

HONG KONG INSTITUTE FOR MONETARY RESEARCH

**ELASTIC ATTENTION, RISK SHARING, AND
INTERNATIONAL COMOVEMENTS**

Wei Li, Yulei Luo and Jun Nie

HKIMR Working Paper No.26/2015

December 2015



Hong Kong Institute for Monetary Research

香港金融研究中心

(a company incorporated with limited liability)

All rights reserved.

Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.

Elastic Attention, Risk Sharing, and International Comovements^{*}

Wei Li^{}**

University of Hong Kong

and

Yulei Luo^{*}**

University of Hong Kong

and

Jun Nie^{**}**

Federal Reserve Bank of Kansas City

December 2015

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 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 also improve the model's predictions for other key moments of the joint dynamics of consumption and income. Finally, we show that the main conclusions of our benchmark model do not change in an extension with capital accumulation.

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

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

^{*} We are grateful to R. Anton Braun, Hongyi Chen, Kaiji Chen, Martin Ellison, 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, and Federal Reserve Bank of Atlanta 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.

^{**} Corresponding author. Faculty of Business and Economics, The University of Hong Kong, Hong Kong. E-mail: verawli@connect.hku.hk.

^{***} 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 of the stylized facts in open economies. One of the major inconsistencies between the models' predictions and the empirical evidence concerns cross-country consumption and income correlations. Specifically, in a canonical open economy model under a 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 it is 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 optimal (or elastic) attention into an otherwise standard small open economy (SOE) model. Specifically, we follow Sims (2010) and assume that 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. They thus optimally choose information-processing capacity and make decisions based on perceived information.³ After solving our benchmark model explicitly, we show that this elastic attention mechanism can help endogenously generate heterogeneous and gradual consumption adjustments to income shocks across countries, and thus make the model fit the data better in explaining international consumption and income correlations as well as some other key stochastic properties of the joint dynamics of consumption and income in individual countries.

By inspecting the mechanism of our benchmark model, we find that there are three competing forces that determine the dynamics of consumption and income in our benchmark model. The first channel is the slow adjustment channel. Specifically, if the home country and the rest of the world have the same degree of slow adjustment, imperfect state observations generate gradual responses of consumption

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

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

³The assumption is also consistent with a psychological theory on elastic attention proposed in Kahneman (1973).

growth to income shocks, and the channel has no impact on cross-country correlations because its impacts on consumption variance and cross-country consumption covariance are cancelled out. The second channel is the common noise channel. The common noise from imperfect observations reduces consumption correlations across countries because it increases consumption volatility while having no effect on the covariance of consumption across countries. The third channel is the elastic attention channel. Consumers facing fixed information-processing costs optimally choose their information-processing capacity and thus the speed of adjustment to consumption. The gap between heterogeneous responses of consumption to income has the potential to lower cross-country consumption correlations. It is worth noting that this 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 underlying mechanism that generates greater heterogeneity and thus lower cross-country consumption correlations.

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.⁴ This paper considers an important application of RI and shows that it can help to better understand international comovements. Our paper is closely related to Luo, Nie, and Young (2014). They examine the effects of model uncertainty due to a concern over model misspecification (robustness) on international consumption correlations in a SOE real business cycles (RBC) model. In their economy, agents make optimal decisions while bearing in mind that the model is misspecified in some unknown way. 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

⁴See Luo (2008) and Mac'kowiak and Wiederholt (2009) for an application of RI to consumption and firm decisions within linear-quadratic-Gaussian setting.

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.

There are other extensions 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 high equity premium and volatility of the exchange rate, consumption correlations in their model are slightly higher than the output correlations. Another competing 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 consumption correlations 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 (FI-RE) 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 main findings about how elastic attention improves the model's performance on the other key stochastic properties of the joint dynamics of consumption and income. Section 5 discusses an extension with endogenous capital accumulation.

Section 6 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 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

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

subject to the flow budget constraint

$$b_{t+1} = Rb_t - c_t + y_t, \quad (2)$$

where $u(c_t) = -(\bar{c} - c_t)^2/2$ is the utility function,⁵ 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 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 0.⁶ Here we assume that the household sector takes y_t as given to keep our model tractable. Incorporating the firm sector and modelling investment decisions explicitly does not change the main results in this paper. 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

⁵Our main results do not change if we adopt a constant-absolute-risk-aversion (CARA) utility function, $u(c_t) = -\exp(-\alpha c_t)/\alpha$, where $\alpha > 0$ is the coefficient of absolute risk aversion. This main reason for this result is that the stochastic property of the joint dynamics of consumption and income are mainly determined by the marginal propensity to consume 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.

⁶Here we ignore the investment and government spending components.

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:

$$c_t = (R - 1)s_t, \quad (3)$$

where

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

is the expected present value of lifetime resources, 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 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 permanent income s_t that can be solved analytically.⁷ Letting s_t be defined as a new state variable, we can reformulate the SOE model as

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

subject to

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

where the time $(t + 1)$ innovation to permanent income can be written as

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

⁷See 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.

$v(s_0)$ 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):

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

which relates changes in consumption to income innovations.⁸ 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:

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

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 a constant-relative-risk-averse (CRRA) utility (e.g., Kollman 1996) for two reasons. First, most existing RI models assume that the

⁸Under FI-RE the expression for the change in individual consumption is the same as that for the change in aggregate consumption.

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 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 (it is not generally a member of a known class of distributions, making it difficult to characterize with a small number of moments).⁹ Second, there is no two-country general equilibrium in which the general equilibrium interest rate is constant.¹⁰ Specifically, consider a simple FI-RE two-country general equilibrium model in which the home country's budget constraint and consumption decision are characterized by (2) and (3), respectively, and the agents in the foreign country solve the same problem in which its variables are denoted with an asterisk.¹¹ In general equilibrium, the bond market-clearing condition is

$$b_t + b_t^* = 0 \text{ for all } t. \quad (8)$$

By Walras' law, (8) implies that the global resource constraint should also hold for all t :

$$c_t + c_t^* = y_t + y_t^* \equiv y_t^w, \quad (9)$$

where y_t^w denotes exogenously given current world output. Using the expected resource constraint,

$$E_{t-1}[y_t^w] = E_{t-1}[c_t] + E_{t-1}[c_t^*] = \frac{1}{R} \left(\frac{1}{\beta} c_{t-1} + \frac{1}{\beta^*} c_{t-1}^* \right),$$

we can easily obtain the expression for the general equilibrium interest rate:

$$R = \frac{1}{E_{t-1}[y_t^w]} \left(\frac{1}{\beta} c_{t-1} + \frac{1}{\beta^*} c_{t-1}^* \right). \quad (10)$$

⁹Note that within the LQ setting, the first two moments are sufficient to characterize the distribution of the state.

¹⁰Note 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 whole distribution of the true state.

¹¹Note that here we relax the assumption that $\beta R = 1$ so that β can be different in the two countries.

However, given that $c_{t-1} = (R - 1)s_{t-1}$ and $c_{t-1}^* = (R - 1)s_{t-1}^*$, the right-hand side of (10) is a time- $(t - 1)$ random variable, i.e., there is no constant R such that (10) holds. It is obvious that this argument also holds for the RI model as both c_{t-1} and c_{t-1}^* are random variables at $t - 1$.

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:

$$y_t = y_t^p + y_t^i, \quad (11)$$

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

$$y_t^i = \bar{y} + \varepsilon_t, \quad (13)$$

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:

$$y_t^* = y_t^{p*} + y_t^{i*}, \quad (14)$$

$$y_t^{p*} = y_{t-1}^{p*} + \varepsilon_t^*, \quad (15)$$

$$y_t^{i*} = \bar{y}^* + \varepsilon_t^*, \quad (16)$$

where ε_t^* and ε_t^* are orthogonal iid shocks with mean 0 and variance ω^{*2} and $\omega_{\varepsilon^*}^2$, respectively.

In addition, we allow contemporaneous correlations between the SOE and the ROW,

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

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

$$* \text{corr}(\Delta y_t, \Delta y_t^*) = \frac{E[\varepsilon_t \varepsilon_t^*] + 2E[\varepsilon_t \varepsilon_t^*]}{\sqrt{\omega^2 + 2\omega \omega_\varepsilon^2} \sqrt{\omega^{*2} + 2\omega \omega_\varepsilon^{*2}}} \quad (17)$$

Using annual GDP data from 1950 – 2010 from the Penn World Tables (version 7.1), we find that income volatility is mainly due to permanent shocks in the G-7 countries. For example, $\omega^2 = 186523.2^2$ and $\omega_\varepsilon^2 = 2.5^2$ for the U.S. That is, $\omega \gg \omega_\varepsilon$. This empirical result is consistent with the literature 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 consumption. He estimates that the transitory income component accounts for 1% to 2% of total variance of consumption. Luo et al. (2015) find that the $\omega_\varepsilon/\omega$ 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.¹² Therefore, we focus on the permanent income component for the rest of our discussion. The cross country income correlation can be approximated as follows:

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

Under the income specification, the innovation to life-time wealth can be reduced to $\zeta_t = \varepsilon_t/(R-1) + \varepsilon_t/R: N(0, \omega_\zeta^2)$ which is approximately equal to $\varepsilon_t/(R-1)$ 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

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

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

¹²For 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.

worth noting that this consumption behavior does not fit the data well. As is 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 show that introducing elastic attention can help generate more realistic impulse responses of consumption to income.

Given that consumption dynamics in the ROW is

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

the international consumption correlation can thus be written as

$$* \text{corr}(\Delta c_t, \Delta c_t^*) \approx * \text{corr}(\varepsilon_t, \varepsilon_t^*) \approx * \text{corr}(\Delta y_t, \Delta y_t^*). \quad (18)$$

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

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.¹³ With finite capacity $\kappa \in (0, \infty)$, the state

¹³Formally, entropy is defined as the expectation of the negative of the (natural) log of the density function, $-E[\ln(f(s))]$. The entropy of a discrete distribution with equal weight on two points is simply $E[\ln(f(s))] = -0.5\ln(0.5) - 0.5\ln(0.5) = 0.69$, 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

variable s following a continuous distribution cannot be observed without error and thus the information set at time $t + 1$, \mathcal{J}_{t+1} , is generated by the entire history of noisy signals $\{s_j^*\}_{j=0}^{t+1}$. Agents with finite capacity will choose a new signal $s_{t+1}^* \in \mathcal{J}_{t+1} = \{s_1^*, s_2^*, \dots, s_{t+1}^*\}$ 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

$$\mathcal{H}(s_{t+1}|\mathcal{J}_t) - \mathcal{H}(s_{t+1}|\mathcal{J}_{t+1}) \leq \kappa, \quad (19)$$

where κ is the typical consumer's information channel capacity, $\mathcal{H}(s_{t+1}|\mathcal{J}_t)$ denotes the entropy of the state prior to observing the new signal at $t + 1$, and $\mathcal{H}(s_{t+1}|\mathcal{J}_{t+1})$ 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, $s_t|\mathcal{J}_t: N(E[s_t|\mathcal{J}_t], \Sigma_t)$, where $\Sigma_t = E_t[(s_t - \hat{s}_t)^2]$, is optimal.¹⁴ In addition, Mac'kowiak and Wiederholt (2009) also show 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. Specifically, within the LQG setting, the information-processing constraint, (19), can be reduced to

$$\ln(R^2\Sigma_t + \omega_\xi^2) - \ln(\Sigma_{t+1}) \leq 2\kappa; \quad (20)$$

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

$$s_t^* = s_t + \xi_t, \quad (21)$$

capacity devoted to monitoring s is $\kappa = 0.69$ nats.

¹⁴This result is often assumed as a matter of convenience in signal extraction models with exogenous noises, and RI can rationalize this assumption.

where $\xi_t: N(0, \Lambda)$, 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): $\Lambda = (\Sigma^{-1} - \Psi^{-1})^{-1}$, where $\Psi = R^2\Sigma + \omega_\xi^2$ is the posterior variance of the state. Note that the specification in (21) is standard in the signal extraction literature and captures the situation where agents happen or choose to have imperfect knowledge of the underlying shocks.¹⁵ Since imperfect observations of the state lead to welfare losses, agents use the processed information to estimate the true state.¹⁶ Specifically, we assume that households use the Kalman filter to update the perceived state $\hat{s}_t = E_t[s_t]$ after observing new signals in the steady state in which the conditional variance of s_t , $\Sigma_t = \text{var}_t(s_t)$, has converged to a constant Σ :

$$\hat{s}_{t+1} = (1 - \theta)(R\hat{s}_t - c_t) + \theta(s_{t+1} + \xi_{t+1}), \quad (22)$$

where $\theta = 1 - \exp(-2\kappa)$ is the Kalman gain (i.e., the observation weight).¹⁷

Combining (5) with (22), we obtain the following equation governing the perceived state \hat{s}_t :

$$\hat{s}_{t+1} = R\hat{s}_t - c_t + \eta_{t+1}, \quad (23)$$

where

$$\eta_{t+1} = \theta R(s_t - \hat{s}_t) + \theta(\zeta_{t+1} + \xi_{t+1}) \quad (24)$$

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

$$s_t - \hat{s}_t = \frac{(1-\theta)\zeta_t}{1-(1-\theta)R \cdot L} - \frac{\theta\xi_t}{1-(1-\theta)R \cdot L} \quad (25)$$

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

¹⁵Note 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 et al. (2014) for a recent application.

¹⁶See Luo (2008) and Luo et al. (2015) for details about the welfare losses due to information imperfections within a partial equilibrium permanent income hypothesis framework.

¹⁷ θ measures how much uncertainty about the state can be removed upon receiving new signals about the state.

$$\eta_{t+1} = \theta \left[\left(\frac{\zeta_{t+1}}{1-(1-\theta)R \cdot L} \right) + \left(\xi_{t+1} - \frac{\theta R \xi_t}{1-(1-\theta)R \cdot L} \right) \right], \quad (26)$$

where $\omega_{\xi}^2 = * var(\xi_{t+1}) = \frac{1}{\theta} \frac{1}{1-(1-\theta)-R^2} \omega_{\zeta}^2$. Expression (26) 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 (26) 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.¹⁸

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:

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

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., $\lambda \rightarrow 0$, there is no informational friction as in the FI-RE model, $\Sigma = 0$; on the contrary, when $\lambda \rightarrow \infty$, $\Sigma \rightarrow \infty$, i.e., no information will be collected. The optimal steady state conditional variance can be solved as

$$\Sigma = \frac{-(1-R(R-1)\tilde{\lambda}) + \sqrt{(1-R(R-1)\tilde{\lambda})^2 + 4R^2\tilde{\lambda}}}{2R^2} \omega_{\zeta}^2, \quad (27)$$

where $\tilde{\lambda} \equiv \lambda/\omega_{\zeta}^2$. \hat{s}_t is governed by the Kalman filtering equation

$$\hat{s}_{t+1} = (1-\theta)(R\hat{s}_t - c_t) + \theta(s_{t+1} + \xi_{t+1}). \quad (28)$$

¹⁸Note that when $\theta = 1$, $* var(\xi_{t+1}) = 0$.

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

$$\theta = 1 - \frac{1}{R^2} \left\{ 1 + \frac{2}{-[1-R(R-1)\tilde{\lambda}] + \sqrt{[1-R(R-1)\tilde{\lambda}]^2 + 4R^2\tilde{\lambda}}} \right\}^{-1}. \quad (29)$$

Figure 1 clearly shows that the value of θ increases with the level of macroeconomic uncertainty measured by ω_ξ^2 (i.e., $\partial\theta/\partial\omega_\xi^2 > 0$). 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).

Under RI, optimal consumption is

$$c_t = (R - 1)\hat{s}_t, \quad (30)$$

and the change in consumption can be expressed as

$$\Delta c_t^{RI} = \theta(R - 1) \left[\frac{\zeta_t}{1 - (1 - \theta)R \cdot L} + \left(\xi_t - \frac{\theta R \xi_{t-1}}{1 - (1 - \theta)R \cdot L} \right) \right]. \quad (31)$$

Similarly, in the ROW, we have

$$\Delta c_t^{*RI} = \theta^*(R - 1) \left[\frac{\zeta_t^*}{1 - (1 - \theta^*)R \cdot L} + \left(\xi_t^* - \frac{\theta^* R \xi_{t-1}^*}{1 - (1 - \theta^*)R \cdot L} \right) \right]. \quad (32)$$

From these expressions, it is clear that consumption adjusts gradually to income shocks instead of fully adjusting immediately. When $\theta < 1$, the true state is no longer observable due to the existence of the consumer's endogenous information noises ξ . Through gradual learning and adjustment, inattention opens up for past income shocks and information noises to affect the current consumption decision. As the consumer pays less attention (smaller θ), these shocks become more important. When $\theta = 1$, the

true state can be observed and past shocks are not informative, hence the above expression reduces to $\Delta c_t = (R - 1)\zeta_t$, which is the same as in the full information model. For different countries, their domestic fundamental uncertainty can affect optimal consumption decisions through θ and its interaction with ξ . Different volatility of income shocks leads to different levels of θ . The higher θ is, the more new information a country processes. Different levels of θ then lead to different adjustments of consumption. This helps to explain why consumption correlation is in general lower than 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.¹⁹

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 (56) yields the expression of the change in aggregate consumption:

$$\Delta c_t = (R - 1) \left[\frac{\theta \zeta_t}{1 - (1 - \theta)R \cdot L} + \theta \left(\bar{\xi}_t - \frac{\theta R \bar{\xi}_{t-1}}{1 - (1 - \theta)R \cdot L} \right) \right], \quad (33)$$

where i denotes a particular individual, $E^i[\cdot]$ is the population average, and $\bar{\xi}_t = E^i[\xi_t]$ is the common noise.

Assume that ξ_t consists of two independent noises: $\xi_t = \bar{\xi}_t + \xi_t^i$, where $\bar{\xi}_t = E^i[\xi_t]$ and ξ_t^i are the common and idiosyncratic components of the error generated by ζ_t , respectively. Define a single

¹⁹Of course, to rationalize the RI hypothesis, we need to assume that the optimality of ex post Gaussianity holds approximately when the utility function is negative exponential. The detailed proof of this result is available from the corresponding author by request.

parameter,

$$\mu \equiv \frac{*sd(\bar{\xi}_t)}{*sd(\xi_t)} \in [0,1],$$

to measure the common source of coded information on the aggregate component (or the relative importance of $\bar{\xi}_t$ vs. ξ_t).²⁰ Idiosyncratic noises are cancelled out after aggregation while the common noise remains.²¹

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

$$*corr(\Delta c, \Delta c^*) \approx \Pi *corr(\Delta y, \Delta y^*), \quad (34)$$

where

$$\begin{aligned} \Pi &= \frac{1}{[1-(1-\theta)(1-\theta^*)R^2]\sigma(\theta, \mu)\sigma(\theta^*, \mu)}, \\ \sigma(\theta, \mu) &= \sqrt{\frac{1}{1-[(1-\theta)R]^2} + \mu^2 \left\{ \frac{1}{(1-(1-\theta)R^2)\theta} - \frac{1}{1-[(1-\theta)R]^2} \right\}}, \\ \sigma(\theta^*, \mu) &= \sqrt{\frac{1}{1-[(1-\theta^*)R]^2} + \mu^2 \left\{ \frac{1}{(1-(1-\theta^*)R^2)\theta^*} - \frac{1}{1-[(1-\theta^*)R]^2} \right\}}. \end{aligned} \quad (35)$$

Proof. See Appendix 7.1.

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:

²⁰In 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.

²¹Black (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.

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 (35) 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 of consumption correlation 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. Habit formation has been modelled directly as a structure of preferences in which psychological factors make consumers prefer gradual adjustment in consumption; consequently, consumption volatility is more painful than it would be in the absence of habits. The key difference between habit formation and RI without noises is that slow adjustment in consumption under habit formation is optimal because consumers are assumed to prefer to smooth not only consumption but also consumption growth, while slow consumption adjustment under RI is optimal because capacity constraints make consumers take more time to acquire and process information. Therefore, habit formation by itself cannot help resolve the cross-country consumption correlation puzzle. This is consistent with the conclusion obtained in a habit formation model proposed by Fuhrer and Klein (2006).

The slow adjustment mechanism can also be generated by assuming that consumers cannot

distinguish the two components of income specified in (11)-(13).²² Specifically, following Muth (1960) and Pischke (1991), given that the change in income is

$$\Delta y_{t+1} = \varepsilon_{t+1} + \varepsilon_{t+1} - \varepsilon_t, \quad (36)$$

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

$$\Delta y_{t+1} = v_{t+1} - \alpha v_t, \quad (37)$$

where the innovation, v_t , with mean 0 and variance ω_v^2 , is not a fundamental driving process -- it contains information on current and lagged permanent and transitory income shocks. Equating the variances and autocorrelation coefficients of the original and derived processes, (36) and (37), we have

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

where $\rho = -\omega_\varepsilon^2 / (\omega^2 + 2\omega_\varepsilon^2)$ and $\alpha \in [0, 1]$ will be large if the variance of the transitory shock ω_ε^2 is large relative to the variance of the permanent shock ω^2 and will converge to 0 as ω_ε^2 approaches 0. Following the same procedure in Section 2.1, we can solve for the expression for the change in aggregate consumption as follows:

$$\Delta c_t = \frac{R - \alpha}{R} \frac{\varepsilon_{t+1}}{1 - \alpha \cdot L}. \quad (38)$$

where the slow adjustment mechanism is captured by the factor $1/(1 - \alpha \cdot L)$. Under incomplete information, the presence of the transitory shock plays a role in strengthening the inertial responses to the aggregate income shock because α is a function of the variance of the transitory shock. If α is a large value, the effect will be initially small but highly persistent. However, given that $\omega^2 \omega_\varepsilon^2$ in our estimation using U.S. data, we can easily calculate that α is close to 0. In other words, given the estimated income process, the propagation mechanism in the IC model is extremely weak, and the expression for the changes in consumption is almost identical to the one we obtained in our benchmark

²²Boz, Daude, and Durdu (2011) incorporate this type of incomplete information into a SOE-RBC model and examine how it affects business cycle dynamics in emerging markets.

model.

The inattentiveness and infrequent adjustment model is proposed in Reis (2006). Specifically, Reis (2006) assumes that during the intervals of inattentiveness, consumption dynamics are determined by the standard determinant consumer's optimizing problem and consumption is determined by the standard stochastic consumer problem at the adjustment dates. Reis then finds that aggregate consumption growth between two consecutive periods, t and $t + 1$, in the model economy can be written as

$$\Delta c_{t+1} = \text{constant} + \Psi(0)e_{t+1} + \Psi(1)e_t + \dots + \Psi(I)e_{t+1-I}, \quad (39)$$

where $\Psi(s) \geq \Psi(s + 1) \geq 0$ for $s = 1, 2, \dots, I$, and $\{e_t\}$ are mutually uncorrelated "news" unpredictable one period ahead. Expression (39) reveals that aggregate consumption exhibits slow adjustment because "news" diffuses across all individuals slowly. This conclusion is therefore also consistent with that obtained in our RI model without noises.

For the common noise channel, Expression (34) 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 4 and 5 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 is different by nature, which leads to different levels of attention. Figure 3 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

$$\Delta c_t^{RI} = \theta(R - 1) \frac{\zeta_t}{1 - (1 - \theta)R}, \quad (40)$$

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

$$\Delta c_t^{*RI} = \theta^*(R - 1) \frac{\zeta_t^*}{1 - (1 - \theta^*)R}. \quad (41)$$

The consumption correlation between the two countries becomes

$$* \text{corr}(\Delta c_t^{RI}, \Delta c_t^{*RI}) \approx \bar{\Pi} * \text{corr}(\Delta y_t, \Delta y_t^*), \quad (42)$$

where

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

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 $\theta = \theta^*$, $\bar{\Pi} = 1$, 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 $\Pi \in [\underline{\Pi}, \bar{\Pi}]$, where

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

We proceed to vary the two parameters, μ and θ^* , to show the implications from different models quantitatively. 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 , $* \text{corr}(\Delta c, \Delta c^*)$ is increasing in θ^* , which is the same as in the representative agent model ($\mu = 1$). As we can see from Table 3, our model

fits the data better for many combinations of the two parameter values. For example, for France, when $\theta = 0.8$ and $\mu = 0.8$ (i.e., 80% of uncertainty is removed upon new signals and 80% of the noise information is remained after aggregation), the RI model predicts that $*corr(\Delta c, \Delta c^*) = 0.42$, which is very close to its empirical counterpart, 0.41.

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:*

$$rv = \frac{*sd(\Delta c)}{*sd(\Delta y)} = \sqrt{\frac{\theta^2}{1-[(1-\theta)R]^2} + \mu^2 \left\{ \frac{\theta}{1-(1-\theta)R^2} - \frac{\theta^2}{1-[(1-\theta)R]^2} \right\}}, \quad (43)$$

the first-order autocorrelation of consumption change is

$$\rho_c(1) = \frac{(1-\mu^2)(1-\theta)R}{1+\mu^2 \left[\frac{1-(1-\theta)^2 R^2}{(1-(1-\theta)R^2)\theta} - 1 \right]}, \quad (44)$$

and the contemporaneous correlation between consumption change and income change is

$$*corr(\Delta c, \Delta y) = \frac{1}{\sqrt{\frac{1}{1-[(1-\theta)R]^2} + \mu^2 \left\{ \frac{1}{(1-(1-\theta)R^2)\theta} - \frac{1}{1-[(1-\theta)R]^2} \right\}}}. \quad (45)$$

Proof. See Appendix 7.2.

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 7, for most of

the combinations, our model fits the data better. Compared to the benchmark model, we 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 (the presence of ξ_t) dominates the slow propagation channel (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 , $rv \in$

$\left[\sqrt{\frac{\theta^2}{1-[(1-\theta)R]^2}}, \sqrt{\frac{\theta}{1-(1-\theta)R^2}} \right]$. 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 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 the 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 a 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 the ROW to Canada since over 70% of Canada's international trade is with the U.S. (Miyamoto and Nguyen, 2014). We will discuss more about this pair of countries in Section 4.3.

Table 1 illustrates 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 that we construct and the U.S. economy, $sd(\Delta c^*)/sd(\Delta y^*) \in [0.57, 0.58]$ and $corr(\Delta c^*, \Delta y^*) \in [0.91, 0.93]$. If we assume there is no common noise in the ROW, our benchmark model can match $sd(\Delta c^*)/sd(\Delta y^*)$ when $\theta^* \in [0.48, 0.50]$ and match $corr(\Delta c^*, \Delta y^*)$ 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.²³ 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) finds $\theta = 0.4$ based on the response of aggregate output to monetary policy shocks. Luo (2008) finds that if $\theta = 0.5$, the otherwise standard permanent income model generates a realistic relative volatility of consumption to labor income. Mac'kowiak and Wiederholt (2009) find that given a total information flow of 133 bits, the decision-maker of the typical firm only allocates 0.76 bits of information flow to tracking aggregate technology and 0.41 bits 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

²³ 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.

total resources hit by the innovation to total resources. 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 $\tilde{\lambda}^*$ can be recovered by solving equation (29). Given that $\tilde{\lambda}^* \equiv \lambda^*/\omega_\zeta^{*2}$ and that both domestic country and ROW consumers are facing a stable information cost, i.e., $\lambda = \lambda^*$, we have

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

Plugging the expression for $\tilde{\lambda}$ back into Equation (29), we can derive θ . Now we are ready to calculate Ξ and then use Equation (34) to determine $*corr(\Delta c_t^{RI}, \Delta c_t^{*RI})$. Following Glick and Rogoff (1995), the interest rate (R) and the depreciation rate (δ) are set to be 1.04 and 0.05, respectively.

4.3 Main Results

Table 2 reports a summary of statistics. The values of $*corr(\Delta y, \Delta y^*)$ (and $*corr(\Delta c, \Delta c^*)$) are the simple correlation coefficients between the annual change in country's real output (and consumption) and the annual change in the rest of the world's real output (and consumption), with the "world" defined as an 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 introducing elastic attention can 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 $\mu = 0.3$. 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 $\theta^* = 0.5$. 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. Table 5 shows the results obtained from the model with the elastic attention channel alone when the common noise channel is muted. It shows 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 correlations. Table 6 shows that these results are robust when we use another version of the Penn World Tables data (PWT8.0).²⁴

Table 7 compares the other key stochastic properties of the joint dynamics of consumption and income: the relative volatility of consumption to income, the first-order autocorrelation of consumption, and the contemporaneous consumption-income correlations 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 countries that we study. 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.

4.4 The Canada-US Case

²⁴For real GDP, we use RGDPna, GDP using national-accounts growth rates.

The Canada-US case is of interest because Canada and the U.S. have one of the world's closest bilateral relationships. Their 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 literature. We now apply our elastic attention model to study US and Canada 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 since 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 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 larger.

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 that we discussed in our theoretical model. Specifically, one is due to information noise that is endogenously induced by inattention. For the Canadian and US people, their information noises have zero covariance but the presence of common noise increases the variance of their own consumption innovations, and therefore decreases the correlation. The other channel is due to different levels of fundamental uncertainty in the two countries. Since the variance of the annual change of output in the U.S. is much larger than that for Canada, the attention level in different countries would be chosen to be different. If we assume that 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. Extension: Capital Accumulation and Endogenous Net Output

5.1 Introducing Capital Accumulation

In this section, we discuss how elastic attention affects cross-country consumption correlations in a SOE model when we consider endogenous capital accumulation. Specifically, we follow Glick and Rogoff (1995), Gruber (2002) and Luo et al. (2014) to model the firm sector, and assume that the production function is

$$y_t = a_t k_t^\alpha - \frac{g}{2} \frac{i_t^2}{k_t}, \quad (46)$$

where k_t is the capital stock, i_t is gross investment, $\frac{g}{2} \frac{i_t^2}{k_t}$ measures the loss of output due to adjustment costs ($g > 0$ is a constant), and a_t is a multiplicative country-specific productivity shock that follows

$$a_{t+1} = (1 - \rho)\bar{a} + \rho a_t + \varepsilon_{t+1}, \quad (47)$$

where $\rho \in [0,1]$ is the persistence coefficient, \bar{a} is the mean of the country-specific productivity shock, and ε_{t+1} is an *iid* Gaussian innovation with mean 0 and variance ω^2 . For simplicity, we assume that the firm has perfect state observations.

The objective of the firm is to choose capital and investment to maximize the following profit function

$$\max \sum_{t=j}^{\infty} \left(\frac{1}{R}\right)^{t-j} \left(a_t k_t^\alpha - \frac{g}{2} \frac{i_t^2}{k_t} - i_t \right). \quad (48)$$

subject to the capital accumulation equation

$$k_{t+1} = (1 - \delta)k_t + i_t, \quad (49)$$

for $t \geq j$. Following the same procedure used in Glick and Rogoff (1995), we can solve for the optimal capital accumulation and investment rules as follows:

$$k_t = \lambda_1 k_{t-1} + \frac{\alpha \bar{k}^{-\alpha}}{g \lambda_2} \sum_{j=t}^{\infty} \left(\frac{1}{\lambda_2}\right)^{j-t} E_{t-1}[a_j] + \Omega, \quad (50)$$

$$i_t = \lambda_1 i_{t-1} + \lambda_i \Delta a_t, \quad (51)$$

where $\Omega = \frac{\alpha \bar{k}^{-\alpha}}{g(1-\lambda_2)L}$ is an irrelevant constant term, L is the lag operator, $\lambda_i = \frac{\rho \alpha \bar{k}^{-\alpha}}{g \lambda_2 (1-\rho)}$, and the eigenvalues, $\lambda_1 \in (0,1)$ and $\lambda_2 > 1$, satisfy $\lambda_1 + \lambda_2 = 1 + R - (\alpha - 1)\alpha \bar{k}^{-\alpha-1}/g$ and $\lambda_1 \lambda_2 = R$. Here we assume that the firm is owned by the household. Given the inelastic labor supply assumption in this model, we are able to model consumption-saving and investment decisions separately at first and then combine the decision rules.²⁵ We can now use (6), (50), and (51), to derive the innovation to consumers' perceived income as follows:²⁶

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

where

$$\Xi = \frac{1}{R-\rho} \left[1 + \frac{\alpha \rho (R+\delta)}{g(R-\lambda_1)(\lambda_2-\rho)} \right]. \quad (52)$$

The above expression shows a linear relationship between the innovation to total wealth and the innovation to aggregate productivity.

For the SOW, we have a similar expression:

$$\zeta_{t+1}^* \equiv \Xi^* \varepsilon_{t+1}^*,$$

where $\Xi^* = \frac{1}{R-\rho^*} \left[1 + \frac{\alpha \rho^* (R+\delta^*)}{g^*(R-\lambda_1^*)(\lambda_2^*-\rho^*)} \right]$ and

$$a_{t+1}^* = (1-\rho^*)\bar{a}^* + \rho^* a_t^* + \varepsilon_{t+1}^*, \quad (53)$$

²⁵See Glick and Rogoff (1995) and Gruber (2002) for detailed discussions on this specification.

²⁶The derivation of this result is available from the corresponding author upon request.

where $\rho^* \in [0,1]$ is the persistence coefficient, \bar{a}^* is the mean of the country-specific productivity shock, and ε_{t+1}^* is an *iid* Gaussian innovation with mean 0 and variance ω^{*2} . As in our benchmark model, we also assume that there is a positive contemporaneous correlation between ε_{t+1} and ε_{t+1}^* :

$$* \text{corr}(\varepsilon_{t+1}, \varepsilon_{t+1}^*) = \phi.$$

Given the productivity processes, the productivity correlation is

$$* \text{corr}(a_{t+1}, a_{t+1}^*) = \Pi_a \phi, \quad (54)$$

where $\Pi_a = \frac{\sqrt{(1-\rho^2)(1-\rho^{*2})}}{1-\rho\rho^*}$. Note that $\Pi_a = 1$ when $\rho = \rho^*$ and $\Pi_a < 1$ when $\rho \neq \rho^*$.

5.2 Theoretical Implications for Cross-Country Consumption Correlations

Combined with the FI-RE model proposed in Section 2.1, the change in consumption in the home country and the ROW can be written as

$$\Delta c_t = (R-1)\Xi\varepsilon_t \text{ and } \Delta c_t^* = (R-1)\Xi^*\varepsilon_t^*,$$

respectively. Using these two expressions, the international consumption correlation can thus be written as

$$* \text{corr}(\Delta c_t, \Delta c_t^*) = * \text{corr}(\varepsilon_t, \varepsilon_t^*) = \frac{1}{\Pi_a} * \text{corr}(a_t, a_t^*). \quad (55)$$

Note that if the estimated productivity persistence parameters, ρ and ρ^* , are different and less than 1, $\Pi_a < 1$ and $* \text{corr}(c_t, c_t^*) > * \text{corr}(a_t, a_t^*)$. This prediction contradicts the empirical evidence, just as in Section 2.3.

As shown in the benchmark model, the change in individual consumption can be expressed as

$$\Delta c_t^{RI} = \theta(R-1) \left[\frac{\zeta_t}{1-(1-\theta)R \cdot L} + \left(\xi_t - \frac{\theta R \xi_{t-1}}{1-(1-\theta)R \cdot L} \right) \right]. \quad (56)$$

Assume that $\xi = \bar{\xi}_t + \zeta_t^i$, where $\bar{\xi}_t = E^i[\xi_t]$ is the common noise and define $\mu \equiv \frac{*sd(\bar{\xi}_t)}{*sd(\xi_t)} \in [0,1]$ to measure the relative importance of the common components of the error generated by ζ_t . After aggregating all consumers, idiosyncratic components are cancelled out. We then have the following expression for the change in aggregate consumption in terms of the productivity shocks:

$$\Delta c_t = (R - 1) \left[\frac{\theta \zeta_t}{1 - (1 - \theta)R \cdot L} + \theta \left(\bar{\xi}_t - \frac{\theta R \bar{\xi}_{t-1}}{1 - (1 - \theta)R \cdot L} \right) \right]. \quad (57)$$

In the case of $\mu = 0$, all the endogenous noises are cancelled out

$$\Delta c_t^{RI} = \theta(R - 1) \frac{\zeta_t}{1 - (1 - \theta)R \cdot L} = \theta(R - 1) \frac{\Xi \varepsilon_t}{1 - (1 - \theta)R \cdot L}. \quad (58)$$

For the ROW, we have a similar expression. It is clear from these expressions that consumption adjusts gradually to productivity shocks. For different countries, the level of their fundamental uncertainty (i.e., the variance of the productivity shock) can affect optimal consumption decisions. Different levels of volatility of productivity shocks lead to different levels of optimal attention, θ , and thus heterogeneous consumption adjustments in response to the productivity shocks.²⁷ Just like in the benchmark model, this mechanism helps to explain why the consumption correlation is lower than the income correlation under optimal attention.

The consumption correlation between the home country and the ROW in the presence of capital accumulation can be written as

$$*corr(\Delta c_t^{RI}, \Delta c_t^{*RI}) = \Pi *corr(\varepsilon_t, \varepsilon_t^*) = \frac{\Pi}{\Pi_a} *corr(a_t, a_t^*), \quad (59)$$

where

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

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

²⁷Note that the higher θ is, the more new information a country can process.

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

It is clear that in the special case when $\mu = 0$,

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

In addition, when $\theta = \theta^*$, $\Pi = 1$, which gives the same result as in the benchmark model. As θ and θ^* drift apart, the consumption correlation becomes smaller relative to the income correlation.

5.3 Quantitative Results

For the calibration exercise, we need to simulate the standard deviation of the permanent income shock ω_ζ , which equals $\Xi\sigma(\varepsilon_t)$. Ξ is given by (52). To determine Ξ and the standard deviation of the productivity shock, $\sigma(\varepsilon_t)$, we first estimate the productivity process (47) with the G7 data. Table 8 displays the estimated persistence, ρ and ρ^* , and other characteristics of the process.

Following the existing literature (e.g., Glick and Rogoff 1995 and Gruber 2002), the interest rate (R) is set to be 1.04 and the share of capital (α) is set to be 0.37, 0.52, 0.32, 0.35, 0.36, 0.46, and 0.34 for Canada, Italy, UK, France, Germany, Japan and the U.S., respectively. The depreciation rate δ is set to be 0.05. In addition, we normalize α_a to 1. Glick and Rogoff (1995) also estimate that $\lambda_1 = 0.9$ and $\lambda_i = 0.36$. Marquez (2004) extends the sample and reestimates $\lambda_i = 0.12, 0.23, 0.14, 0.24, 0.21, 0.44$ and 0.33 for Canada, Italy, UK, France, Germany, Japan and the U.S., respectively. Here we adopt Marquez's results and use $\lambda_i = \frac{\rho\alpha\bar{k}^\alpha}{g(\lambda_2-\rho)}$ to recover the adjustment cost coefficient g for each country. Ξ can then be calculated accordingly. The same calculation procedure applies to the rest of the world, where the adjustment cost coefficient g is assumed to be 1, and the share of capital α is set to 0.33 as in Gruber (2002).

Table 9 reports simulated results alongside their empirical counterparts. Entries in the first four rows are comovements between a SOE (Canada, Italy, UK, France) and the ROW, where income and

consumption of the ROW are constructed using weighted averages of the G7 excluding the country in question. In the last row, the U.S. is treated as the ROW to Canada. Furthermore, the first two columns show the correlations of first-differenced GDP and first-differenced consumption based on the data. In the third column, the FI-RE model predicts that the consumption correlations are almost as high as their corresponding income correlations. The performance of the model is improved when we assume that consumers have limited and elastic attention. It is clear that the model generates much lower consumption correlations which fit the data quite well. Using the Canada-US case as an example, the consumption correlation is 0.80 under FI-RE, but is reduced to 0.56 when considering elastic attention, which matches the data perfectly.²⁸ In other cases, RI still has the potential to reduce the consumption correlations and match the data well.

6. 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 correlations in four small open economies in the G7. In addition, we find that the elastic attention assumption can better explain other key stochastic properties of the joint dynamics of consumption and income. Finally, we show that considering endogenous capital accumulation does not change the main results we obtain in our benchmark model.

²⁸Here we set $\theta^* = 0.7$. The results are robust to different values of θ^* .

7. Appendix

7.1 Deriving International Consumption Correlations under RI

Given μ and the change in aggregate consumption expression (57), the variance of aggregate consumption growth can be written as

$$\begin{aligned}
 * \text{var}(\Delta c) &= \theta^2(R-1)^2 \left\{ \frac{1}{1-(1-\theta)^2 R^2} \omega_\xi^2 + \mu^2 \left[1 + \frac{\theta^2 R^2}{1-(1-\theta)^2 R^2} \right] \omega_\xi^2 \right\} \\
 &= \theta^2(R-1)^2 \left\{ \frac{1}{1-(1-\theta)^2 R^2} + \mu^2 \left[1 + \frac{\theta^2 R^2}{1-(1-\theta)^2 R^2} \right] \frac{1}{\theta \frac{1}{1-(1-\theta)-R^2}} \right\} \omega_\xi^2 \\
 &= \theta^2(R-1)^2 \left\{ \frac{1}{1-[(1-\theta)R]^2} + \mu^2 \left[\frac{1}{(1-(1-\theta)R^2)\theta} - \frac{1}{1-[(1-\theta)R]^2} \right] \right\} \omega_\xi^2,
 \end{aligned} \tag{60}$$

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

$$* \text{var}(\Delta c^*) = \theta^{*2}(R-1)^2 \left\{ \frac{1}{1-[(1-\theta^*)R]^2} + \mu^2 \left[\frac{1}{(1-(1-\theta^*)R^2)\theta} - \frac{1}{1-[(1-\theta^*)R]^2} \right] \right\} \omega_{\xi^*}^2 \tag{61}$$

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 (57):

$$* \text{cov}(\Delta c, \Delta c^*) = \frac{\theta \theta^* (R-1)^2}{1-(1-\theta)(1-\theta^*)R^2} E[\zeta_t \zeta_t^*]. \tag{62}$$

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

$$* \text{corr}(\Delta c, \Delta c^*) = \frac{* \text{cov}(\Delta c, \Delta c^*)}{* \text{var}(\Delta c) * \text{var}(\Delta c^*)} \approx \frac{1}{[1-(1-\theta)(1-\theta^*)R^2] \sigma(\theta, \mu) \sigma(\theta^*, \mu)} * \text{corr}(\Delta y, \Delta y^*), \tag{63}$$

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

$$* \text{corr}(\Delta y_t, \Delta y_t^*) \approx \frac{E[\varepsilon_t \varepsilon_t^*]}{\omega \omega^*}.$$

7.2 Deriving Other Stochastic Properties of Consumption under RI

Using the income processes proposed in Section 2.2, we have $\Delta y_t \approx \varepsilon_t \approx (R-1)\zeta_t$ and $* \text{var}(\Delta y) \approx \omega^2 \approx (R-1)^2 \omega_\xi^2$. Using the variance of consumption growth we have derived in (60), we can compute the moments as follows:

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

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

and

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

References

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

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

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

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

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

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

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

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

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

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.

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

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.

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

Gruber, Joseph W. (2002), " Productivity Shocks, Habits, and the Current Account," International Finance Discussion Papers 733. Board of Governors of the Federal Reserve System.

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

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.

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

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

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

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

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

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

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

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.

Mac'kowiak, Bartosz and Mirko Wiederholt (2009), " Optimal Sticky Prices under Rational Inattention," *American Economic Review* 99(3), 769-803.

Marquez, Jaime (2004), " Productivity, Investment, and Current Accounts: Reassessing the Evidence," *Review of World Economics* 140(2), 283-301.

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

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.

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

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

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

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

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

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

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.

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.

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

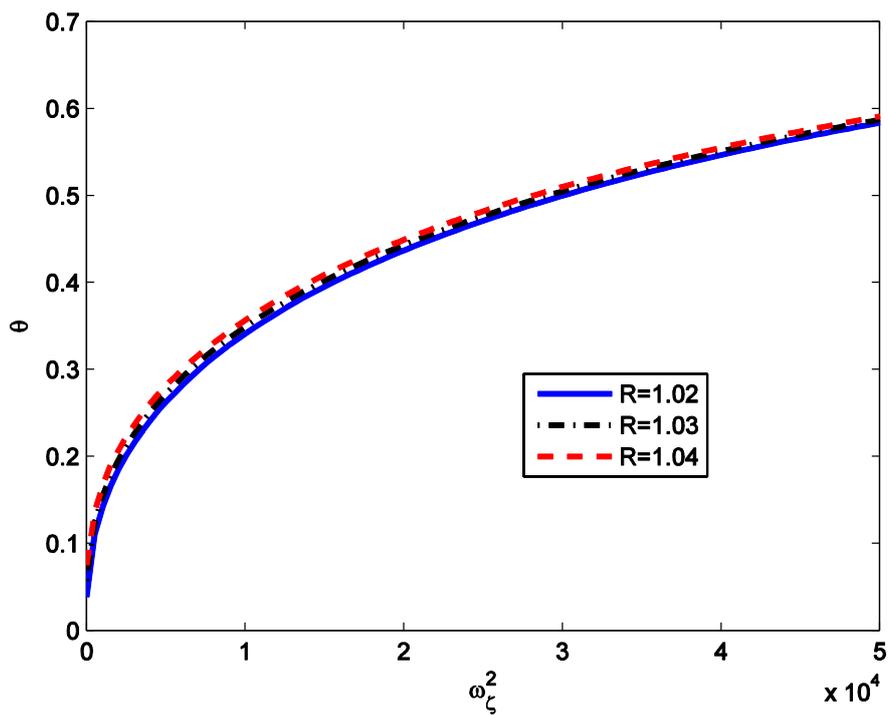


Figure 1: Effect of Fundamental Uncertainty on the Kalman Gain ($\lambda = 6.4 \times 10^4$)

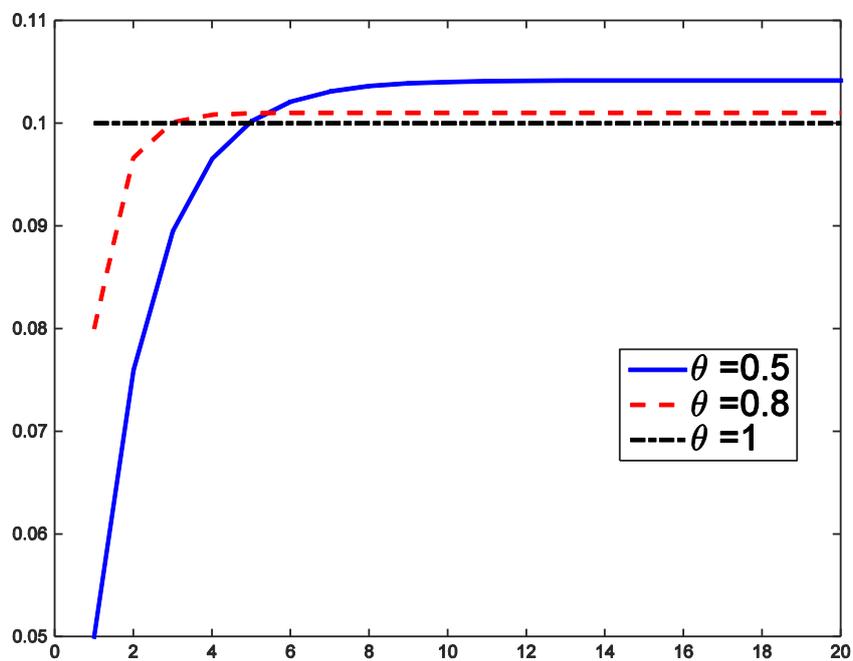


Figure 2: Impulse Responses of Consumption to Income Shock

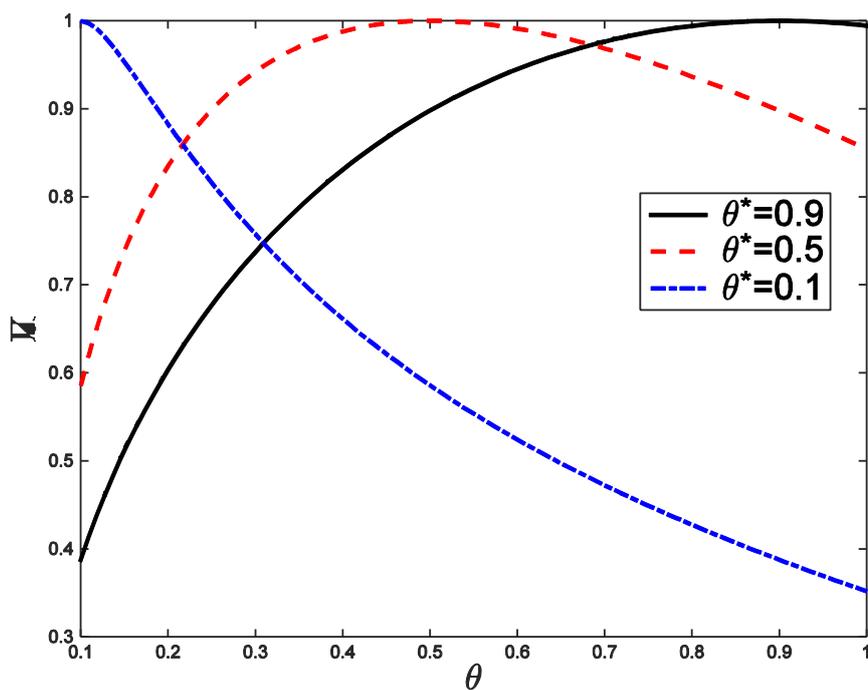


Figure 3: Effects of Elastic Attention on Consumption Correlation ($\mu = 0$)

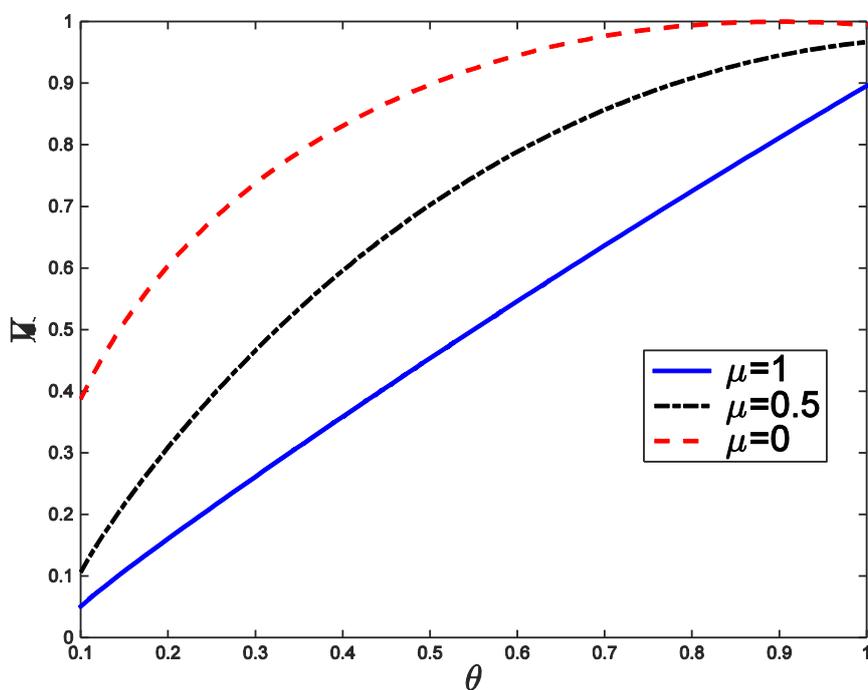


Figure 4: Effects of the Common Noise ($\theta^* = 0.9$)

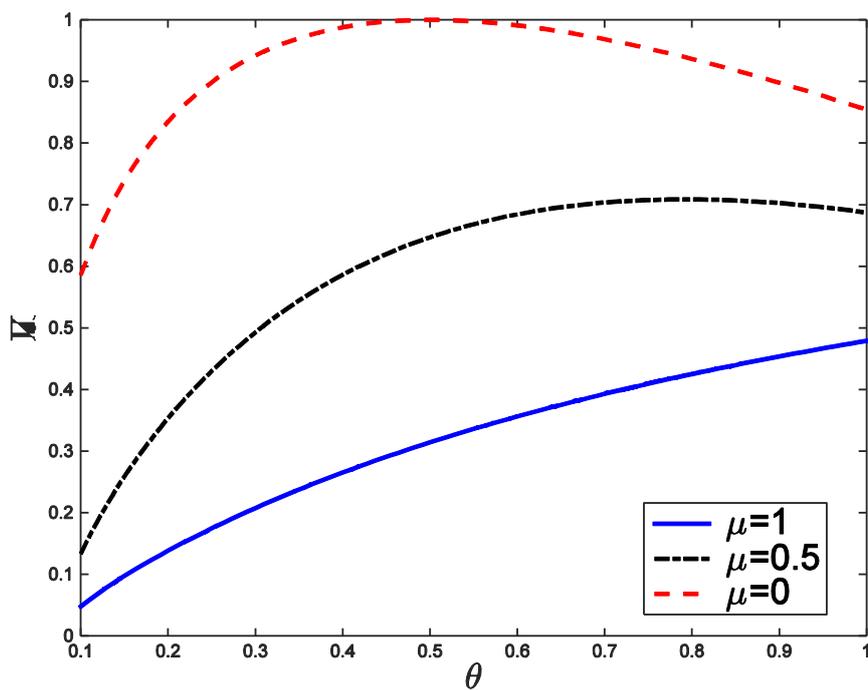


Figure 5: Effects of the Common Noise ($\theta^* = 0.5$)

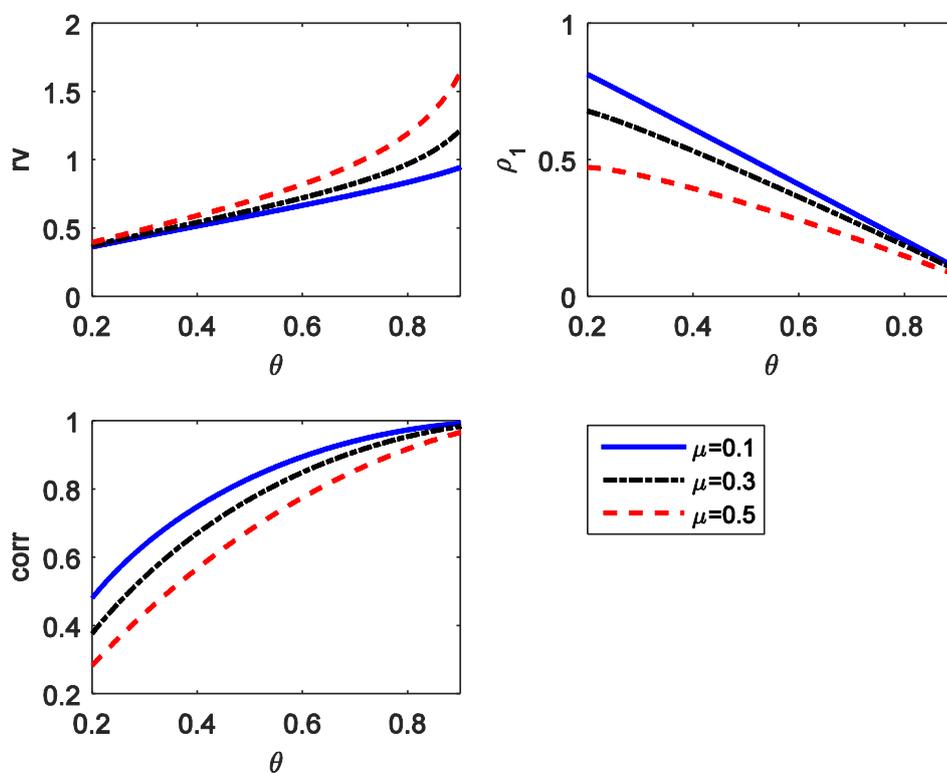


Figure 6: Stochastic Properties of Consumption and Income

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

	$\text{corr}(\Delta y, \Delta y^*)$	$\text{corr}(\Delta c, \Delta c^*)$	$sd(\Delta c)/sd(\Delta y)$	$\text{corr}(\Delta c, \Delta y)$	$\text{autocorr}(\Delta c)$
Canada	0.83 (0.05)	0.56 (0.08)	0.52 (0.07)	0.75 (0.07)	0.60 (0.08)
France	0.51 (0.17)	0.33 (0.13)	0.45 (0.07)	0.76 (0.06)	0.49 (0.11)
UK	0.75 (0.08)	0.70 (0.07)	0.73 (0.05)	0.93 (0.02)	0.56 (0.15)
Italy	0.44 (0.20)	0.11 (0.16)	0.48 (0.10)	0.75 (0.03)	0.46 (0.10)
Canada-US	0.87 (0.04)	0.58 (0.08)			
US			0.57 (0.05)	0.91 (0.02)	0.64 (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
($\mu = 0.2$)	0.56	0.83	0.61	0.54	0.50	0.46	0.43
($\mu = 0.3$)	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
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
($\mu = 0.2$)	0.33	0.51	0.41	0.37	0.34	0.32	0.3
($\mu = 0.3$)	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
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
($\mu = 0.2$)	0.70	0.75	0.67	0.62	0.59	0.56	0.53
($\mu = 0.3$)	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
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
($\mu = 0.2$)	0.11	0.44	0.37	0.34	0.32	0.30	0.28
($\mu = 0.3$)	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
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
($\mu = 0.2$)	0.58	0.87	0.53	0.46	0.42	0.38	0.35
($\mu = 0.3$)	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

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: Theoretical $\text{corr}(\Delta c, \Delta c^*)$ from different models - elastic attention channel only ($\mu = 0$)

	Data	RE		RI			
		$\theta^* = 1$	$\theta^* = 0.9$	$\theta^* = 0.8$	$\theta^* = 0.7$	$\theta^* = 0.6$	$\theta^* = 0.5$
Canada	0.56	0.83	0.66	0.62	0.60	0.59	0.59
France	0.33	0.51	0.43	0.41	0.40	0.39	0.39
UK	0.70	0.75	0.70	0.67	0.65	0.64	0.64
Italy	0.11	0.44	0.39	0.37	0.36	0.36	0.35
Canada-US	0.58	0.87	0.61	0.56	0.55	0.54	0.55

Table 6: Summary of data and model predictions PWT8.0 - elastic attention channel only ($\mu = 0$)

	$\text{corr}(\Delta y, \Delta y^*)^D$	$\text{corr}(\Delta c, \Delta c^*)^D$	$\text{corr}(\Delta c, \Delta c^*)^{RE}$		$\text{corr}(\Delta c, \Delta c^*)^{RI}$		
			$\theta^* = 1$	$\theta^* = 0.9$	$\theta^* = 0.8$	$\theta^* = 0.7$	$\theta^* = 0.6$
Canada	0.85	0.41	0.85	0.69	0.64	0.62	0.62
France	0.55	0.41	0.55	0.47	0.45	0.44	0.43
UK	0.75	0.67	0.75	0.70	0.67	0.66	0.65
Italy	0.39	0.19	0.39	0.35	0.34	0.33	0.32
Canada-US	0.87	0.38	0.87	0.61	0.56	0.55	0.55

Table 7: Comparing consumption moments from different models

	Data	RE	RI($\mu = 0$)		RI($\mu = 0.1$)		RI($\mu = 0.3$)	
			$\theta^* = 0.9$	$\theta^* = 0.8$	$\theta^* = 0.9$	$\theta^* = 0.8$	$\theta^* = 0.9$	$\theta^* = 0.8$
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	0.60	0.54
$autocorr(\Delta c)$	0.49	0	0.59	0.71	0.57	0.67	0.44	0.43
$corr(\Delta c, \Delta y)$	0.76	1	0.81	0.70	0.80	0.68	0.72	0.58
UK								
$sd(\Delta c)/sd(\Delta y)$	0.73	1	0.63	0.52	0.63	0.53	0.67	0.59
$autocorr(\Delta c)$	0.56	0	0.46	0.61	0.45	0.58	0.36	0.43
$corr(\Delta c, \Delta y)$	0.93	1	0.89	0.79	0.88	0.78	0.83	0.70
Italy								
$sd(\Delta c)/sd(\Delta y)$	0.48	1	0.58	0.48	0.58	0.49	0.63	0.56
$autocorr(\Delta c)$	0.46	0	0.53	0.67	0.52	0.64	0.40	0.44
$corr(\Delta c, \Delta y)$	0.75	1	0.85	0.74	0.84	0.73	0.77	0.64
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

Table 8: Endogenous output - estimation and calibration results for different countries

	ρ_a	ρ_a^*	$corr(\varepsilon, \varepsilon^*)$	$corr(a, a^*)$	Ξ	Ξ^*	Π_a
<i>Canada</i>	0.62	0.52	0.80	0.78	4.61	5.97	0.99
<i>France</i>	0.58	0.52	0.48	0.47	6.24	5.97	1.00
<i>UK</i>	0.64	0.52	0.65	0.67	5.23	5.97	0.98
<i>Italy</i>	0.41	0.52	0.32	0.37	4.43	5.97	0.99
<i>Canada-US</i>	0.62	0.51	0.85	0.83	4.61	6.78	0.99

Table 9: Endogenous output - summary of data and model predictions ($\mu = 0$)

	$\text{corr}(\Delta y, \Delta y^*)^D$	$\text{corr}(\Delta c, \Delta c^*)^D$	$\text{corr}(\Delta c, \Delta c^*)^{RE}$		$\text{corr}(\Delta c, \Delta c^*)^{RI}$		
			$\theta^* = 1$	$\theta^* = 0.9$	$\theta^* = 0.8$	$\theta^* = 0.7$	$\theta^* = 0.6$
Canada	0.83	0.56	0.80	0.62	0.58	0.56	0.55
France	0.51	0.33	0.48	0.43	0.41	0.40	0.40
UK	0.75	0.70	0.65	0.63	0.62	0.61	0.61
Italy	0.44	0.11	0.32	0.12	0.12	0.13	0.14
Canada-US	0.87	0.58	0.85	0.60	0.56	0.54	0.54

Table 10: Endogenous output - theoretical corr ($\Delta c, \Delta c^*$) from different models

	Data	RE ($\theta^* = 1$)	RI ($\theta^* = 0.9$)	RI ($\theta^* = 0.8$)	RI ($\theta^* = 0.7$)
Canada					
($\mu = 0$)	0.56	0.80	0.62	0.58	0.56
($\mu = 0.1$)	0.56	0.80	0.60	0.55	0.53
($\mu = 0.2$)	0.56	0.80	0.57	0.50	0.46
($\mu = 0.3$)	0.56	0.80	0.52	0.43	0.38
($\mu = 0.9$)	0.56	0.80	0.26	0.18	0.13
France					
($\mu = 0$)	0.33	0.48	0.43	0.41	0.40
($\mu = 0.1$)	0.33	0.48	0.43	0.40	0.39
($\mu = 0.2$)	0.33	0.48	0.41	0.38	0.36
($\mu = 0.3$)	0.33	0.48	0.39	0.35	0.31
($\mu = 0.9$)	0.33	0.48	0.24	0.17	0.13
UK					
($\mu = 0$)	0.70	0.65	0.63	0.62	0.61
($\mu = 0.1$)	0.70	0.65	0.63	0.61	0.60
($\mu = 0.2$)	0.70	0.65	0.62	0.60	0.57
($\mu = 0.3$)	0.70	0.65	0.60	0.57	0.53
($\mu = 0.9$)	0.70	0.65	0.45	0.34	0.26
Italy					
($\mu = 0$)	0.11	0.32	0.12	0.12	0.13
($\mu = 0.1$)	0.11	0.32	0.10	0.09	0.08
($\mu = 0.2$)	0.11	0.32	0.07	0.05	0.05
($\mu = 0.3$)	0.11	0.32	0.05	0.04	0.03
($\mu = 0.9$)	0.11	0.32	0.02	0.01	0.01
Canada-US					
($\mu = 0$)	0.58	0.85	0.60	0.56	0.54
($\mu = 0.1$)	0.58	0.85	0.59	0.53	0.51
($\mu = 0.2$)	0.58	0.85	0.54	0.47	0.42
($\mu = 0.3$)	0.58	0.85	0.48	0.39	0.34
($\mu = 0.9$)	0.58	0.85	0.23	0.15	0.11