information is costless, i.e., $\lambda \to 0$, there is no informational friction as in the FI-RE model, $\Sigma = 0$; on the contrary, when $\lambda \to \infty$, $\Sigma \to \infty$, i.e., no information will be collected. The optimal steady state conditional variance can be solved as

$$\Sigma = \frac{-(1-R(R-1)\widetilde{\lambda}) + \sqrt{(1-R(R-1)\widetilde{\lambda})^2 + 4R^2\widetilde{\lambda}}}{2R^2}\omega_\zeta^2, \tag{24}$$

where $\widetilde{\lambda} \equiv \lambda/\omega_\zeta^2$ and we use the fact that $\beta R = 1$. $\widehat{s}_t$ is governed by the Kalman filtering equation

$$\widehat{s}_{t+1} = (1-\theta)(R\widehat{s}_t - c_t) + \theta(s_{t+1} + \xi_{t+1}). \tag{25}$$

The Kalman gain θ measures how much uncertainty can be removed. It is positively related to the capacity chosen to process information. Following Luo and Young (2014), θ can be obtained

$$\theta = 1 - \frac{1}{R^2}\left\{1 + \frac{2}{-\left[1-R(R-1)\widetilde{\lambda}\right] + \sqrt{\left[1-R(R-1)\widetilde{\lambda}\right]^2 + 4R^2\widetilde{\lambda}}}\right\}^{-1}. \tag{26}$$

[22]Note that when $\theta = 1$, $\text{var}(\xi_{t+1}) = 0$.

Figure 1 clearly shows that the value of θ increases with the level of macroeconomic uncertainty measured by ω_ζ^2 (i.e., $\partial\theta/\partial\omega_\zeta^2 > 0$). That is, the higher the income uncertainty, the more capacity is devoted to monitoring the evolution of the state. With a fixed information-processing cost λ, the agent is allowed to adjust the optimal level of capacity and attention in such a way that the marginal cost of information-processing for the problem at hand remains constant. This result is consistent with the concept of "elastic" capacity proposed in Kahneman (1973). Coibion and Gorodnichenko (2015) use the SPF forecast survey data to test the degree of information rigidities and find that the information rigidities were decreasing with the volatility of the macroeconomic conditions. Specifically, they find that information rigidities were falling from the late 1960s to the start of the Great Moderation (1983 − 1984) and have declined since then, and argue that one should be wary of treating the degree of information rigidities as a structural parameter because it responds to changes in macroeconomic conditions. In the following analysis, we will show that elastic attention can also be used to explain the cross-country behavior of consumption and income.

Under RI, optimal consumption of a typical consumer in the home country is

$$c_t = (R-1)\widehat{s}_t, \tag{27}$$

and the change in consumption can be expressed as

$$\Delta c_t^{RI} = \theta(R-1)\left[\frac{\zeta_t}{1-(1-\theta)R\cdot L} + \left(\xi_t - \frac{\theta R\xi_{t-1}}{1-(1-\theta)R\cdot L}\right)\right], \tag{28}$$

where L is the lag operator. Here we assume that consumers have sufficient information-processing ability (i.e., $\theta > 1 - 1/R$) to ensure that the Δc_t^{RI} sequence converges. Similarly, in the ROW, we have

$$\Delta c_t^{*RI} = \theta^*(R-1)\left[\frac{\zeta_t^*}{1-(1-\theta^*)R\cdot L} + \left(\xi_t^* - \frac{\theta^* R\xi_{t-1}^*}{1-(1-\theta^*)R\cdot L}\right)\right]. \tag{29}$$

From the MA(∞) operators, $1-(1-\theta)R\cdot L$ and $1-(1-\theta^*)R\cdot L$, in these expressions, it is clear that consumption adjusts gradually to income shocks instead of adjusting immediately and completely. When $\theta < 1$, the true state is no longer observable due to the presence of the consumer's endogenous RI-induced noises ξ. Through gradual learning and adjustment, inattention interacts with past income shocks and noises and then affects the current consumption decision and dynamics. As the consumer pays less attention (smaller θ), these shocks become more important. When $\theta = 1$, the true state can be observed perfectly and past shocks are not informative. Consequently, the above expression reduces to $\Delta c_t = (R-1)\zeta_t$, which is the same as that obtained in the FI-RE model. For different countries, the level of domestic fundamental uncertainty can affect optimal consumption decisions through optimal attention θ and its interaction with the noises. In other words, different levels of income volatility lead to different degrees of θ. The higher θ is, the more new information a country processes. Different values of θ then lead to different adjustment speeds of consumption. This helps explain why the consumption correlation is in general lower than the corresponding income correlation.

It is worth noting that a constant-absolute-risk-aversion (CARA) version of the RI-SOE model and our benchmark model are observationally equivalent in the sense that they lead to the same dynamics of aggregate consumption and savings. The key reason is that the CARA specification introduces a constant precautionary saving term into the consumption function but has no impact on the MPC out of expected total wealth.[23]

3.2 Aggregation

Since the economy consists of a continuum of identical consumers, we now need to discuss the aggregation problem. Sims (2003) argues that a considerable part of the idiosyncratic responses is common across individuals despite the heterogeneity of information noises induced by each individual's own inattention. Aggregating across all individual consumers facing the same aggregate income process using (28) yields the expression of the change in aggregate consumption:

$$\Delta c_t = (R-1)\left[\frac{\theta\zeta_t}{1-(1-\theta)R\cdot L} + \theta\left(\overline{\xi}_t - \frac{\theta R\overline{\xi}_{t-1}}{1-(1-\theta)R\cdot L}\right)\right], \tag{30}$$

where i denotes a particular individual, $E^i[\cdot]$ is the population average, and $\overline{\xi}_t = E^i[\xi_t]$ is the common noise. Figure 2 illustrates the impulse reponses of aggregate consumption the income shock for various degree of attention.

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. Define a single parameter,

$$\mu \equiv \frac{\mathrm{sd}\left(\overline{\xi}_t\right)}{\mathrm{sd}\left(\xi_t\right)} \in [0,1],$$

to measure the common source of coded information on the aggregate component (or the relative importance of $\overline{\xi}_t$ vs. ξ_t).[24] Idiosyncratic noises are cancelled out after aggregation while the common noise remains.[25] In Section 4.2, we will use the national media concentration measures documented in Noam (2016) to calibrate the value of μ, and then explore how the presence of the common noise can help improve the model's predictions on the international consumption correlation as well as the joint dynamics of consumption and income.

[23] The detailed proof of this result is available from the corresponding author by request.

[24] In a recent paper, Angeletos and La'O (2010) show how dispersed information about the aggregate productivity shock contributes to significant noise in the business cycle and helps explain cyclical variations in Solow residuals and labor wedges.

[25] Black (2010) also argues that an idiosyncratic shock along a given dimension (for different types of agents) might not be independent from agent to agent and can have a substantial aggregate effect. See Part III of Black (2010) for a detailed discussion on the law of large numbers.

3.3 Implications for Cross-Country Consumption Correlations

The following proposition summarizes how the aggregation factor affects the cross-country consumption correlation:

Proposition 1 *Given μ, the cross-country consumption correlation can be written as*

$$\operatorname{corr}(\Delta c, \Delta c^*) \approx \Pi \operatorname{corr}(\Delta y, \Delta y^*), \tag{31}$$

where

$$\Pi = \frac{1}{[1-(1-\theta)(1-\theta^*)R^2]\sigma(\theta,\mu)\sigma(\theta^*,\mu)}, \tag{32}$$

$$\sigma(\theta,\mu) = \sqrt{\frac{1}{1-[(1-\theta)R]^2} + \mu^2\left\{\frac{1}{[1-(1-\theta)R^2]\theta} - \frac{1}{1-[(1-\theta)R]^2}\right\}},$$

$$\sigma(\theta^*,\mu) = \sqrt{\frac{1}{1-[(1-\theta^*)R]^2} + \mu^2\left\{\frac{1}{[1-(1-\theta^*)R^2]\theta^*} - \frac{1}{1-[(1-\theta^*)R]^2}\right\}}.$$

Proof. See Appendix 6.1. ■

When $\theta = \theta^* = 1$, $\Pi = 1$, which means the consumption correlation should be as high as the income correlation, which contradicts the empirical findings. Introducing elastic attention ($\lambda > 0$ and $\theta < 1$) has three distinct effects on the consumption correlation:

1. *The slow propagation channel* ($\theta = \theta^* < 1$): If we shut down the endogenous RI-induced noises, RI only introduces slow adjustment into the model. The mechanism reduces the variance of consumption growth and the covariance of cross-country consumption by the same magnitude such that the consumption correlation remains the same. Note that in this case, $\mu = 0$ and (32) reduces to

$$\Pi = \frac{\sqrt{1-[(1-\theta)R]^2}\sqrt{1-[(1-\theta^*)R]^2}}{1-(1-\theta)(1-\theta^*)R^2},$$

   which equals 1 when $\theta = \theta^*$.

2. *The common noise channel*: The presence of the endogenous noise (ξ), which is uncorrelated across countries, increases the variance of consumption without changing the cross-country covariance, and thus reduces the consumption correlation.

3. *The elastic attention channel*: The consumption correlation is further reduced by the difference between θ and θ^*. As θ and θ^* deviate further from each other, the consumption correlation becomes smaller relative to the income correlation.

For the slow adjustment channel, it is worth noting that this channel is similar to that of three other models: the habit formation model, the model with incomplete information about income, and the inattentiveness and infrequent adjustment model. The main reason is that all of these hypotheses lead to slow adjustment in aggregate consumption.[26]

For the common noise channel, Expression (31) shows that the higher the value of μ (i.e., the common noise is more important), the higher the variance of consumption growth and the lower the international consumption correlation. Figures 3 and 4 show that the consumption correlation is decreasing with μ for any given values of θ.

The last channel is identified uniquely in our elastic attention model, in which the attention levels, θ and θ^*, are optimally chosen by consumers based on their own domestic countries' income uncertainty. Income uncertainty in two economies are different by nature, which leads to different levels of optimal attention. Figure 5 shows how Π varies with the value of θ for a given θ^*. To further explore the impact of elastic attention on the consumption correlation, we consider a special case in which all noises are idiosyncratic ($\mu = 0$). In this case, we have the following expression for the change in aggregate consumption in the home country:

$$\Delta c_t^{RI} = \theta(R-1)\frac{\zeta_t}{1-(1-\theta)R\cdot L}, \tag{33}$$

Similarly, we can obtain the change in ROW's aggregate consumption

$$\Delta c_t^{*RI} = \theta^*(R-1)\frac{\zeta_t^*}{1-(1-\theta^*)R\cdot L}. \tag{34}$$

The consumption correlation between the two countries becomes

$$\operatorname{corr}(\Delta c_t^{RI}, \Delta c_t^{*RI}) \approx \overline{\Pi}\operatorname{corr}(\Delta y_t, \Delta y_t^*), \tag{35}$$

where

$$\overline{\Pi} = \frac{\sqrt{(1-(1-\theta)^2R^2)(1-(1-\theta^*)^2R^2)}}{1-(1-\theta)(1-\theta^*)R^2} \leq 1.$$

Since $\mu = 0$, the endogenous noise component disappears. The heterogeneity across countries introduced by elastic capacity depends on the difference between θ and θ^*. When $\theta = \theta^*$, $\overline{\Pi} = 1$, which gives the same predictions as the standard FI-RE model. As the difference between θ and θ^* increases, the consumption correlation becomes smaller relative to the income correlation.

Comparing the implications of the consumption correlation obtained in the only-common-noise case ($\mu = 1$), and the no-common-noise case ($\mu = 0$), we have $\Pi \in [\underline{\Pi}, \overline{\Pi}]$, where

$$\underline{\Pi} = \frac{\sqrt{\theta\theta^*(1-(1-\theta)R^2)(1-(1-\theta^*)R^2)}}{1-(1-\theta)(1-\theta^*)R^2} \text{ and } \overline{\Pi} = \frac{\sqrt{(1-(1-\theta)^2R^2)(1-(1-\theta^*)^2R^2)}}{1-(1-\theta)(1-\theta^*)R^2}.$$

[26]See the online appendix for detailed comparisons between the RI model with these three models with slow adjustment in aggregate consumption.

Using these explicit expressions, we can explore the implications of elastic attention on the correlations for different countries by varying the values of μ and θ^*. We have two interesting findings. First, given the difference between θ and θ^*, the consumption correlation decreases with μ. A higher value of μ means that the common noise plays a more important role in reducing the correlation. Second, given μ, $\operatorname{corr}(\Delta c, \Delta c^*)$ is increasing in θ^*, which is the same as in the representative agent model ($\mu = 1$). In the quantitative analysis conducted in the next section, we will show that our model can fit the data better for many plausibel combinations of the three key parameter values, θ, θ^*, and μ.

3.4 Implications for Other Stochastic Properties of Consumption

Given the exogenous income process and the consumption rule, we can readily obtain other key stochastic properties of the joint dynamics of consumption and income under elastic attention. The following proposition summarizes the implications of elastic attention for the relative volatility, persistence, and correlation with output of consumption in the home country:

Proposition 2 *Under RI, the relative volatility of consumption change to income change (i.e., the excess smoothness ratio) can be written as:*

$$\mathrm{rv} = \frac{\mathrm{sd}(\Delta c)}{\mathrm{sd}(\Delta y)} = \sqrt{\frac{\theta^2}{1 - [(1-\theta)R]^2} + \mu^2 \left\{ \frac{\theta}{1 - (1-\theta) R^2} - \frac{\theta^2}{1 - [(1-\theta)R]^2} \right\}}, \tag{36}$$

the first-order autocorrelation of consumption change is

$$\rho_c(1) = \frac{\left(1-\mu^2\right)(1-\theta)R}{1 + \mu^2 \left[\frac{1-(1-\theta)^2 R^2}{(1-(1-\theta)R^2)\theta} - 1 \right]}, \tag{37}$$

and the contemporaneous correlation between consumption change and income change is

$$\operatorname{corr}(\Delta c, \Delta y) = \frac{1}{\sqrt{\frac{1}{1-[(1-\theta)R]^2} + \mu^2 \left\{ \frac{1}{(1-(1-\theta)R^2)\theta} - \frac{1}{1-[(1-\theta)R]^2} \right\}}}. \tag{38}$$

Proof. See Appendix 6.2. ■

Using these explicit expressions, Figure 6 illustrates how RI affects the three key stochastic properties of consumption and income. It is clear from this figure that for given values of μ, the relative consumption volatility and the consumption-output correlation are increasing with the degree of attention, and the first-order autocorrelation of consumption is decreasing with the degree of attention. The intuition for these results is that the slow adjustment channel dominates the common noise channel when the aggregation factor is not very high. In addition, as we can also see from Table 5, for most of the combinations, our model fits the data better. Compared to the benchmark model, we can obtain a positive first-order consumption autocorrelation and relative volatility of consumption to income less than 1.

When μ is sufficiently high, the consumption correlation is decreasing in θ, because the noise channel governed by the ξ_t term dominates the slow propagation channel governed by the $1-(1-\theta)R\cdot L$ term. The volatility of consumption is decreasing in θ due to less induced noises. When μ is not sufficiently high, rv is increasing in θ since the slow propagation channel takes control and increases volatility as θ goes up. Note that in the representative agent model, the excess smoothness ratio is $\sqrt{\frac{\theta}{1-(1-\theta)R^2}} \geq 1$. Imperfectly observing the state reduces the ability to smooth consumption and thus results in excess volatility of consumption. On the other hand, $rv = \sqrt{\frac{\theta^2}{1-(1-\theta)^2R^2}} \leq 1$ when $\mu = 0$. Therefore, given θ,

$$\text{rv} \in \left[\sqrt{\frac{\theta^2}{1-[(1-\theta)R]^2}}, \sqrt{\frac{\theta}{1-(1-\theta)\,R^2}}\right].$$

For example, if $\theta = 40\%$ and $\mu = 0.1$, the model predicts that $rv = 0.52$, which is close to its empirical counterpart in the U.S. data (around 0.54).

Given a fixed μ, the autocorrelation of consumption growth is decreasing with θ since the response of consumption to the noise has a negative relationship with consumption growth over time. The consumption-income correlation is increasing in θ given a fixed μ. A reduction in μ leads to a higher autocorrelation and a higher consumption-income correlation. The intuition is that more idiosyncratic noises are cancelled out and the noise channel reduces the variance of consumption growth by a smaller amount, while the covariance between consumption growth and income growth remains the same.

4 Quantitative Implications

4.1 Data

We use annual data between 1950 and 2010 from Penn World Tables (PWT), both version 7.1 and version 8.0, to study the consumption-income correlation between each of the four smaller economies of the G7 (Canada, Italy, UK and France) and the rest of the world economy. All variables are at 2005 constant prices. The ROW economy with respect to each country is constructed using the weighted average of the G7 countries excluding the domestic country. The correlation between Canada and the U.S. is also studied as a special case. The U.S. is treated as ROW to Canada since over 70% of Canada's international trade is with the U.S. (See, for example, Miyamoto and Nguyen 2014). We will discuss more about this pair of countries in Section 4.3.

Table 1 demonstrates the international consumption correlation puzzle. The puzzle persists in both per capita data and aggregate data. In this paper, we choose to discuss aggregate data, which is more consistent with our discussion of aggregation in the model. Table 2 summarizes the key empirical findings. The numbers in parentheses are GMM-corrected standard errors.

4.2 Parameter Values

We choose the fixed Kalman gain for the rest of the world, θ^*, to fit the consumption-income dynamics within the ROW. For the four ROWs we constructed and the U.S. economy, we obtain that $\operatorname{sd}(\Delta c^*)/\operatorname{sd}(\Delta y^*) \in [0.57, 0.58]$ and $\operatorname{corr}(\Delta c^*, \Delta y^*) \in [0.91, 0.93]$. If we assume there is no common noise in the ROW, our benchmark model can match $\operatorname{sd}(\Delta c^*)/\operatorname{sd}(\Delta y^*)$ when $\theta^* \in [0.48, 0.50]$ and match $\operatorname{corr}(\Delta c^*, \Delta y^*)$ when $\theta^* \in [0.60, 0.62]$. In the following analysis, for tractability, we assume that the value of θ^* is set to fall in the range of $[0.5, 1]$. It is worth noting that a less-than-one value of θ can be rationalized by examining the welfare effects of limited capacity.[27] In the RI literature, to explain the observed aggregate fluctuations and the effects of monetary policy on the macroeconomy, the calibrated values of θ are lower and deviate more from the FI-RE case. For example, Adam (2007) found $\theta = 0.4$ based on the response of aggregate output to monetary policy shocks. Luo (2008) find that if $\theta = 0.5$, the otherwise standard permanent income model generates a realistic relative volatility of consumption to labor income. Maćkowiak and Wiederholt (2015) find that given a total information flow of 133 bits, the decision-maker of the typical firm only allocates 0.76 bit of information flow to tracking aggregate technology and 0.41 bit to tracking monetary policy. Therefore, the exogenous capacity given in our model can be regarded as a shortcut to small fractions of consumers' total capacity used to monitor their total resources hit by the innovation to total resources. Coibion and Gorodnichenko (2015) proposed an empirical model to test information rigidities including the RI specification for a variety of economic agents such as consumers, firms, and financial market participants for whom forecast data are available. They find that the RI specification is likely to play a pervasive role in determining macroeconomic dynamics, and the model can fit the data quite well when $\theta = 0.5$ in their forecasting model. It is worth noting that although the value of θ in this range is not a large number and is well below the total information-processing ability of human beings, it is not unreasonable in practice for ordinary consumers because they also face many other competing demands on capacity.

The value of $\widetilde{\lambda}^*$ can be recovered by solving equation (26). Given that $\widetilde{\lambda}^* \equiv \lambda^*/\omega_\zeta^{*2}$ and that both domestic country and ROW consumers are facing a stable information cost, i.e., $\lambda = \lambda^*$, we have

$$\widetilde{\lambda} = \frac{\lambda}{\omega_\zeta^2} = \frac{\widetilde{\lambda}^* \omega_\zeta^{*2}}{\omega_\zeta^2}.$$

Plugging the expression for $\widetilde{\lambda}$ back into Equation (26), we can get θ.

[27] See Luo (2008) and Luo et al. (2015) for details about welfare losses due to imperfect observations within the linear-quadratic-Gaussian permanent income framework; they are uniformly small.

We do not specify a value for μ at first and study how the consumption correlations change as we vary μ from 0 to 0.9. Recall that a higher μ means a larger part of information noises are common across individuals. In the next step, we will calibrate the value for μ according to national media concentration measures. Intuitively, if a country's media is highly concentrated and lack of variety, individuals receive information from limited sources and are likely to be biased towards the same direction. Media concentration measures are all borrowed from Noam (2016). While there are a few alternative measures, the following are the indexes related to our definition of μ: (1) pooled C1, (2) average C1, and (3) content C1, which are explained in detail below.

Pooled C1 is the largest media company's share of the overall national media market. If the top 1 company takes up, for example, 20% of the market, the pooled C1 would be 20. The overall market is divided to 13 industries, newspapers, magazines, broadcast TV, search engines, etc., in Noam (2016). For each industry, C1 measures the market share of the largest company in that industry. Average C1 is the weighted average adjusting for different sizes of the 13 media industries. Pooled C1 is lower than average C1 for all countries as pooled index calculates market shares of the larger overall market. Content C1 refers to C1 index of content creators.

Finally, following Glick and Rogoff (1995), the interest rate (R) and the depreciation rate (δ) are set to be 1.04 and 0.05, respectively. Now we are ready to calculate Π and then use Equation (31) to determine $\text{corr}(\Delta c_t^{RI}, \Delta c_t^{*RI})$.

4.3 Main Results

Table 2 reports the summary of statistics. The values of $\text{corr}\left(\Delta y, \Delta y^*\right)$ (and $\text{corr}\left(\Delta c, \Delta c^*\right)$) are the simple correlation coefficients between the annual change in country's real output (and consumption) and the annual change of the rest of the world's real output (and consumption), with the "world" defined as the output-weighted average of the rest of the G7 countries in the Penn World Tables (version 7.1). The Canada-US correlations are the output (or consumption) correlations between Canada and the U.S.

Table 3 compares the cross-country consumption correlations between the FI-RE and RI models with different values of the Kalman gain of the ROW (θ^*) and the aggregation factor (μ). (The first column summarizes the empirical findings from the data.) Our key result here is that introducing RI improves the performance of the model in terms of matching the cross-country consumption and income correlations. As shown in the second column of Table 3, the FI-RE model predicts that the consumption correlations are almost as high as the income correlations. In contrast, as shown in the last five columns corresponding to different values of θ^*, we can see that when introducing elastic attention, a small deviation from the FI-RE case can significantly improve the model's predictions and generate much lower consumption correlations, which fit the data better. For example, the FI-RE model predicts that the consumption correlation between Canada and the ROW is 0.83 and can be reduced to its empirical counterpart, 0.56, in the elastic attention model when $\theta^* = 0.9$ and $\mu = 0.3$ (i.e., 90% of uncertainty is removed upon new signals and 30% of the noise information is remained after aggregation.) In addition, we can see that in the $\mu = 0$ case (i.e., all of the RI-induced noises are cancelled out after aggregation), the correlation is reduced to 0.66 when $\theta^* = 0.9$ and 0.59 when $\theta^* = 0.5$. For France, when $\theta = 0.8$ and $\mu = 0.3$, the elastic attention model predicts that $\text{corr}(\Delta c, \Delta c^*) = 0.33$, which is the same as the empirical counterpart. Given the value of μ, it is clear from the table that the correlation is decreasing with the degree of attention because the elastic attention channel becomes more and more important as the degree of inattention increases. Furthermore, we can see from the table that the correlation is also decreasing with μ because the common noise channel becomes more important as μ increases. In all cases, elastic attention helps reduce the consumption correlation and makes the model match the data better.

It is worth noting that in this exercise the optimal degree of attention is different for each country. That is, we fix the value of λ, the cost of obtaining information, and the degree of attention, θ, is implied by the optimal decision, (26). Table 4 reports the implied values of θ in Table 3 for each country in each case. The rows with $\mu = 0$ in Table 3 show the results obtained from the model with the elastic attention channel alone when the common noise channel is muted. They show that the consumption correlations are in general lower for all values of θ^*. Even small deviations from the FI-RE model ($\theta^* = 0.9$) drives down the consumption correlation.

In the rows with $\mu =$pooled C1/100 in Table 3, μ is calibrated using one of the national media concentration measures, pooled C1, in each of the small open economies. For simplicity, here we assume $\mu^* = \mu$ for the rest of the world. The results clearly show that for calibrated values of μ using the measure of national media concentration, our rational inattention model does a good job in matching the observed cross-country consumption correlations even when the degree of inattention is low, i.e., when the RI model only slightly deviates from the FI-RE model. For example, when $\theta = 90\%$ and using the pooled C1 measure, the model predicts that the values of $\text{corr}\,(\Delta c, \Delta c^*)$ are 0.56, 0.38, and 0.69 for Canada, France, and UK, respectively, which are very close to the empirical counterparts.[28]

Table 5 compares the other key stochastic properties of the joint dynamics of consumption and income: (i) the relative volatility of consumption to income, (ii) the first-order autocorrelation of consumption, and (iii) the contemporaneous consumption-income correlation in individual countries between the FI-RE and RI models when θ^* varies from 0.8 to 0.9. Our key result here is that RI significantly improves the performance of the model in terms of these consumption moments; for different combinations of the key parameters, θ^* and μ, each model economy has more realistic consumption dynamics. Quantitatively, we can see that the improvements are significant for all

[28]The conclusion is robust when we use the other two measures of national media concentration (average C1 and content C1). The results are reported in Tables A1 and A2 in the online appendix.

countries we studied. For example, in Canada, the relative consumption volatility falls from 1 in the FI-RE case to 0.48 in the elastic attention case in which $\theta^* = 0.9$ and $\mu = 0$, which is much closer to its empirical counterpart, 0.52. The autocorrelation rises from 0 to 0.67, which is closer to its empirical counterpart, 0.60. The consumption-income correlation falls from 1 to 0.75, which is exactly its empirical counterpart. These findings are consistent with our theoretical results obtained in Section 3.4. The improvements of the model's predictions remain significant when the calibration exercise is repeated to include OECD countries.[29]

4.4 The Canada-US Case

The Canada-US case is of interest because Canada and the U.S. have one of the world's closest bilateral relationships. Total trade of goods (imports and exports) in 2014 amounted to 750.8 billion dollars. In addition, Canada is a typical small open economy studied in the international economics literature. We now apply our elastic attention model to study their correlations, treating the U.S. as the rest of the world to Canada. The U.S. is a reasonable approximation of the rest of the world to Canada because their relationship is highly asymmetric. First, Canada relies on the U.S. as its principal trading partner. 71% of Canada's total goods trade was conducted with the U.S in 2014. Over the period $1973 - 2012$, on average, 75% of Canadian exports and 68% of imports were traded with the U.S. (See Minamoto and Nguyen, 2014; data from Statistics Canada.) Second, the U.S. is overwhelmingly larger than Canada, with an economy more than 10 times the size of the Canadian economy.

The cross-country consumption correlation puzzle also exists in this special group. Specifically, the correlation coefficient of the change in annual real output between Canada and the U.S. is 0.87, while the corresponding consumption correlation is only 0.58. From the final five rows of Table 3, we can see that the FI-RE model predicts that the consumption correlation should be approximately the same as the output correlation, 0.87. By contrast, in a small deviation from the FI-RE case in which we assume that the typical consumer in the U.S. has limited capacity $\theta^* = 0.8$, the elastic attention model predicts that the consumption correlation between Canada and US is only 0.56, which is much closer to its empirical counterpart (0.58). It is also clear from the same table that many combinations of θ^* and μ have the potential to explain the empirical correlation well. In this special case, the low consumption correlation between the two countries is due to the two channels we discussed in our theoretical model. Specifically, one is due to the presence of the common noise that is endogenously induced by inattention. For the Canadian and the US people,

[29] We report the results in Tables A3 and A4 in the online appendix. The Rest of the World economy is the weighted average of the rest of OECD countries instead. For example, the rational inattention model with $\theta^* = 0.7$ and $\mu = 0.2$ predicts the consumption correlation between France and the ROW to be 0.38 and coincides with data, which is improved from FI-RE model's prediction, 0.51.

their information noises have zero covariance but the presence of the common noise increases the variance of their own consumption innovations, and therefore decreases the correlation. The other channel is due to the different levels of fundamental uncertainty in the two countries. Since the variance of annual change of output in the U.S. is larger than in Canada, the optimal attention level in the two countries would be generally different. If we assume the attention level for the U.S. is 0.8 and the two countries face the same marginal information-processing cost, the attention level for Canada would be lower, which further lowers the consumption correlation.

5 Conclusion

We have examined how introducing optimal attention (or elastic attention) into an otherwise standard small open economy model changes international consumption-income correlations and the joint dynamics of consumption and income. Specifically, we have shown that a rational inattention model with agents whose attention is elastic to exogenous income processes has the potential to better explain the observed international diversification and the consumption and income correlation in four small open economies in the G7. In addition, we found that the elastic attention assumption can also better explain the other key stochastic properties of the joint dynamics of consumption and income.

6 Appendix

6.1 Deriving International Consumption Correlations under RI

Given μ and the change in aggregate consumption expression (30), the variance of aggregate consumption growth can be written as

$$\begin{aligned}\operatorname{var}(\Delta c) &= \theta^2(R-1)^2\left\{\frac{1}{1-(1-\theta)^2R^2}\omega_\zeta^2+\mu^2\left[1+\frac{\theta^2R^2}{1-(1-\theta)^2R^2}\right]\omega_\xi^2\right\}\\ &= \theta^2(R-1)^2\left\{\frac{1}{1-(1-\theta)^2R^2}+\mu^2\left[1+\frac{\theta^2R^2}{1-(1-\theta)^2R^2}\right]\frac{1}{\theta}\frac{1}{1/(1-\theta)-R^2}\right\}\omega_\zeta^2\\ &= \theta^2(R-1)^2\left\{\frac{1}{1-[(1-\theta)R]^2}+\mu^2\left[\frac{1}{(1-(1-\theta)R^2)\theta}-\frac{1}{1-[(1-\theta)R]^2}\right]\right\}\omega_\zeta^2, \qquad (39)\end{aligned}$$

where we use the fact that $\omega_\xi^2=\operatorname{var}(\xi_{t+1})=\frac{1}{\theta}\frac{1}{1/(1-\theta)-R^2}\omega_\zeta^2$. Similarly, we can derive

$$\operatorname{var}(\Delta c^*)=\theta^{*2}(R-1)^2\left\{\frac{1}{1-[(1-\theta^*)R]^2}+\mu^2\left[\frac{1}{(1-(1-\theta^*)R^2)\theta}-\frac{1}{1-[(1-\theta^*)R]^2}\right]\right\}\omega_{\zeta^*}^2$$

for the ROW. Since RI-induced noises are assumed to be uncorrelated with fundamental shocks and across countries, the covariance between Δc and Δc^* does not depend on the information noise part of the expression in (30):

$$\operatorname{cov}(\Delta c, \Delta c^*) = \frac{\theta\theta^* (R-1)^2}{1-(1-\theta)(1-\theta^*)R^2} E\left[\zeta_t \zeta_t^*\right].$$

Therefore, the correlation between Δc and Δc^* can be written as:

$$\operatorname{corr}(\Delta c, \Delta c^*) = \frac{\operatorname{cov}(\Delta c, \Delta c^*)}{\operatorname{var}(\Delta c)\operatorname{var}(\Delta c^*)} \approx \frac{1}{[1-(1-\theta)(1-\theta^*)R^2]\sigma(\theta,\mu)\sigma(\theta^*,\mu)} \operatorname{corr}(\Delta y, \Delta y^*), \quad (40)$$

where $\sigma(\theta,\mu) = \sqrt{\frac{1}{1-[(1-\theta)R]^2} + \mu^2 \left\{ \frac{1}{(1-(1-\theta)R^2)\theta} - \frac{1}{1-[(1-\theta)R]^2} \right\}}$ and we use the facts that $\zeta_t \approx \frac{1}{R-1}\varepsilon_t$ and $\operatorname{corr}(\Delta y_t, \Delta y_t^*) \approx \frac{E[\varepsilon_t \varepsilon_t^*]}{\omega\omega^*}$.

6.2 Deriving Other Stochastic Properties of Consumption under RI

Given the income processes proposed in Section 2.2, we have $\Delta y_t \approx \varepsilon_t \approx (R-1)\,\zeta_t$ and $\operatorname{var}(\Delta y) \approx \omega^2 \approx (R-1)^2 \omega_\zeta^2$. With the variance of consumption growth we have derived in (39), we can compute the moments as follows:

$$\frac{\operatorname{sd}(\Delta c)}{\operatorname{sd}(\Delta y)} = \sqrt{\frac{\theta^2}{1-[(1-\theta)R]^2} + \mu^2 \left\{ \frac{\theta}{1-(1-\theta)R^2} - \frac{\theta^2}{1-[(1-\theta)R]^2} \right\}},$$

$$\begin{aligned}
\frac{\operatorname{cov}(\Delta c_t, \Delta c_{t-1})}{\operatorname{var}(\Delta c)} &= \frac{\left(1-\mu^2\right) \frac{(1-\theta)R}{1-(1-\theta)^2 R^2}}{\frac{1}{1-[(1-\theta)R]^2} + \mu^2 \left[\frac{1}{(1-(1-\theta)R^2)\theta} - \frac{1}{1-[(1-\theta)R]^2} \right]} \\
&= \frac{\left(1-\mu^2\right)(1-\theta)R}{1+\mu^2 \left[\frac{1}{\theta} + \frac{1}{1-(1-\theta)R^2} - 2 \right]},
\end{aligned}$$

and

$$\begin{aligned}
\frac{\operatorname{cov}(\Delta c, \Delta y)}{\operatorname{sd}(\Delta c)\operatorname{sd}(\Delta y)} &= \frac{\theta (R-1)^2}{\theta(R-1)\sqrt{\frac{1}{1-[(1-\theta)R]^2} + \mu^2 \left[\frac{1}{(1-(1-\theta)R^2)\theta} - \frac{1}{1-[(1-\theta)R]^2} \right]}(R-1)} \\
&= \frac{1}{\sqrt{\frac{1}{1-[(1-\theta)R]^2} + \mu^2 \left[\frac{1}{(1-(1-\theta)R^2)\theta} - \frac{1}{1-[(1-\theta)R]^2} \right]}}.
\end{aligned}$$

References

[1] Adam, Klaus (2007), "Optimal Monetary Policy with Imperfect Common Knowledge," *Journal of Monetary Economics* 54(2), 276-301.

[2] Angeletos, George-Marios and Jennifer La'O (2010), "Noisy Business Cycles," *NBER Macroeconomics Annual 2009* Vol. 24, 319-378.

[3] Backus, David K., Patrick J. Kehoe, and Finn E. Kydland (1992), "International Real Business Cycles," *Journal of Political Economy* 84(1), 84-103.

[4] Bai, Yan and Zhang Jing (2012), "Financial Integration and International Risk Sharing," *Journal of International Economics* 86(1), 17-32.

[5] Bai, Yan and José-Víctor Ríos-Rull (2015), "Demand Shocks and Open Economy Puzzles, " *American Economic Reviews: Papers and Proceedings* 105(5), 644-649.

[6] Black, Fischer (2010), *Exploring General Equilibrium.* The MIT Press.

[7] Boz, Emine, Christian Daude, and C. Bora Durdu (2011), "Emerging Market Business Cycles: Learning about the Trend," *Journal of Monetary Economics* 58, 616-631.

[8] Chen, Kan and Mario Crucini (2014), "Trends and Cycles in Small Open Economies: Making the Case for a General Equilibrium Approach," Working Paper.

[9] Cheremukhin, Anton and Antonella Tutino (2016), "Information Rigidities and Asymmetric Business Cycles," forthcoming in *Journal of Economic Dynamics and Control.*

[10] Colacito, Ric, Max Croce, Steven Ho, and Philip Howard (2014) "BKK the EZ Way. International Long-Run Growth News and Capital Flows,"Working Paper.

[11] Cochrane, J. (2005), *Asset Pricing*, Princeton University Press.

[12] Coibion, Oliver and Yuriy Gorodnichenko (2012), "What Can Survey Forecasts Tell Us About Informational Rigidities?," *Journal of Political Economy* 120(1), 116-159.

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

[14] Devereux, Michael, Allan Gregory, and Gregor Smith (1992), "Realistic Cross-Country Consumption Correlations in a Two-Country, Equilibrium, Business Cycle Model," *Journal of International Money and Finance* 11(1), 3-16.

[15] Engel, Charles and Jian Wang (2011), "International Trade in Durable Goods: Understanding Volatility, Cyclicality, and Elasticities, " *Journal of International Economics* 83(1), 37-52.

[16] Fuhrer, Jeffrey C. and Michael W. Klein (2006), "Risky Habits: On Risk Sharing, Habit Formation, and the Interpretation of International Consumption Correlations," *Review of International Economics* 14(4), 722-740.

[17] Glick, Reuven and Kenneth Rogoff (1995), "Global vs. Country-Specific Productivity Shocks and the Current Account," *Journal of Monetary Economics* 35(1), 159-192.

[18] Hall, Robert E. (1978), "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy* 86(6), 971-987.

[19] Hansen, Lars Peter (1987), "Calculating Asset Prices in Three Example Economies," in Bewley, Truman F. (ed.), Advances in Econometrics, Fifth World Congress, Cambridge University Press.

[20] Hong, Harrison, Walter Torous, and Rossen Valkanov (2007), "Do Industries Lead Stock Markets?" *Journal of Financial Economics* 83(2), 367-396.

[21] Kahneman, Daniel (1973), *Attention and Effort*, Prentice-Hall Press.

[22] Kehoe, Patrick and Fabrizio Perri (2002), "International Business Cycles with Endogenous Incomplete Markets," *Econometrica* 70(3), 907-928.

[23] Kollmann, Robert (1996), "Incomplete Asset Markets and the Cross-Country Consumption Correlation Puzzle," *Journal of Economic Dynamics and Control* 20(5), 945-962.

[24] Lewis, Karen K. (1996), "What Can Explain the Apparent Lack of International Consumption Risk Sharing?" *Journal of Political Economy* 104(2), 267-297.

[25] Luo, Yulei (2008), "Consumption Dynamics under Information Processing Constraints," *Review of Economic Dynamics* 11(2), 366-385.

[26] Luo, Yulei and Eric R. Young (2014), "Signal Extraction and Rational Inattention," *Economic Inquiry* 52(2), 811-829.

[27] Luo, Yulei and Eric R. Young (2016), "Induced Uncertainty, Market Price of Risk, and the Dynamics of Consumption and Wealth," *Journal of Economic Theory* 163, 1-41.

[28] Luo, Yulei, Jun Nie, and Eric R. Young (2014), "Robust Control, Informational Frictions, and International Consumption Correlations," *European Economic Review* 67, 1-27.

[29] Luo, Yulei, Jun Nie, and Eric R. Young (2015), "Slow Information Diffusion and the Inertial Behavior of Durables Consumption," *Journal of the European Economic Association*, 1-36.

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

[31] Maćkowiak, Bartosz and Mirko Wiederholt (2015), "Business Cycle Dynamics under Rational Inattention," *Review of Economic Studies* 82(4): 1502-1532.

[32] Matejka, Filip and Alisdair McKay (2015), "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *American Economic Review* 105(1), 272-98.

[33] Miyamoto, Wataru and Nguyen, Thuy Lan (2014), "Understanding the Cross Country Effects of US Technology Shocks," Working Paper.

[34] Noam, Eli M. and The International Media Concentration Collaboration (2016), *Who Owns the World's Media? Media Concentration and Ownership around the World*, Oxford University Press.

[35] Obstfeld, Maurice (1994), "Are Industrial-Country Consumption Risks Globally Diversified?" in Leiderman, Leonardo and Assaf Razin (eds.), *Capital Mobility: The Impact on Consumption, Investment and Growth.* Cambridge University Press, Cambridge.

[36] Obstfeld, Maurice, and Rogoff, Kenneth (1996), *Foundations of International Macroeconomics*, The MIT Press.

[37] Pischke, Jörn-Steffen (1995), "Individual Income, Incomplete Information, and Aggregate Consumption," *Econometrica*, 63, 805-840.

[38] Quah, Danny (1990), "Permanent and Transitory Movements in Labor Income: An Explanation for "Excess Smoothness" in Consumption," *Journal of Political Economy* 98, 449-475.

[39] Reis, Ricardo (2006), "Inattentive Consumers," *Journal of Monetary Economics* 53(8), 1761-1800.

[40] Sims, Christopher A. (2003), "Implications of Rational Inattention," *Journal of Monetary Economics* 50(3), 665-690.

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

[42] 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.

[43] Stockman, Alan and Linda Tesar (1995), "Tastes and Technology in a Two-Country Model of the Business Cycle: Explaining International Comovements," *American Economic Review* 85(1), 168-185.

[44] Wen, Yi (2007), "By Force of Demand: Explaining International Comovements," *Journal of Economic Dynamics and Control* 31(1), 1-23.

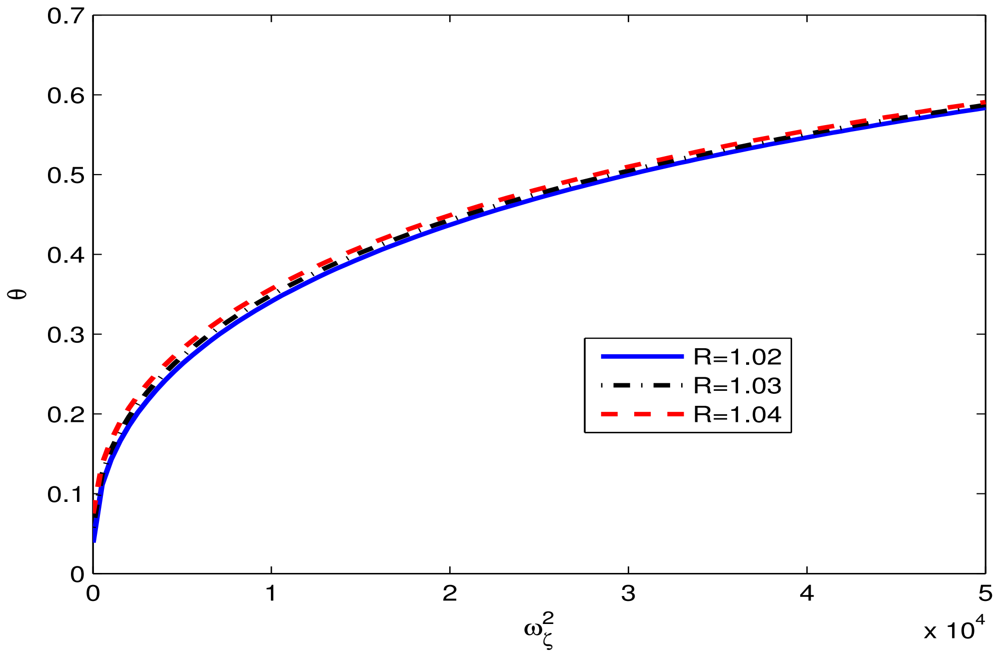

Figure 1: Effect of Fundamental Uncertainty on the Kalman Gain ($\lambda = 6.4 \times 10^4$)

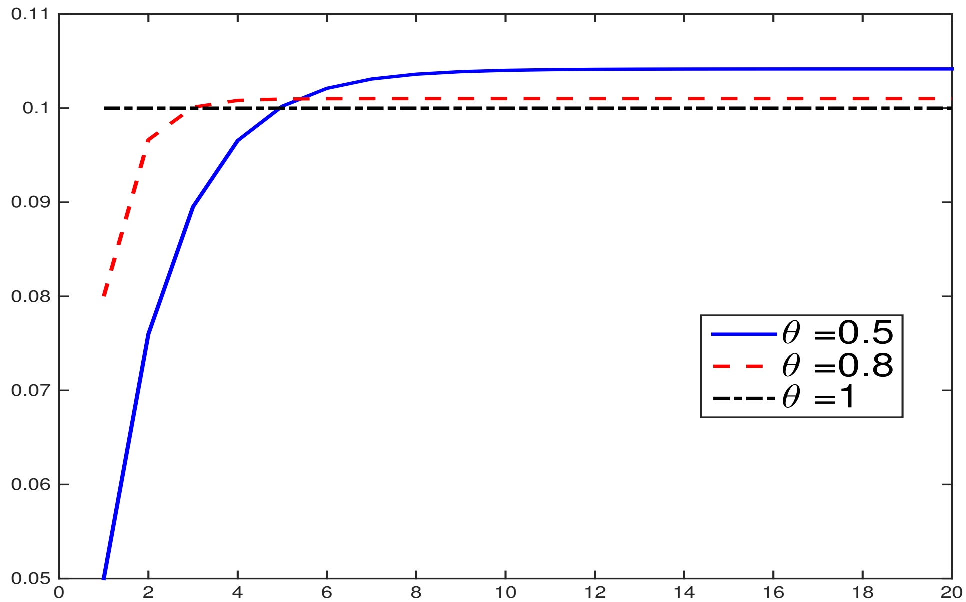

Figure 2: Impulse Responses of Aggregate Consumption to Income Shock

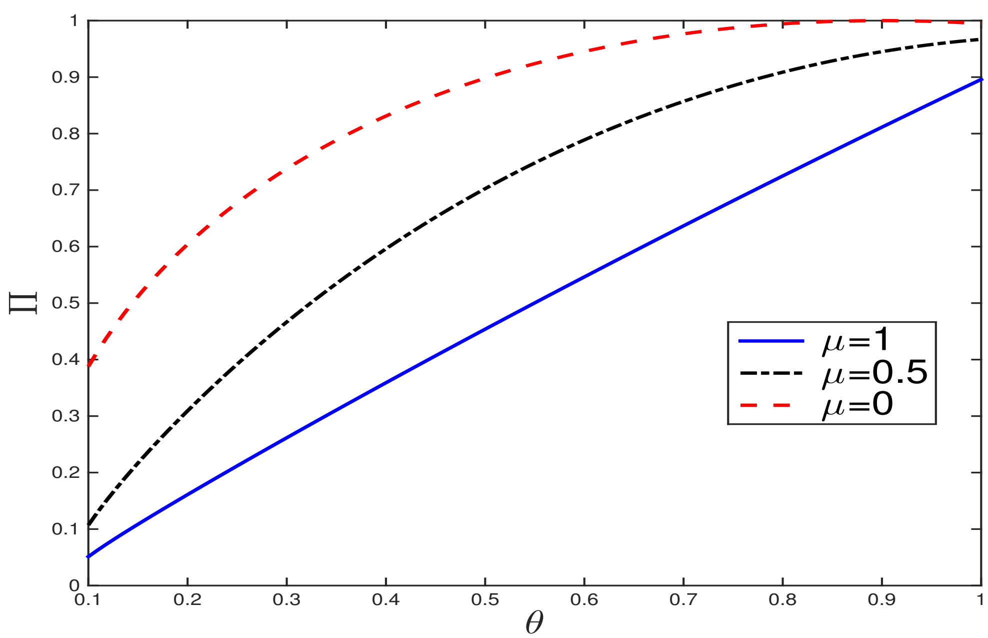

Figure 3: Effects of the Common Noise ($\theta^* = 0.9$)

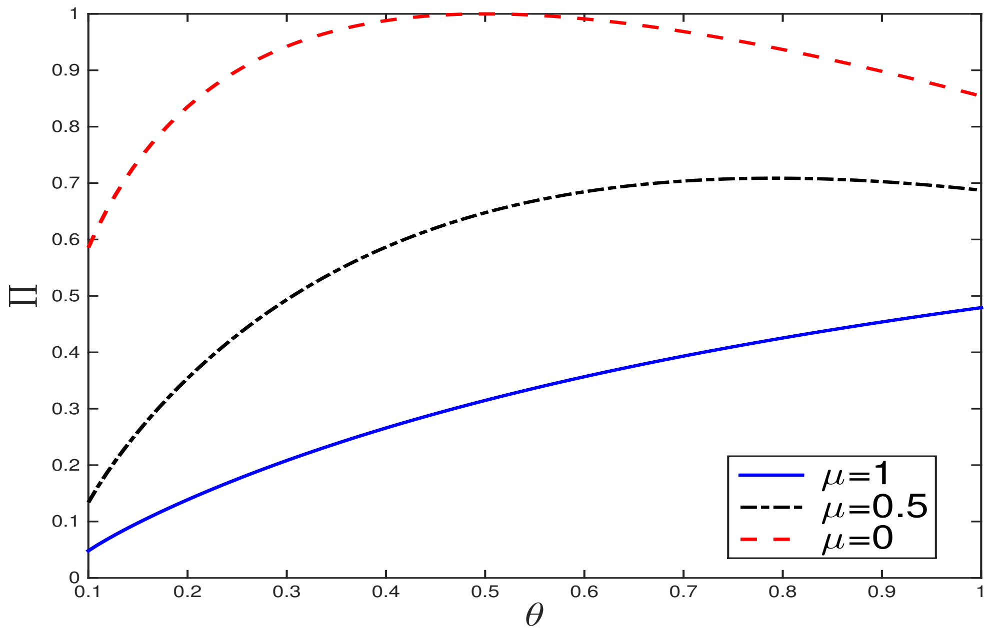

Figure 4: Effects of the Common Noise ($\theta^* = 0.5$)

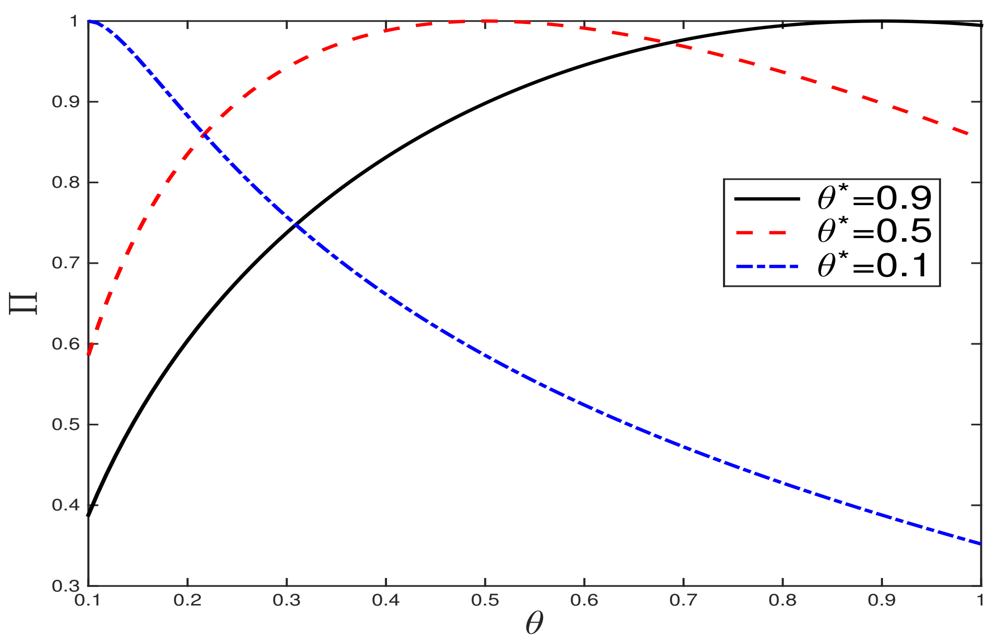

Figure 5: Effects of Elastic Attention on Consumption Correlation ($\mu = 0$)

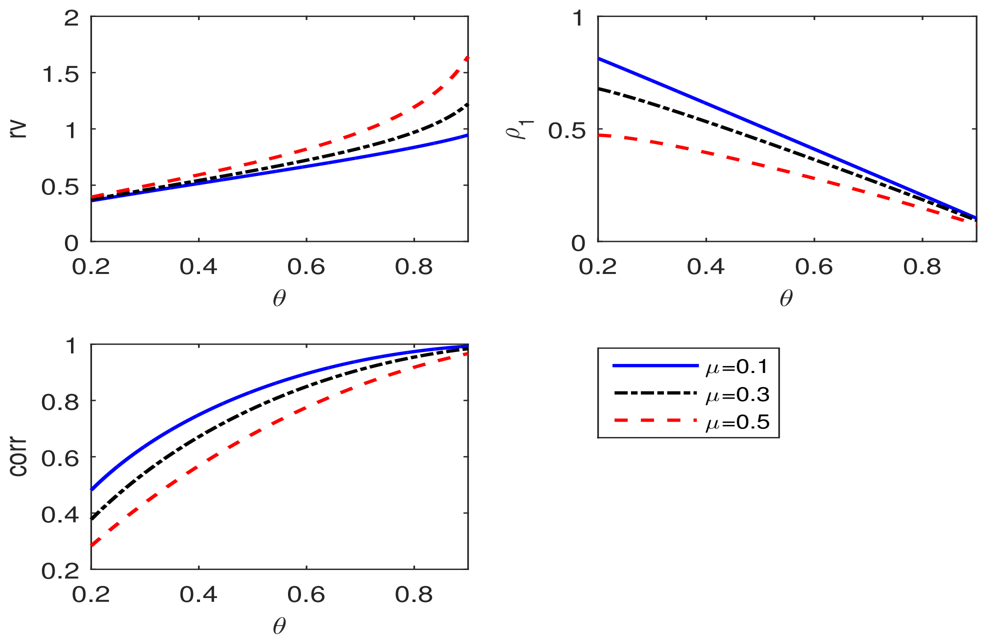

Figure 6: Stochastic Properties of Consumption and Income

Table 1: The BKK puzzle

	$\text{corr}(\Delta y, \Delta y^*)$	$\text{corr}(\Delta c, \Delta c^*)$
Canada	0.83	0.56
France	0.51	0.33
UK	0.75	0.70
Italy	0.44	0.11
Canada-US	0.87	0.58

* The numbers $\text{corr}(\Delta y, \Delta y^*)$ and $\text{corr}(\Delta c, \Delta c^*)$ are the simple correlation coefficients between the annual change of a country's real output (or consumption) and the annual change of the rest of the world's real output (or consumption), with the "world" defined as the output-weighted average of the rest G7 countries in the Penn World Table (version 7.1). Canada-US correlations are between Canada and the U.S..

Table 2: Summary of statistics

	corr $(\Delta y, \Delta y^*)$	corr $(\Delta c, \Delta c^*)$	$sd(\Delta c)/sd(\Delta y)$	$corr(\Delta c, \Delta y)$	$autocorr(\Delta c)$
Canada	0.83	0.56	0.52	0.75	0.60
	(0.05)	(0.08)	(0.07)	(0.07)	(0.08)
France	0.51	0.33	0.45	0.76	0.49
	(0.17)	(0.13)	(0.07)	(0.06)	(0.11)
UK	0.75	0.70	0.73	0.93	0.56
	(0.08)	(0.07)	(0.05)	(0.02)	(0.15)
Italy	0.44	0.11	0.48	0.75	0.46
	(0.20)	(0.16)	(0.10)	(0.03)	(0.10)
Canada-US	0.87	0.58			
	(0.04)	(0.08)			
US			0.57	0.91	0.64
			(0.05)	(0.02)	(0.09)

Table 3: Theoretical corr $(\Delta c, \Delta c^*)$ from different models

	Data	RE ($\theta^* = 1$)	RI ($\theta^* = 0.9$)	RI ($\theta^* = 0.8$)	RI ($\theta^* = 0.7$)	RI ($\theta^* = 0.6$)	RI ($\theta^* = 0.5$)
Canada							
($\mu = 0$)	0.56	0.83	0.66	0.62	0.60	0.59	0.59
($\mu = 0.1$)	0.56	0.83	0.65	0.59	0.57	0.55	0.54
($\mu = 0.3$)	0.56	0.83	0.56	0.47	0.42	0.37	0.33
($\mu = 0.9$)	0.56	0.83	0.29	0.20	0.15	0.12	0.09
(μ=pooled C1/100)	0.56	0.83	0.56	0.47	0.42	0.37	0.33
France							
($\mu = 0$)	0.33	0.51	0.43	0.41	0.40	0.39	0.39
($\mu = 0.1$)	0.33	0.51	0.43	0.40	0.38	0.37	0.36
($\mu = 0.3$)	0.33	0.51	0.38	0.33	0.30	0.27	0.24
($\mu = 0.9$)	0.33	0.51	0.22	0.15	0.11	0.09	0.07
(μ=pooled C1/100)	0.33	0.51	0.38	0.33	0.29	0.26	0.23
UK							
($\mu = 0$)	0.70	0.75	0.70	0.67	0.65	0.64	0.64
($\mu = 0.1$)	0.70	0.75	0.69	0.65	0.63	0.62	0.61
($\mu = 0.3$)	0.70	0.75	0.64	0.58	0.53	0.48	0.43
($\mu = 0.9$)	0.70	0.75	0.41	0.30	0.23	0.17	0.13
(μ=pooled C1/100)	0.70	0.75	0.69	0.65	0.63	0.62	0.60
Italy							
($\mu = 0$)	0.11	0.44	0.39	0.37	0.36	0.36	0.35
($\mu = 0.1$)	0.11	0.44	0.38	0.36	0.35	0.34	0.33
($\mu = 0.3$)	0.11	0.44	0.35	0.31	0.28	0.25	0.23
($\mu = 0.9$)	0.11	0.44	0.21	0.15	0.11	0.09	0.07
(μ=pooled C1/100)	0.11	0.44	0.32	0.27	0.23	0.20	0.17
Canada-US							
($\mu = 0$)	0.58	0.87	0.61	0.56	0.55	0.54	0.55
($\mu = 0.1$)	0.58	0.87	0.58	0.53	0.50	0.49	0.48
($\mu = 0.3$)	0.58	0.87	0.47	0.39	0.33	0.29	0.26
($\mu = 0.9$)	0.58	0.87	0.22	0.15	0.11	0.08	0.06
(μ=pooled C1/100)	0.58	0.87	0.47	0.39	0.33	0.29	0.26

Table 4: Calibrated θ

	$\theta^* = 0.9$	$\theta^* = 0.8$	$\theta^* = 0.7$	$\theta^* = 0.6$	$\theta^* = 0.5$
Canada	0.36	0.26	0.21	0.17	0.14
France	0.43	0.31	0.25	0.20	0.17
UK	0.56	0.42	0.33	0.27	0.21
Italy	0.49	0.36	0.28	0.23	0.19
Canada-US	0.26	0.19	0.15	0.13	0.11

Table 5: Other consumption moments: $sd(\Delta c)/sd(\Delta y)$, $autocorr(\Delta c)$ and $corr(\Delta c, \Delta y)$

	Data	RE	RI($\mu = 0$)		RI($\mu = 0.1$)		RI($\mu =$pooled C1/100)	
			$\theta^* = 0.9$	$\theta^* = 0.8$	$\theta^* = 0.9$	$\theta^* = 0.8$	$\theta^* = 0.9$	$\theta^* = 0.8$
Canada								
$sd(\Delta c)/sd(\Delta y)$	0.52	1	0.48	0.41	0.49	0.42	0.56	0.52
$autocorr(\Delta c)$	0.60	0	0.67	0.77	0.63	0.71	0.44	0.43
$corr(\Delta c, \Delta y)$	0.75	1	0.75	0.64	0.73	0.62	0.64	0.50
France								
$sd(\Delta c)/sd(\Delta y)$	0.45	1	0.54	0.45	0.54	0.46	0.61	0.55
$autocorr(\Delta c)$	0.49	0	0.59	0.71	0.57	0.67	0.42	0.44
$corr(\Delta c, \Delta y)$	0.76	1	0.81	0.70	0.80	0.68	0.71	0.58
UK								
$sd(\Delta c)/sd(\Delta y)$	0.73	1	0.63	0.52	0.63	0.53	0.63	0.53
$autocorr(\Delta c)$	0.56	0	0.46	0.61	0.45	0.58	0.45	0.58
$corr(\Delta c, \Delta y)$	0.93	1	0.89	0.79	0.88	0.78	0.88	0.78
Italy								
$sd(\Delta c)/sd(\Delta y)$	0.48	1	0.58	0.48	0.58	0.49	0.68	0.63
$autocorr(\Delta c)$	0.46	0	0.53	0.67	0.52	0.64	0.31	0.32
$corr(\Delta c, \Delta y)$	0.75	1	0.85	0.74	0.84	0.73	0.72	0.57
Canada (as in Canada-US)								
$sd(\Delta c)/sd(\Delta y)$	0.52	1	0.41	0.35	0.42	0.37	0.52	0.50
$autocorr(\Delta c)$	0.60	0	0.77	0.84	0.71	0.75	0.43	0.38
$corr(\Delta c, \Delta y)$	0.75	1	0.64	0.54	0.62	0.51	0.50	0.38