Learning in International Markets and a Rational Expectation Approach to the Contagion Puzzle

By Steven Wei Ho*

Job Market Paper

Abstract: I develop a general equilibrium model in which agents from two countries do not observe directly the long-run growth prospects of their economies. Instead, the agents rationally learn the hidden components through the Kalman filter applied to international consumption data. Learning endogenously produces: (i) a rational explanation of international contagion phenomenon, defined as changes in one country’s asset prices in response to foreign news, that occurs in the absence of domestic news, (ii) large and time-varying international equity risk premia, and (iii) a resolution of the forward premium anomaly, defined as the tendency of high interest rate currencies to appreciate.

Keywords: Bayesian Learning, Contagion, Forward Premium Anomaly, Long-Run Risk, Dynamic Stochastic General Equilibrium

The most current version of this article can be found at http://areas.kenan-flagler.unc.edu/finance/phd/host/jmp.pdf

* I thank the support from my advisors Riccardo Colacito, David Ravenscraft and Mariano Croce. I am also grateful to Christian Lundblad, Pab Jotikasthira and Sydney Ludvigson. The faculty and Ph.D. students at Kenan-Flagler have provided useful suggestions. I also thank Jose Fillat and Karl Diether for comments. Department of Finance, Kenan-Flagler Business School, University of North Carolina at Chapel Hill. Email: Steven_Ho@kenan-flagler.unc.edu
1. Introduction

Many of the statistical challenges that plague econometricians presumably also plague market participants. Naive application of rational expectations equilibrium concepts may endow investors with too much knowledge about future growth prospects. Hansen et al. (2008)

Do agents know the long-run prospect of the economy? I address this question by investigating the international asset pricing implications of an economy in which agents have to learn the long-run prospect of the economy from international historical growth data.

A growing body of the literature has documented that long-run growth prospects influence both domestic Bansal and Yaron (2004) and international asset prices Colacito and Croce (2011). All these models are built around the assumption that consumption is exposed to two sources of shocks perfectly observable to the agents. Specifically, short-run risk is modeled as an i.i.d shock that only affects the economy for one period, whereas long-run risk is modeled as a small, but highly persistent autoregressive process, which is thus expected to affect long-run growth rates. A key question within the long-run risk paradigm is the ability of agents to perfectly detect such a small long-run component.

Earlier papers have investigated this question in a one-country setting (Ai, 2010; Bansal and Shaliastovich, 2011; Croce et al., 2012; Johannes et al., 2010). In this article, I study a two-country economy, each populated by one consumer. The dynamics of consumption growth within each country feature the aforementioned long-run risks. The agents do not know the actual value of each country’s specific long-run risk, in the sense that they cannot can accurately break down consumption news into a component which is going to affect the economy for many periods, and a component which represent short-run i.i.d shocks that is only going to affect the consumption stream for one period. The long-run components of the two countries are correlated (Colacito and Croce, 2011), thus innovation in one country can provide information regarding the long-run component of the other country.

The Bayesian learning mechanism employed in this article is the Kalman filter (Kalman and Bucy, 1961; Kalman, 1960). Due to the recursive nature of the Kalman filter, it can also be viewed as the simplest dynamic Bayesian Net-
work (Haykin, 2001; Murphy, 2002). An intuitive interpretation of the Kalman filter approach is that in each period the agent forms a prior regarding the joint probability distribution of the two latent variables, which are the home and foreign long-run risks, and come up with a prediction for the following period; as time goes by these predictions are compared with the realized observations. The agent adjusts the estimates of the latent variables and turns her posterior distribution into a new prior to be used in the subsequent period. Under such a setup, agent from one country would recursively Bayesian update her estimate of her country-specific long-run component by utilizing all available historical information in consumption growth in both countries, and optimizes her asset pricing decisions accordingly.

I find that learning can account for prominent features of domestic and international financial markets. Specifically, the learning model presented in this article offers a solution of the puzzling contagion phenomenon in the context of a dynamic stochastic equilibrium framework with complete markets and no arbitrage. Contagion is broadly defined as the propagation of shocks in excess of what can be explained by fundamentals (Forbes and Rigobon, 2000). There is no universally accepted explanation for why contagion occurs in equilibrium (Bekaert and Harvey, 2003; Karolyi and Stulz, 1996). This phenomenon is puzzling as one could potentially arbitrage when asset prices deviate from fundamentals. Current explanations offered include investor herding behavior (Cipriani and Guarino, 2008), and the spread of fear or “market sentiments” (Fromlet, 2001). However, such psychological arguments do not prevent an agent, who is perfectly rational and unmoved by sentiments, from undertaking arbitrage. I propose a rational expectations explanation for why contagion may occur in equilibrium.

I define the true long-run components of each country as the economic fundamentals, which are unobservable and thus have to be learned, taking the history of the consumption stream of both countries as observables. The agents of the two countries are endowed with the same information. When agents learn using Kalman filter, they form an estimate of the entire joint distribution of the two latent variables characterized by the means and the variance-covariance matrix. As a result, innovations in consumption streams of both countries, are two sources of information that will be incorporated into the estimate of the long-run persistent component of either one country. Since the growth rates of consumption and
their long-run components are correlated across countries, the estimate of home’s economic fundamental will be revised in response to foreign innovations, even in the absence of any new piece of information in the home country. Equivalently, Bayesian learning gives rise to contagion as one country may revise its long-run growth expectations solely in response to news coming from abroad.\footnote{Note that at each point in time, equilibrium prices are determined using agents’ estimates of the distribution of fundamentals given all historical information up to that point. This means that an econometrician who observes the entire data series may observe temporary mis-pricing, conditional on her information set which is larger than the agents’.
}

The model is also able to provide a rational explanation for the forward premium anomaly, i.e. the well-documented tendency of high interest rate currencies to appreciate (Fama, 1984), which is at odds with the prediction of the uncovered interest rate parity relationship. The model in this article can provide a solution due to its ability to endogenously generate time-varying volatility through learning. Take for example the case in which the home country agent believes that the estimate of her long-run component is very unreliable, i.e. it features a large variance. This variance of the estimation error will steadily decrease through learning in the transition toward the steady-state Kalman filter value. During this process, the consumption profile of the home agent becomes less uncertain and thus the interest rate in this country gets relatively larger. By no arbitrage, the currency of the home country become more valuable, being associated to a safer consumption profile, and it is thus expected to appreciate, despite its higher interest rate.

Furthermore, I show that within each country, learning increases the equity risk premium by as much as 22%. This finding confirms and extends the analysis of Croce et al. (2012) to the case in which investors exploit the cross-sectional dimension of countries in addition to the domestic time-series of consumption. Additionally, when the Kalman filter is off steady-state, the model endogenously generates time-varying uncertainty, as the recursive application of the Bayesian updating changes the accuracy of the estimated long-run risks through time. This means that the model features a time-varying risk-premium.

The rest of the article is organized as follows. Section 2 reviews the related literature. In Section 3, I describe the model setup, the information structure involved, and the learning. In Section 4, I present the analytical asset pricing solutions of learning under different information structures. In Section 5, I provide both the
theoretical arguments and numerical simulation results which would explain contagion in international markets and the forward-premium puzzle. Section 6 concludes the article.

2. Related Literature

In the Bansal and Yaron (2004) model, there is a small but highly persistent predictable long-run component in consumption growth, which is subject to long-run shocks. They are called long-run shocks since due to the highly persistent nature of the long-run component, even small innovations in this component will induce large cash flow movements in a long time horizon. Thus, the low-frequency movements in consumption growth rates are called the long-run risk (Bansal et al., 2010).

A notable ingredient of the long-run risk literature, which I also employ here, is the Epstein-Zin-Weil preference (Epstein and Zin, 1989; Weil, 1989). Under this recursive preference, agents in each period would optimize the tradeoff between utility of current period and continuation utility derived from all future periods (Backus et al., 2005). The recursive preference allows the separation of coefficient of relative risk aversion $\gamma$ and intertemporal elasticity of substitution (IES) $\psi$, which can be simultaneously large, whereas in standard CRRA preference $\psi = 1/\gamma$. This feature is desirable as an individual’s willingness to take financial risk, does not have to be necessarily associated with the inverse of her inclination to substitute today’s consumption with future consumption in response to change in intertemporal prices (Chen et al., 2011). When IES is larger than $1/\gamma$, agents prefer early resolution of uncertainty and exposure to the long-run risk carries high risk premium.

Hansen and Sargent (2010) consider the model in which agents are concerned with model mis-specification and Bayesian updates the parameters of several different sub-models. The agent will also update model-mixing probability and decide which sub-model is the most likely. This setup generates elevated risk premium since agent’s distrust of her model adds to the price of risk.

Croce et al. (2012) investigates the role of information in consumption based long-run risk model and is able to explain the downward slope of equity term structure. When investors can identify both the short-run and long-run components
of consumption risk, the standard long-run risk model can generate a sizable equity market risk premium only if the equity term structure slopes up. However, when investors cannot distinguish short-run and long-run components of consumption risk, the model is able to generate both a large equity market risk premium and a downward sloping equity term structure, which is what we observe in data.

Bansal and Shaliastovich (2011) focuses on the endogenous choice in learning. The agents may either pay a cost to learn the true long-run component or use Kalman filtering based on historical data. The actions of investors to learn about the true state can explain the asset-price jumps. The model implies income volatility can predict future jumps in returns.

Ai (2010) studies the implication of public information quality about persistent productivity shocks in a model with Kreps-Porteus preferences. The production-based long-run risk model with learning implies that when information quality is low equity premium is high and volatility of the risk-free rate is low.

3. Model Setup

3.1. Preferences

There are two countries in the economy which are denoted home and foreign. Markets are complete. As in Colacito and Croce (2011), perfect home bias is imposed in the setup so that representative agent of each country would only consume the goods that she is endowed with. Since markets are complete, agents can still trade Arrow-Debreu securities and they have no home bias over financial assets. The first order conditions with respect to the purchase of those financial assets (No-Arbitrage equations) pin down the value of all assets and the growth rate of the exchange rate. For compactness of representation, only home’s preferences and prices are characterized in this section. The foreign counterparts are denoted with identical expression with a superscript “*” added to designate the variables as foreign.

Following the long-run risk literature which emphasizes the importance of long-run economic growth prospects (Bansal and Yaron, 2004), the home representative agent has the Epstein-Zin-Weil (1989) preference:

$$U_t = \left\{ (1 - \delta)C_t^{1 - \frac{1}{\delta}} + \delta \left( E_t \left[ U_{t+1}^{1-\gamma} \right] \right)^{\frac{1-1/\sigma}{1-\gamma}} \right\}^{\frac{1-1/\sigma}{1-1/\sigma}}$$
A noteworthy feature of the Epstein-Zin-Weil preference is the separation of the coefficient of intertemporal elasticity of substitution (IES) \( \psi \) and the coefficient of relative risk aversion \( \gamma \), which can be simultaneously larger than one; whereas in standard CRRA preference \( \psi = 1/\gamma \). When \( IES > 1/\gamma \), agents prefer early resolution of uncertainty (Epstein and Zin, 1991). For convenience, we can define composite parameter \( \theta = \frac{1-\gamma}{1-1/\psi} \). Denote the ex-dividend price dividend ratio of an asset that pays a consumption stream \( C_t \) at end of period \( t \) as \( W_{c,t} = P_t^C/C_t \). Also denote the ex-dividend price dividend ratio of an asset that pays a consumption stream \( D_t \) at end of period \( t \) as \( W_{d,t} = P_t^D/D_t \). \( M_{t+1} \) is the pricing kernel and \( R_{f,t} \) is the return on a one-period risk-free asset at time \( t \). Given the Epstein-Zin-Weil preference, the optimal consumption choice yields the following asset pricing equations:

\[
\begin{align*}
E_t[M_{t+1}R_{c,t+1}] = 1 & \quad \Rightarrow \quad R_{c,t+1} = \frac{(P_{t+1}^C + C_{t+1})/P_t^C}{(P_t^C/C_t)} \\
E_t[M_{t+1}R_{d,t+1}] = 1 & \quad \Rightarrow \quad R_{d,t+1} = \frac{(P_{t+1}^D + D_{t+1})/P_t^D}{(P_t^D/D_t)}
\end{align*}
\]

\[
\begin{align*}
M_{t+1} &= \left( \delta \frac{C_{t+1}}{C_t} \right)^{-1/\psi} R_{c,t+1}^{\theta - 1} \\
R_{f,t} &= (E_t[M_{t+1}])^{-1}
\end{align*}
\]

One can show the log pricing kernel is a function of log consumption growth \( \Delta c_t \), and the log return on an asset which pays the consumption stream \( r_{c,t+1} \):

\[
m_{t+1} = \frac{1 - \gamma}{1 - 1/\psi} \log \delta - \frac{1 - \gamma}{\psi - 1} \Delta c_{t+1} + \frac{1}{1 - 1/\psi} r_{c,t+1}
\]

The Epstein-Zin-Weil preference and its basic asset pricing implications are shared by the three information structures describe below.

### 3.2. Full Information

Let the lower-case letters denote the variables in logarithms. The processes for log consumption and log dividend growth in the case of full information are specified as follows:
\[ \Delta c_t = \mu + z_t + \sigma_{c,h} \cdot \varepsilon_{c,t} \]
\[ \Delta c_t^* = \mu + z_t^* + \sigma_{c,f} \cdot \varepsilon_{c,t}^* \]
\[ z_t = \rho \cdot z_{t-1} + \sigma_{z,h} \cdot \varepsilon_{z,t} \]
\[ z_t^* = \rho \cdot z_{t-1}^* + \sigma_{z,f} \cdot \varepsilon_{z,t}^* \]
\[ \Delta d_t = \mu_d + \lambda \cdot z_t + \sigma_{d,h} \cdot \varepsilon_{d,t} \]
\[ \Delta d_t^* = \mu_d + \lambda \cdot z_t^* + \sigma_{d,f} \cdot \varepsilon_{d,t}^* \]
\[ (\varepsilon_{c,t}, \varepsilon_{c,t}^*, \varepsilon_{z,t}, \varepsilon_{z,t}^*, \varepsilon_{d,t}, \varepsilon_{d,t}^*) \sim N.i.i.d(0, \Omega_c) \]

\[ \Omega_c = \begin{bmatrix} 1 & \rho_c & 0 & 0 & 0 & 0 \\ \rho_c & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \rho_z & 0 & 0 \\ 0 & 0 & \rho_z & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \rho_d \\ 0 & 0 & 0 & 0 & \rho_d & 1 \end{bmatrix} \]

The small persistent predictable components \( z_t, z_t^* \) capture the small but persistent component of expected consumption and dividend growth rate in home and foreign countries, respectively. Note the timing convention here is that the predictive component \( z_t \) enters the consumption growth at time \( t \) rather than time \( t+1 \). In Bansal and Yaron (2004), \( z_t \) is the conditional expectation of consumption of next period, whereas in this timing convention the conditional expectation of consumption of next period is \( \rho z_t \). Let \( \varepsilon_{c,t}, \varepsilon_{c,t}^* \) denote the short-run shocks to home and foreign consumption. \( \varepsilon_{z,t}, \varepsilon_{z,t}^* \) are the long-run shocks to the highly persistent component. They are called the long-run shocks because due to the highly persistent nature of \( z_t \), even a small innovation can impact cash flows for a very long time and thus agents demand a high risk premia for the long-run shocks when \( IES > 1/\gamma \). \( \varepsilon_{d,t}, \varepsilon_{d,t}^* \) are the dividend shocks and \( \lambda \) can be interpreted as the leverage ratio (Abel, 1999). As in Bansal and Yaron (2004), dividend is not exposed to short-run shock. Indeed, for the short-run, long-run, dividend shocks within one country, \( \varepsilon_{c,t}, \varepsilon_{z,t}, \varepsilon_{d,t} \), each type of shock is orthogonal to any other type of shock. On the other hand, the cross-country correlations short-run, long-run, dividend shocks are denoted as \( \rho_c, \rho_z, \) and \( \rho_d \), respectively.
In the full information case, agents observe all variables of interest and fully take equation 4 into account when making asset pricing decisions.

3.3. Limited Information-Learning from Consumption Stream

Under the information structure of limited information with learning from consumption stream, agents are assumed to observe all historical data of consumption growth, however the agents do not observe the underlying long-run persistent components $z_t$ and $z^*_t$. As a consequence of limited information, the agents do not know if the innovation in consumption is due to $\varepsilon_{c,t}$ or $\varepsilon_{z,t}$. Because the linear state space of equation 4 is jointly Gaussian, the agents can use the Kalman filter (Harvey, 1989) to infer the underlying latent variables, namely $z_t$ and $z^*_t$. The Kalman filter is the optimal linear filter for Gaussian systems. Due to the recursive nature of the Kalman filter, it can also be viewed as the simplest dynamic Bayesian Network (Haykin, 2001; Murphy, 2002).

Under this information structure the agent is only learning from consumption stream but not dividend stream, the case of learning from both consumption and dividend stream would be investigated in the next section. It would still be interesting to price dividend under the current scenario, however, a question arises as to how to specify the exogenous dividend growth process. Suppose the dividend process is as in equation 4, then the agent must be able to infer the latent variables not only from the consumption stream, but also the dividend stream which contains additional information regarding true values of the latent variables. In order to achieve internal consistency with the information structure, dividend processes equation 4 are replaced with the following:

\begin{align*}
\Delta d_t &= \mu_d + \lambda \cdot (\Delta c_t - \mu) \\
\Delta d^*_t &= \mu_d + \lambda \cdot (\Delta c^*_t - \mu)
\end{align*}

The setup of the exogenous endowment processes are otherwise identical to the full information case. Note that the dividend process does not add any additional information for the filtering as it is a linear transformation of the consumption process, thus the agent only updates the filtered latent variable in response to consumption innovation.
As in Bansal and Shaliastovich (2011), I assume that $\Omega_c$ is known by the agent. In addition to knowledge about the structure of cash flow processes, agents of both countries share the same information set at time $t$ which is:

$$I_t^c = \{\Delta c_{t-1}^c\}_{i=0,1,...} \cup \{\Delta c_{t-1}^*\}_{i=0,1,...} \cup \Omega_c$$

Note that $\Omega_c$ is not time-varying. In fact, learning with off-steady-state Kalman filter can generate endogenous time-varying volatility; more details on this and the Forward Premium Anomaly can be found in section 5.2.

Let the filtered state vector for the case of learning from consumption stream be denoted by $\hat{z}_t$, $\hat{z}_t^*$

$$\hat{z}_t = E_t[z_t|I_t^c]$$
$$\hat{z}_t^* = E_t[z_t^*|I_t^c]$$

The variance covariance matrix of the filtering errors is:

$$P_t^c = \begin{bmatrix}
Phh_t^c & Phf_t^c \\
Phf_t^c & Pff_t^c
\end{bmatrix}$$

$$Phh_t^c = E_t[(z_t - \hat{z}_t)^2|I_t^c]$$
$$Phf_t^c = E_t[(z_t - \hat{z}_t)(z_t^* - \hat{z}_t^*)|I_t^c]$$
$$Pff_t^c = E_t[(z_t^* - \hat{z}_t^*)^2|I_t^c]$$

Given the filtered latent variables, the structure of the cash flow process implies the following innovation representation:

$$\Delta c_t = \mu + \hat{z}_t + \nu_{c,t}$$
$$\Delta c_t^* = \mu + \hat{z}_t^* + \nu_{c,t}^*$$
$$\Delta d_t = \mu_d + \lambda \cdot (\Delta c_t - \mu)$$
$$\Delta d_t^* = \mu_d + \lambda \cdot (\Delta c_t^* - \mu)$$

where the innovations in consumption in home and foreign countries are defined as:
\[
\begin{align*}
\nu_{c,t} &= \sigma_{c,h} \cdot \varepsilon_{c,t} + z_t - \tilde{z}_t \\
\nu^*_{c,t} &= \sigma_{c,f} \cdot \varepsilon^*_{c,t} + z^*_t - \tilde{z}^*_t
\end{align*}
\]

The filtering problem and its solutions are solved in Appendix A.

3.4. Limited Information-Learning from Consumption and Dividend Stream

The setup of the exogenous endowment processes are identical to the full information case. Notably the dividend process is:

\[
\begin{align*}
\Delta d_t &= \mu_d + \lambda \cdot z_t + \sigma_{d,h} \cdot \varepsilon_{d,t} \\
\Delta d_t^* &= \mu_d + \lambda \cdot z_t^* + \sigma_{d,f} \cdot \varepsilon_{d,t}^*
\end{align*}
\]

Both the consumption process and the dividend process provide information for the filtering of the state vector. The agents use Kalman filter to update the filtered latent variables in response to both consumption innovation and dividend innovation. Agents of both countries share the same information set at time t which is:

\[
I_t^d = \left\{ \{\Delta c_{t-i}\}_{i=0,1,...}, \{\Delta c^*_{t-i}\}_{i=0,1,...}, \{\Delta d_{t-i}\}_{i=0,1,...}, \{\Delta d^*_{t-i}\}_{i=0,1,...}, \Omega_c \right\}
\]

Let the filtered state vector for the case of learning from both consumption and dividend stream be denoted by \( \tilde{z}_t, \tilde{z}_t^* \)

\[
\begin{align*}
\tilde{z}_t &= E_t[z_t|I_t^d] \\
\tilde{z}_t^* &= E_t[z_t^*|I_t^d]
\end{align*}
\]

The variance covariance matrix of the filtering errors is:

\[
P_t^d = \begin{bmatrix}
Phh_t^d & Phf_t^d \\
Phf_t^d & Pff_t^d
\end{bmatrix}
\]

\[
\begin{align*}
Phh_t^d &= E_t[(z_t - \tilde{z}_t)^2|I_t^d] \\
Phf_t^d &= E_t[(z_t - \tilde{z}_t)(z_t^* - \tilde{z}_t)|I_t^d] \\
Pff_t^d &= E_t[(z_t^* - \tilde{z}_t)^2|I_t^d]
\end{align*}
\]

Given the filtered latent variables, the structure of the cash flow process implies the following innovation representation:
\[ \Delta c_t = \mu + \tilde{z}_t + \nu_{c,t} \]
\[ \Delta c_t^* = \mu + \tilde{z}_t^* + \nu_{c,t}^* \]
\[ \Delta d_t = \mu_d + \lambda \cdot \tilde{z}_t + \nu_{d,t} \]
\[ \Delta d_t^* = \mu_d + \lambda \cdot \tilde{z}_t^* + \nu_{d,t}^* \]

where the innovations in consumption and dividend, in home and foreign countries are defined as:
\[
\begin{align*}
\nu_{c,t} &= \sigma_{c,h} \cdot \varepsilon_{c,t} + z_t - \tilde{z}_t \\
\nu_{c,t}^* &= \sigma_{c,f} \cdot \varepsilon_{c,t}^* + z_t^* - \tilde{z}_t^* \\
\nu_{d,t} &= \sigma_{d,h} \cdot \varepsilon_{d,t} + \lambda \cdot (z_t - \tilde{z}_t) \\
\nu_{d,t}^* &= \sigma_{d,f} \cdot \varepsilon_{d,t}^* + \lambda \cdot (z_t^* - \tilde{z}_t^*)
\end{align*}
\]

The filtering problem and its solutions are solved in Appendix B

4. Analytical Model Solution

4.1. Full Information

The Euler equations 2 yield the policy function for price-consumption and price-dividend ratio as a function of the true long-run persistent components \( z_t, z_t^* \). Appendix C derives the log-linearized solutions for full information. In vector form, we have that for the home country:

\[ m_{t+1} = \frac{1}{\psi} z_{t+1} + \Gamma_m \eta_{t+1} \]
\[ r_{c,t+1} = \frac{1}{\psi} z_{t+1} + \Gamma_c \eta_{t+1} \]
\[ r_{d,t+1} = \frac{1}{\psi} z_{t+1} + \Gamma_d \eta_{t+1} \]

\[ E_t [r_{c,t+1}^{ex}] = -\Gamma_m S \Gamma_c - \frac{1}{2} \Gamma_c \Gamma_c' \]
\[ E_t [r_{d,t+1}^{ex}] = -\Gamma_m S \Gamma_d - \frac{1}{2} \Gamma_d \Gamma_d' \]
where \( \eta_t = (\sigma_{c,h} \cdot \varepsilon_{c,t}, \sigma_{c,f} \cdot \varepsilon_{c,t}^*, \sigma_{z,h} \cdot \varepsilon_{z,t}, \sigma_{z,f} \cdot \varepsilon_{z,t}^*, \sigma_{d,h} \cdot \varepsilon_{d,t}, \sigma_{d,f} \cdot \varepsilon_{d,t}^*) \)' is the vector of the shocks.

### 4.2. Limited Information-Learning from Consumption Stream

One can show the one-step-ahead evolution equations for the variances and covariances of the filtering errors are:

\[
P_{h \hat{h}}^c = \frac{\rho^2 \sigma_{c,h}^2 \left( \Phi_{h \hat{h}}^c \left( \rho^2 \sigma_{c,f}^2 + \Phi_{f \hat{f}}^c \right) \right)}{\left( \sigma_{c,h}^2 + \Phi_{h \hat{h}}^c \right) \left( \sigma_{c,f}^2 + \Phi_{f \hat{f}}^c \right) - \left( \rho \sigma_{c,h} \sigma_{c,f} \right)^2} + \sigma_{z,h}^2
\]
\[
P_{f f}^c = \frac{\rho^2 \sigma_{c,f}^2 \left( \rho^2 - 1 \right) \Phi_{f f}^c \sigma_{c,h} \sigma_{c,f} + \rho \left( \Phi_{f f}^c - \Phi_{h \hat{h}}^c \right)}{\left( \rho^2 - 1 \right) \Phi_{f f}^c \sigma_{c,h} \sigma_{c,f} + \rho \left( \Phi_{f f}^c - \Phi_{h \hat{h}}^c \right)} + \sigma_{z,f}^2
\]

One can derive the 2 by 2 Kalman filter in this case:

\[
K^c = \begin{bmatrix}
K_{11}^c & K_{12}^c \\
K_{21}^c & K_{22}^c
\end{bmatrix}
\]

where

\[
K_{11}^c = \frac{\Phi_{f f}^c (- \Phi_{h \hat{h}}^c - \rho \sigma_{c,h} \sigma_{c,f} \sigma_{c,f}) \left( \sigma_{c,f}^2 + \Phi_{f f}^c \right) + \Phi_{h \hat{h}}^c \left( \sigma_{c,f}^2 + \Phi_{f f}^c \right)}{\left( \sigma_{c,h}^2 + \Phi_{h \hat{h}}^c \right) \left( \sigma_{c,f}^2 + \Phi_{f f}^c \right) - \left( \rho \sigma_{c,h} \sigma_{c,f} \right)^2}
\]
\[
K_{12}^c = \frac{\Phi_{f f}^c \left( \sigma_{c,h}^2 + \Phi_{h \hat{h}}^c \right) \left( \sigma_{c,f}^2 + \Phi_{f f}^c \right) - \left( \rho \sigma_{c,h} \sigma_{c,f} \right)^2}{\left( \sigma_{c,h}^2 + \Phi_{h \hat{h}}^c \right) \left( \sigma_{c,f}^2 + \Phi_{f f}^c \right) - \left( \rho \sigma_{c,h} \sigma_{c,f} \right)^2}
\]
\[
K_{21}^c = \frac{\Phi_{f f}^c \left( \sigma_{c,h}^2 + \Phi_{h \hat{h}}^c \right) \left( \sigma_{c,f}^2 + \Phi_{f f}^c \right) - \left( \rho \sigma_{c,h} \sigma_{c,f} \right)^2}{\left( \sigma_{c,h}^2 + \Phi_{h \hat{h}}^c \right) \left( \sigma_{c,f}^2 + \Phi_{f f}^c \right) - \left( \rho \sigma_{c,h} \sigma_{c,f} \right)^2}
\]
\[
K_{22}^c = \frac{\Phi_{f f}^c \left( \sigma_{c,h}^2 + \Phi_{h \hat{h}}^c \right) \left( \sigma_{c,f}^2 + \Phi_{f f}^c \right) - \left( \rho \sigma_{c,h} \sigma_{c,f} \right)^2}{\left( \sigma_{c,h}^2 + \Phi_{h \hat{h}}^c \right) \left( \sigma_{c,f}^2 + \Phi_{f f}^c \right) - \left( \rho \sigma_{c,h} \sigma_{c,f} \right)^2}
\]

the one-step-ahead state evolution equations for the filtered home and foreign long-run persistent components have the following expressions:

\[
\hat{z}_t = \rho \cdot \hat{z}_{t-1} + K_{11}^c \cdot (\sigma_{c,h} \cdot \varepsilon_{c,t} + z_t - \hat{z}_t) + K_{12}^c \cdot (\sigma_{c,f} \cdot \varepsilon_{c,t}^* + z_t^* - \hat{z}_t)
\]
\[
\hat{z}_t^* = \rho \cdot \hat{z}_{t-1}^* + K_{21}^c \cdot (\sigma_{c,h} \cdot \varepsilon_{c,t} + z_t - \hat{z}_t) + K_{22}^c \cdot (\sigma_{c,f} \cdot \varepsilon_{c,t}^* + z_t^* - \hat{z}_t)
\]

The Kalman Gains in equation 29-30 can be steady-state Kalman Gain calculated
by solving the Riccati Equation A.18, or they could be off-steady-state in which case the Kalman Gains themselves are “state” variables, evolving according to equations 25-28 from initial conditions.

Under limited information, the Euler equations 2 yield the policy function for price-consumption and price-dividend ratio as a function of the filtered long-run persistent components \( \hat{z}_t, \hat{z}^*_t \). Utilizing the mapping machinery in Appendix D, one could show that once we define

\[
\eta_t^c = \begin{pmatrix}
\nu_{c,t} \\
\nu^*_{c,t} \\
K_{11}c \nu_{c,t} + K_{12}c \nu^*_{c,t} \\
K_{21}c \nu_{c,t} + K_{22}c \nu^*_{c,t} \\
\lambda \cdot \nu_{c,t} \\
\lambda \cdot \nu^*_{c,t}
\end{pmatrix}
\]

and \( S_{LC} = \Sigma_{LC} P_{LC} \Sigma'_{LC} \) (D.6), we have the following asset pricing results:

\[
\begin{align*}
\hat{m}_{t+1} &= m - \frac{1}{\psi} \hat{z}_{t+1} + \gamma \nu_{c,t+1} - \frac{\left( \gamma - \frac{1}{\psi} \right) \kappa_c \left( \nu_{c,t+1} K_{11}c + K_{12}c \nu^*_{c,t+1} \right)}{1 - \rho \kappa_c} \\
\hat{r}_{c,t+1} &= \hat{r}_c + \frac{1}{\psi} \hat{z}_{t+1} + \nu_{c,t+1} + \frac{\left( 1 - \frac{1}{\psi} \right) \kappa_c \left( \nu_{c,t+1} K_{11}c + K_{12}c \nu^*_{c,t+1} \right)}{1 - \rho \kappa_c} \\
\hat{r}_{d,t+1} &= \hat{r}_d + \frac{1}{\psi} \hat{z}_{t+1} + \lambda \nu_{c,t+1} - \frac{\left( \lambda - \frac{1}{\psi} \right) \kappa_d \left( \nu_{c,t+1} K_{11}c + K_{12}c \nu^*_{c,t+1} \right)}{1 - \rho \kappa_d}
\end{align*}
\]

\[
E_t[r_{c,t+1}^{ex}] = -\Gamma_m S_{LC} \Gamma'_c - \frac{1}{2} \Gamma_c S_{LC} \Gamma'_c \\
E_t[r_{d,t+1}^{ex}] = -\Gamma_m S_{LC} \Gamma'_d - \frac{1}{2} \Gamma_d S_{LC} \Gamma'_d
\]

4.3. Limited Information-Learning from Consumption and Dividend Stream

The expressions for the evolution equation of Kalman Gain and filtering error are quite involved, and are included in Appendix B.

The one-step-ahead state evolution equations for the filtered home and foreign long-run persistent components are:
\[
\tilde{z}_t = \rho \cdot \tilde{z}_{t-1} + K_{d11}^d \cdot \nu_{c,t} + K_{d12}^d \cdot \nu_{c,t}^* + K_{d13}^d \cdot \nu_{d,t} + K_{d14}^d \cdot \nu_{d,t}^*
\]

(32)

\[
\tilde{z}_t^* = \rho \cdot \tilde{z}_{t-1}^* + K_{d21}^d \cdot \nu_{c,t} + K_{d22}^d \cdot \nu_{c,t}^* + K_{d23}^d \cdot \nu_{d,t} + K_{d24}^d \cdot \nu_{d,t}^*
\]

(33)

The Kalman Gains in equation 32-33 can be steady-state Kalman Gain calculated by solving the Riccati Equation B.22, or they could be off-steady-state and evolve from initial conditions according to equations B.7-B.14.

Under limited information, the Euler equations 2 yields the policy function for price-consumption and price-dividend ratio as a function of the filtered long-run persistent components \(\tilde{z}_t, \tilde{z}_t^*\). Utilizing the mapping machinery in Appendix D, one could show that once we define

\[
\eta_t^d = \begin{pmatrix}
\nu_{c,t+1} \\
\nu_{c,t+1} \nu_{c,t+1}^* \\
\nu_{c,t+1} + K_{d11}^d + K_{d13}^d \nu_{d,t+1} + K_{d12}^d \nu_{c,t+1}^* + K_{d14}^d \nu_{d,t+1}^* \\
\nu_{c,t+1} + K_{d12}^d \nu_{c,t+1} + K_{d13}^d \nu_{d,t+1} + K_{d14}^d \nu_{d,t+1}^* \\
\nu_{c,t+1} + K_{d12}^d \nu_{c,t+1} + K_{d13}^d \nu_{d,t+1} + K_{d14}^d \nu_{d,t+1}^* \\
\nu_{c,t+1} + \nu_{d,t+1} \\
\nu_{c,t+1} + \nu_{d,t+1}^*
\end{pmatrix}
\]

and \(S_{LCD} = \Sigma_{LCD}P_{LCD}\Sigma_{LCD}'\) (D.9),

we have the following asset pricing results:

\[
m_{t+1} = \bar{m} - \frac{1}{\psi} \tilde{z}_{t+1} - \gamma \nu_{c,t+1} - \frac{\left(\gamma - \frac{1}{\psi}\right) \kappa_c \left(\nu_{c,t+1} K_{d11}^d + K_{d13}^d \nu_{d,t+1} + K_{d12}^d \nu_{c,t+1}^* + K_{d14}^d \nu_{d,t+1}^*\right)}{1 - \rho \kappa_c}
\]

(34)

\[
r_{c,t+1} = \bar{r}_c + \frac{1}{\psi} \tilde{z}_{t+1} + \nu_{c,t+1} + \frac{\left(1 - \frac{1}{\psi}\right) \kappa_c \left(\nu_{c,t+1} K_{d11}^d + K_{d13}^d \nu_{d,t+1} + K_{d12}^d \nu_{c,t+1}^* + K_{d14}^d \nu_{d,t+1}^*\right)}{1 - \rho \kappa_c}
\]

\[
r_{d,t+1} = \bar{r}_d + \frac{1}{\psi} \tilde{z}_{t+1} + \nu_{c,t+1} + \nu_{d,t+1} - \frac{\left(\lambda - \frac{1}{\psi}\right) \kappa_d \left(\nu_{c,t+1} K_{d11}^d + K_{d13}^d \nu_{d,t+1} + K_{d12}^d \nu_{c,t+1}^* + K_{d14}^d \nu_{d,t+1}^*\right)}{1 - \rho \kappa_d}
\]

and the conditional expectations of excess returns are given by
\[ E_t[r_{ct+1}^{xe}] = -\Gamma_m S_{LCD} \Gamma'_c - \frac{1}{2} \Gamma_c S_{LCD} \Gamma'_c \]
\[ E_t[r_{dt+1}^{xe}] = -\Gamma_m S_{LCD} \Gamma'_d - \frac{1}{2} \Gamma_d S_{LCD} \Gamma'_d \]

5. Results

5.1. Contagion

In this section I will present both theoretical and simulation results. By utilizing the mapping machinery in Appendix D and the log-linearization results developed in Appendix C, one can derive the following theoretical expressions for the pricing kernel as in section 4,

For the information structure of learning from consumption stream, we have:

\[
m_{t+1} = \bar{m} - \frac{1}{\psi} \tilde{z}_{t+1} - \gamma \nu_{c,t+1} - \left( \frac{\gamma - \frac{1}{\psi}}{1 - \rho \kappa_c} \right) \kappa_c \left( \nu_{c,t+1} K_{c,11}^c + K_{c,12}^c \nu_{c,t+1}^* \right) \]

\[
m_{t+1}^* = \bar{m}^* - \frac{1}{\psi} \tilde{z}_{t+1}^* - \gamma \nu_{c,t+1}^* - \left( \frac{\gamma - \frac{1}{\psi}}{1 - \rho \kappa_c} \right) \kappa_c \left( \nu_{c,t+1} K_{c,21}^c + K_{c,22}^c \nu_{c,t+1}^* \right) \]

For the information structure of learning from consumption and dividend stream, we have:

\[
m_{t+1} = \bar{m} - \frac{1}{\psi} \tilde{z}_{t+1} - \gamma \nu_{c,t+1} - \left( \frac{\gamma - \frac{1}{\psi}}{1 - \rho \kappa_c} \right) \kappa_c \left( \nu_{c,t+1} K_{d,11}^d + K_{d,13}^d \nu_{d,t+1} + K_{d,12}^d \nu_{c,t+1}^* \right) \]

\[
m_{t+1}^* = \bar{m}^* - \frac{1}{\psi} \tilde{z}_{t+1}^* - \gamma \nu_{c,t+1}^* - \left( \frac{\gamma - \frac{1}{\psi}}{1 - \rho \kappa_c} \right) \kappa_c \left( \nu_{c,t+1} K_{d,21}^d + K_{d,23}^d \nu_{d,t+1} + K_{d,22}^d \nu_{c,t+1}^* \right) \]

There is no unique definition in international asset pricing literature as to what contagion is. It can be viewed as foreign influence on US equity risk premium (Chan et al., 1992). It can be viewed as the documented phenomenon that shocks to Nikkei Index can impact S&P 500 and vice-versa (Karolyi and Stulz, 1996). It can be defined as how regional markets respond to information in one country, such as the crisis Asia in 1997 and Latin American in 1994 (Bekaert and Harvey, 2003). However, Bekaert et al. (2005) maintains that contagion should be more
than just what is revealed as increased correlation during crisis, since statistically it is natural for correlation to go up when volatility is high (Ang and Bekaert, 2002). The preferred definition of contagion is the propagation of shocks in excess of what can be explained by fundamentals (Forbes and Rigobon, 2000). This is the definition of contagion I use in this article, and I define economic fundamentals as the long-run persistent components.

Under full information, the pricing kernel of home only depends on home’s long-run component. Under limited information with learning, one could readily see how contagion occurs by observing equation 35 and equation 36. The home pricing kernel will respond to foreign innovation in consumption $\nu^{*}_{c,t+1}$, and when learning from dividend stream as well, home pricing kernel will also respond to foreign innovation in dividend $\nu^{*}_{d,t+1}$. Similarly, foreign pricing kernel will respond to home innovations. Since markets are complete and there is no arbitrage, all assets are priced using the pricing kernel. Thus home asset price will drop in response to foreign bad news and rise in response to foreign good news, even when there is no shock to home fundamental.

From simulation results, the Impulse Response Functions (IRF) plots can be used to illustrate this point. Figure 1 shows that under full information, home pricing kernel, and associated asset prices such as risk premium and risk free rate does not respond at all to foreign shock. However, when there is learning involved, both figure 2 and figure 3 show that either short-run or long-run or dividend shock from foreign country can impact home’s pricing kernel and returns for an extended period of time.

From the point of view of an agent who uses Kalman filtering to Bayesian update the fundamentals, when innovation in consumption stream occurs, the agent does not know the current economic fundamental and cannot break down the innovation into a short-run component which would affect the economy for only one-period and a long-run component which would affect the economy for many periods. The agent can only optimize asset pricing decisions using her belief of the distribution of the fundamental, formed through using all historical information up to that point in time. However, from the point of view of an econometrician who observe the entire data series and estimate the underlying economic fundamentals at each point in time, she is in a better position to tell if the shock was short term or long
term. In other words, there is arbitrage to eyes of the econometrician but not to the agents, since the information set of the econometrician who observe the entire history is larger than the information set of the rational agents who at each point in time observe information only up to that point.

5.2. Forward Premium Puzzle

For compactness of presentation, I will focus on the information structure of learning from consumption stream; all the results also follow for the case of learning from consumption and dividend stream. To solve the Forward Premium Puzzle, we need to establish the relationship between $P_{hh,t}^c$ and the off-steady-state Kalman Gains. It can be shown in the comparative statics results that $\frac{\partial K_{11,t}}{\partial P_{hh,t-1}^c}$ is positive and orders of magnitude larger than $\frac{\partial K_{12,t}}{\partial P_{hh,t-1}^c}$, $\frac{\partial K_{21,t}}{\partial P_{hh,t-1}^c}$, $\frac{\partial K_{22,t}}{\partial P_{hh,t-1}^c}$. This is intuitive because off steady-state, $P_{hh,t-1}^c$ is the agent’s belief of the variance of the filtering error of $\hat{z}_{t-1}$. $P_{hh,t-1}^c$ has the largest effect on $K_{11,t}^c$ which measures how the agent update her estimate of home long-run persistent component in response to home consumption innovation. The intuition behind $\frac{\partial K_{11,t}}{\partial P_{hh,t-1}^c} > 0$ is that if $P_{hh,t-1}^c$ is high, then the agent doesn’t trust her previous estimate $\hat{z}_{t-1}$ as much, thus there is a need to revise it based on new innovation, and the Kalman Gain $K_{11,t}^c$ is high for next period; conversely, if $P_{hh,t-1}^c$ is low, then the agent is already quite confident with her estimate $\hat{z}_{t-1}$ which is based on all historical information, and therefore there is little need to revise it, thus $K_{11,t}^c$ is low.

Suppose the agents of both countries start off at an off steady-state prior regarding the filtering error variance of home country’s, but not foreign country’s long-run persistent component. Let $P_{hh,t=0}^c = 150\% \times P_{hh,ss}^c$, where $P_{hh,ss}^c$ is the steady state prior obtained from Riccati equation A.18. Since the agents are learning rationally, through Kalman filter the agents will find out this prior was set to be too high. The consequence of the Kalman filter dynamics is that $\lim_{t-\infty} P_{hh,t}^c = P_{hh,ss}^c$. In other words, $P_{hh,t}^c$ is decreasing over time on a deterministic path. We can see from how this generates endogenous time-varying volatility in the pricing kernel from equation 35. From section 4.2, we also have the following analytical log-linearized asset pricing equations:
The expected change of next period’s exchange rate growth is:  
\[ E_t[\Delta e_{t+1}] = E_t[m_{t+1}^* - m_{t+1}] \]

which by properties of log-normal distribution and equation 35 is:

\[
E_t[\Delta e_{t+1}] = \overline{m}^* - \overline{m} + E_t \left[ \frac{1}{\psi} (z_{t+1}^* - z_{t+1}) \right] - \frac{(\gamma - \frac{1}{\psi}) \kappa_e (\nu_{c,t+1}K_{21,t+1}^e - K_{11,t+1}) + \nu_{c,t+1}^* [K_{22,t+1}^e - K_{12,t+1}^e]}{1 - \rho \kappa_e} \right] \\
+ \frac{1}{2} V_t \left[ \frac{1}{\psi} (z_{t+1}^* - z_{t+1}) \right] - \frac{(\gamma - \frac{1}{\psi}) \kappa_e (\nu_{c,t+1}K_{21,t+1}^e - K_{11,t+1}) + \nu_{c,t+1}^* [K_{22,t+1}^e - K_{12,t+1}^e]}{1 - \rho \kappa_e} \right]
\]

I include the second order terms as well for the forward premium \( f_t - s_t = r_{f,t} - r_{f,t}^* \)

(Backus et al., 2002).

\[
\begin{align*}
(41) \quad r_{f,t+1} & = r_f \cdot \overline{m} + \frac{1}{\psi} E_t (z_{t+1}^* - z_{t+1}) + \frac{1}{2\psi^2} (V_t[z_{t+1}^*] - V_t[z_{t+1}]) \\
(42) \quad r_{f,t+1} & = r_f \cdot \overline{m} + \frac{1}{\psi} E_t (z_{t+1}^* - z_{t+1}) + \frac{1}{2\psi^2} (V_t[z_{t+1}^*] - V_t[z_{t+1}]) \\
(43) \quad & = r_f \cdot \overline{m} + \frac{1}{\psi} E_t (z_{t+1}^* - z_{t+1}) + \frac{1}{2\psi^2} (P_{ff,t}^e - P_{hh,t}^e)
\end{align*}
\]

The \( \beta_{U1P} \) coefficient is obtained from the following regression:  
\( \Delta e_{t+1} = \alpha + \beta_{U1P} \cdot (r_{f,t} - r_{f,t}^*) + \varepsilon_t \). Notice the \( \frac{1}{\psi} E_t (z_{t+1}^* - z_{t+1}) \) term which appears in both \( r_{f,t+1} - r_{f,t+1}^* \) and \( E_t[\Delta e_{t+1}] \); one can readily see that without time-varying volatility, the \( \beta_{U1P} \) coefficient is equal to one (Backus et al., 2002).
Since the comparative statics results show the predominant effect of $P_{hh,t-1}$ is on $K_{11,t}$, I will group the nuisance terms as $Y_t$ and simplify the above expression as:

$$E_t[\Delta e_{t+1}] = m^* - m + \frac{1}{\psi} E_t(\tilde{z}_{t+1} - \tilde{z}_{t+1}^*) + V_t \left[ \frac{\left(\gamma - \frac{1}{\psi}\right) \kappa_c \left(-\nu_c t + 1 K_{11,t}^c\right)}{1 - \rho \kappa_c} + Y_t \right]$$

Thus we can see as time goes by, $P_{hh,t-1}$ is decreasing, $K_{11,t}$ is also decreasing, and therefore $E_t[\Delta e_{t+1}]$ is decreasing. The intuition is that as the estimate $\tilde{z}_t$ gets more reliable, there is less risk in home, and home’s currency appreciates. For the interest rate differential however, since $P_{hh,t}$ is decreasing, $r_{f,t+1} - r_{f,t+1}^*$ is increasing. The intuition is that as the estimate $\tilde{z}_t$ gets more reliable, home has less incentive to invest and demands higher risk free rate.

Thus the above dynamics generates negative $\beta_{UIP}$ coefficient. This provides a solution to the forward premium puzzle.

5.3. Numerical Results

I performed third order numerical approximations for 25 chains of simulations, each with 700 periods. The parameters used in the calibrations are listed in table 1. Variables of interest have the following symbols: $z_t$ is the long-run persistent component; $\Delta ln C_t$ is log consumption growth; $\Delta ln D_t$ is log dividend growth; $m$ is pricing kernel; $\Delta e_t$ is log exchange rate growth; $r_{f,t}$ is log risk free rate; $r_{d,t}$ is log return on dividend stream; $r_{ex,t}$ is equity risk premium; $ln \frac{P_{t}}{C_{t}}$ is the log price consumption ratio ; $r_{c,t}$ is the log return on the asset which pays the consumption stream.

As shown in table 8, learning can elevate risk premium since under learning, the pricing kernel is more volatile due to uncertainty regarding the underlying latent variables. Table 9 shows that the pricing kernel is less volatile under learning from both consumption and dividend stream, compared to learning from consumption stream only, since learning from multiple information sources increases the accuracy of estimated latent variables, which is documented in the filtering error covariance matrices, table 6 and table 7. It is notable that the exchange rate volatility does not go up much compared to the full-information case, since learning generates higher cross-country correlation of the estimated long-run components. In the
data between US and UK from 1979 to 2006, the volatility of exchange rate is around 11% (Colacito and Croce, 2011).

Another consequence of learning is that the cross-country correlation of risk free rate $r_{f,t}$, and return on the asset which pays consumption stream $r_{c,t}$, are now higher, as shown in table 11. This effect is more prominent for the case of learning from consumption, and less prominent for the case of learning from consumption and dividend. This is also due to that learning from multiple sources of information enables the agents to better infer the latent variables with accuracy and thus move closer toward full information case.

The fact that agents cannot distinguish short-run from long-run shock can also be observed in the correlation matrix of mapped shocks in table 4 and table 5, which were obtained through the mapping procedure developed in Appendix D. While the true correlation matrix under full information (table 3) has zero correlation between different types of shocks, that is no longer true under learning.

6. Concluding Remarks

This article documents the asset pricing implications of learning in consumption-based long-run risk literature in the context of international markets. There are two countries and the long-run prospects of them are not directly observable, though they are known to be correlated. Agents recursively Bayesian update their estimates of long-run persistent components through Kalman filter. Learning provides an explanation for the contagion phenomenon, defined as changes in one country’s asset prices in response to foreign news, that occurs in the absence of domestic news. Learning generates higher correlation of pricing kernels across two countries, and higher risk premium without increasing volatility of exchange rate. When evaluated off steady-state, learning can generate time-varying volatility in risk premium and can also provide an explanation for the forward premium puzzle.

It has been observed in data that correlation across countries is higher in bear markets than in bull markets (Ang and Bekaert, 2002). For future research, one possible extension is to incorporate pessimistic beliefs in the learning process, which will generate even richer asset pricing dynamics. It may be also interesting to extend the model to three countries.
Appendix A. Kalman filter Derivation for the Information Structure of Learning from Consumption Stream

The consumption and dividend processes are assumed to be the following:

\begin{align}
\Delta c_t &= \mu + z_t + \sigma_{c,h} \cdot \varepsilon_{c,t} \\
\Delta c^*_t &= \mu + z^*_t + \sigma_{c,f} \cdot \varepsilon^*_{c,t} \\
z_t &= \rho \cdot z_{t-1} + \sigma_{z,h} \cdot \varepsilon_{z,t} \\
z^*_t &= \rho \cdot z^*_{t-1} + \sigma_{z,f} \cdot \varepsilon^*_{z,t} \\
\Delta d_t &= \mu_d + \lambda \cdot (z_t + \sigma_{c,h} \cdot \varepsilon_{c,t}) \\
\Delta d^*_t &= \mu_d + \lambda \cdot (z^*_t + \sigma_{c,f} \cdot \varepsilon^*_{c,t})
\end{align}

\[
\text{Var} \begin{pmatrix}
\sigma_{c,h} \cdot \varepsilon_{c,t} \\
\sigma_{c,f} \cdot \varepsilon^*_{c,t} \\
\sigma_{z,h} \cdot \varepsilon_{z,t} \\
\sigma_{z,f} \cdot \varepsilon^*_{z,t}
\end{pmatrix} = \begin{bmatrix}
\sigma_{c,h}^2 & \rho_c \cdot \sigma_{c,h} \sigma_{c,f} & 0 & 0 \\
\rho_c \cdot \sigma_{c,h} \sigma_{c,f} & \sigma_{c,f}^2 & 0 & 0 \\
0 & 0 & \sigma_{z,h}^2 & \rho_z \cdot \sigma_{z,h} \sigma_{z,f} \\
0 & 0 & \rho_z \cdot \sigma_{z,h} \sigma_{z,f} & \sigma_{z,f}^2
\end{bmatrix}
\]

Define:

\[
A^c = \begin{bmatrix}
\rho & 0 \\
0 & \rho
\end{bmatrix} \quad H^c = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}
\]

\[
Q^c = \begin{bmatrix}
\sigma_{c,h}^2 & \rho_c \cdot \sigma_{c,h} \sigma_{c,f} \\
\rho_c \cdot \sigma_{c,h} \sigma_{c,f} & \sigma_{c,f}^2
\end{bmatrix} \quad R^c = \begin{bmatrix}
\sigma_{z,h}^2 & \rho_z \cdot \sigma_{z,h} \sigma_{z,f} \\
\rho_z \cdot \sigma_{z,h} \sigma_{z,f} & \sigma_{z,f}^2
\end{bmatrix}
\]

The state vector is \( \begin{pmatrix} z_t \\ z^*_t \end{pmatrix} \).

The measurement vector is \( \begin{pmatrix} \Delta c_t \\ \Delta c^*_t \end{pmatrix} \).

The filtered state vector for the case of learning from consumption process is \( \begin{pmatrix} \hat{z}_t \\ \hat{z}^*_t \end{pmatrix} \).

The variance covariance matrix of the filtering errors is:

\[
P^c_t = \begin{bmatrix}
Phh_t & Phf_t^c \\
Phf^c_t & Pf_f^c_t
\end{bmatrix}
\]

Then by applying the standard Kalman filter update equation

\[
P^c_t = A^c P^c_{t-1} A^c' - A^c P^c_{t-1} H^c [H^c P^c_{t-1} H^c' + R^c]^{-1} H^c P^c_{t-1} A^c' + Q^c
\]

One can show the one-step-ahead evolution equations for the variances of the filtering errors are:
We have,

\[ \nu = \nu_{LC,t} \]

The innovations in consumption in home and foreign countries are defined as:

\[ \left\{ \begin{array}{l}
\nu_{c,t} = \sigma_{c,h} \cdot \varepsilon_{c,t} + z_t - \hat{z}_t \\
\nu^*_{c,t} = \sigma_{c,f} \cdot \varepsilon^*_{c,t} + z^*_t - \hat{z}^*_t
\end{array} \right. \]

We have,

\[ \Delta c_t = \mu + \hat{z}_t + \nu_{c,t} \]
\[ \Delta c^*_t = \mu + \hat{z}^*_t + \nu^*_{c,t} \]
\[ \Delta d_t = \mu_d + \lambda \cdot (\Delta c_t - \mu) \]
\[ \Delta d^*_t = \mu_d + \lambda \cdot (\Delta c^*_t - \mu) \]
And the one-step-ahead state evolution equations for the filtered home and foreign long-run persistent components are:

\[
\begin{align*}
\hat{z}_t &= \rho \cdot \hat{z}_{t-1} + K_{11}^c \cdot \nu_{c,t} + K_{12}^c \cdot \nu^*_{c,t} \\
\hat{z}^*_t &= \rho \cdot \hat{z}^*_{t-1} + K_{21}^c \cdot \nu_{c,t} + K_{22}^c \cdot \nu^*_{c,t}
\end{align*}
\]  

(A.16)  
(A.17)  

The steady state Kalman filter is the solution to the following Discrete Algebraic Riccati Equation:

\[
A_c^e . P_{ss}^e . A_c^{eT} - A_c^e . P_{ss}^e . H_c^{eT} . \left[ H_c^e . P_{ss}^e . H_c^{eT} + R_c^e \right]^{-1} . H_c^e . P_{ss}^e . A_c^{eT} + Q_c^e
\]  

(A.18)
Appendix B. Kalman filter Derivation for the Information Structure of Learning from Consumption Stream and Dividend Stream

The consumption and dividend processes are assumed to be the following:

(B.1) \[ \begin{align*}
\Delta c_t &= \mu + z_t + \sigma_{c,h} \cdot \varepsilon_{c,t} \\
\Delta c^*_t &= \mu + z^*_t + \sigma_{c,f} \cdot \varepsilon^*_{c,t} \\
z_t &= \rho \cdot z_{t-1} + \sigma_{z,h} \cdot \varepsilon_{z,t} \\
z^*_t &= \rho \cdot z^*_{t-1} + \sigma_{z,f} \cdot \varepsilon^*_{z,t}
\end{align*} \]

(B.2) \[ \begin{align*}
\Delta d_t &= \mu + \lambda \cdot z_t + \sigma_{d,h} \cdot \varepsilon_{d,t} \\
\Delta d^*_t &= \mu + \lambda \cdot z^*_t + \sigma_{d,f} \cdot \varepsilon^*_{d,t}
\end{align*} \]

Ω = Var \[ \begin{pmatrix}
\sigma_{c,h} \cdot \varepsilon_{c,t} \\
\sigma_{c,f} \cdot \varepsilon^*_{c,t} \\
\sigma_{z,h} \cdot \varepsilon_{z,t} \\
\sigma_{z,f} \cdot \varepsilon^*_{z,t} \\
\sigma_{d,h} \cdot \varepsilon_{d,t} \\
\sigma_{d,f} \cdot \varepsilon^*_{d,t}
\end{pmatrix} = \\
\begin{bmatrix}
\sigma_{c,h}^2 & \rho c \sigma_{c,h} \sigma_{c,f} & 0 & 0 & 0 & 0 \\
\rho c \sigma_{c,h} \sigma_{c,f} & \sigma_{c,f}^2 & 0 & 0 & 0 & 0 \\
0 & 0 & \sigma_{z,h}^2 & \rho z \sigma_{z,h} \sigma_{z,f} & 0 & 0 \\
0 & 0 & \rho z \sigma_{z,h} \sigma_{z,f} & \sigma_{z,f}^2 & 0 & 0 \\
0 & 0 & 0 & 0 & \sigma_{d,h}^2 & \rho d \sigma_{d,h} \sigma_{d,f} \\
0 & 0 & 0 & 0 & \rho d \sigma_{d,h} \sigma_{d,f} & \sigma_{d,f}^2
\end{bmatrix}
\]

Define:

\[ A^d = \begin{bmatrix} \rho & 0 \\ 0 & \rho \end{bmatrix}, \quad H^d = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \]

\[ Q^d = \begin{bmatrix} \sigma_{c,h}^2 & \rho c \cdot \sigma_{c,h} \sigma_{c,f} \\ \rho c \cdot \sigma_{c,h} \sigma_{c,f} & \sigma_{c,f}^2 \end{bmatrix}, \quad P^d = \begin{bmatrix} \sigma_{z,h}^2 & \rho z \cdot \sigma_{z,h} \sigma_{z,f} & 0 & 0 \\
\rho z \cdot \sigma_{z,h} \sigma_{z,f} & \sigma_{z,f}^2 & 0 & 0 \\
0 & 0 & \sigma_{d,h}^2 & \rho d \sigma_{d,h} \sigma_{d,f} \\
0 & 0 & \rho d \sigma_{d,h} \sigma_{d,f} & \sigma_{d,f}^2
\end{bmatrix} \]

The state vector is \[ \begin{pmatrix} z_t \\ z^*_t \end{pmatrix} \].

The measurement vector is \[ \begin{pmatrix} \Delta c_t \\ \Delta c^*_t \\ \Delta d_t \\ \Delta d^*_t \end{pmatrix} \].

The filtered state vector for the case of learning from both consumption and dividend processes is \[ \begin{pmatrix} \tilde{z}_t \\ \tilde{z}^*_t \end{pmatrix} \]. Both the consumption process and the dividend process provide information for the filtering of the state vector.

The variance covariance matrix of the filtering errors is:
\[
P_t^d = \begin{bmatrix}
Ph_{f_t}^d & Ph_{f_t}^d \\
Ph_{h_t}^d & Pf_{f_t}^d
\end{bmatrix}
\]

Then by applying the standard Kalman filter update equation

\[
P_t^d = A_t^d P_{t-1}^d A_t^d - A_t^d P_{t-1}^d H_t^d \left[ H_t^d P_{t-1}^d H_t^d + R_t^d \right]^{-1} H_t^d P_{t-1}^d A_t^d + Q_t^d
\]

One can show the one-step-ahead evolution equations for the variances of the filtering errors are:

(B.4)

\[
P_{h_t}^d = \left[ P_{f f_{t-1}}^d \left( \rho_{h_t}^2 \sigma_{d,h}^2 \sigma_{c,h}^2 - \left( \rho_{d}^2 - 1 \right) \lambda^2 \sigma_{c,f}^2 - \left( \rho_{d}^2 - 1 \right) \sigma_{d,h}^2 \right) + \rho_{h_t}^2 \left( \rho_{h_t}^2 - 1 \right) \lambda^2 \sigma_{c,h}^2 \left( \rho_{c,h}^2 + \sigma_{h}^2 \right) - 2 \rho_{c,h} \rho_{h_t} \sigma_{c,h} \sigma_{c,f} \sigma_{c,h} \left( \rho_{h_t}^2 - 1 \right) \lambda^2 \sigma_{c,f} \sigma_{c,h} - \left( \rho_{d}^2 - 1 \right) \sigma_{d,h}^2 \sigma_{h}^2 \right) \right] + Ph_{h_t}^d + Pf_{f_t}^d \left( \rho_{h_t}^2 \sigma_{d,h}^2 \sigma_{c,h}^2 - \left( \rho_{d}^2 - 1 \right) \lambda^2 \sigma_{c,f}^2 - \left( \rho_{d}^2 - 1 \right) \sigma_{d,h}^2 \right)
\]

(B.5)

\[
Ph_{f_t}^d = \left[ \rho_{c,h} \rho_{d,h} \sigma_{c,h} \left( \rho_{d}^2 - 1 \right) \sigma_{d,h}^2 \sigma_{c,h}^2 \left( \rho_{c,h}^2 + \sigma_{h}^2 \right) + Pf_{f_{t-1}}^{d} \left( \rho_{h_t}^2 \sigma_{d,h}^2 \sigma_{c,h}^2 - \left( \rho_{d}^2 - 1 \right) \lambda^2 \sigma_{c,f}^2 - \left( \rho_{d}^2 - 1 \right) \sigma_{d,h}^2 \sigma_{h}^2 \right) \right] + Ph_{f_t}^d + Pf_{f_t}^d \left( \rho_{h_t}^2 \sigma_{d,h}^2 \sigma_{c,h}^2 - \left( \rho_{d}^2 - 1 \right) \lambda^2 \sigma_{c,f}^2 - \left( \rho_{d}^2 - 1 \right) \sigma_{d,h}^2 \sigma_{h}^2 \right)
\]
\[ P_{ff}^d = P_{ff}^{d-1} \left( \sigma_{d,f}^2 \sigma_{c,f}^2 \left( - \left( \rho_d^2 - 1 \right) \sigma_{d,h}^2 - \left( \rho_c^2 \sigma_{c,h}^2 + \Delta_{ff} \right) \right) - \left( \rho_d^2 - 1 \right) \lambda^2 \right) \]

\[ + \sigma_{c,f}^2 \left( \sigma_{d,h}^2 \left( - \left( \rho_d^2 - 1 \right) \sigma_{d,h}^2 + \Delta_{hh} \right) - \left( \rho_d^2 - 1 \right) \lambda^2 \right) \]

One can derive the 2 by 4 Kalman filter in this case:

\[ K^d = \begin{bmatrix} K_{11}^d & K_{12}^d & K_{13}^d & K_{14}^d \\ K_{21}^d & K_{22}^d & K_{23}^d & K_{24}^d \end{bmatrix} \]

where

\[ K_{11}^d = \sigma_{d,h} \left( \left( \sigma_{d,h}^2 - 1 \right) \sigma_{d,h} \left( \Delta_{ff} \right) \right) + \lambda^2 \]

\[ K_{12}^d = \left( \sigma_{d,h}^2 - 1 \right) \lambda^2 \]

\[ K_{13}^d = K_{14}^d = 0 \]

\[ \sigma_{d,h} \left( \left( \sigma_{d,h}^2 - 1 \right) \left( \Delta_{ff} \right) \right) + \lambda^2 \]

\[ \sigma_{d,h} \left( \left( \sigma_{d,h}^2 - 1 \right) \left( \Delta_{ff} \right) \right) + \lambda^2 \]

\[ \sigma_{d,h} \left( \left( \sigma_{d,h}^2 - 1 \right) \left( \Delta_{ff} \right) \right) + \lambda^2 \]

\[ \sigma_{d,h} \left( \left( \sigma_{d,h}^2 - 1 \right) \left( \Delta_{ff} \right) \right) + \lambda^2 \]
\[ K_{15}^{d} = \left[ \sigma_{c,h} \lambda \left( (\rho_{c}^{2} - 1) \left( \begin{array}{c} \text{B.9} \\ \text{B.10} \\ \text{B.11} \\
 \end{array} \right) \right) \sigma_{c,h} \sigma_{f} \right] \]

\[ = \left(\rho_{c}^{2} - 1 \right) \left( \begin{array}{c} \text{B.9} \\
 \end{array} \right) \left[ \left( \begin{array}{c} \text{B.10} \\
 \end{array} \right) \right] \left( \begin{array}{c} \text{B.11} \\
 \end{array} \right) \]

\[ = \left(\rho_{c}^{2} - 1 \right) \left( \begin{array}{c} \text{B.9} \\
 \end{array} \right) \left[ \left( \begin{array}{c} \text{B.10} \\
 \end{array} \right) \right] \left( \begin{array}{c} \text{B.11} \\
 \end{array} \right) \]

\[ = \left(\rho_{c}^{2} - 1 \right) \left( \begin{array}{c} \text{B.9} \\
 \end{array} \right) \left[ \left( \begin{array}{c} \text{B.10} \\
 \end{array} \right) \right] \left( \begin{array}{c} \text{B.11} \\
 \end{array} \right) \]

\[ = \left(\rho_{c}^{2} - 1 \right) \left( \begin{array}{c} \text{B.9} \\
 \end{array} \right) \left[ \left( \begin{array}{c} \text{B.10} \\
 \end{array} \right) \right] \left( \begin{array}{c} \text{B.11} \\
 \end{array} \right) \]

\[ = \left(\rho_{c}^{2} - 1 \right) \left( \begin{array}{c} \text{B.9} \\
 \end{array} \right) \left[ \left( \begin{array}{c} \text{B.10} \\
 \end{array} \right) \right] \left( \begin{array}{c} \text{B.11} \\
 \end{array} \right) \]

\[ = \left(\rho_{c}^{2} - 1 \right) \left( \begin{array}{c} \text{B.9} \\
 \end{array} \right) \left[ \left( \begin{array}{c} \text{B.10} \\
 \end{array} \right) \right] \left( \begin{array}{c} \text{B.11} \\
 \end{array} \right) \]

\[ = \left(\rho_{c}^{2} - 1 \right) \left( \begin{array}{c} \text{B.9} \\
 \end{array} \right) \left[ \left( \begin{array}{c} \text{B.10} \\
 \end{array} \right) \right] \left( \begin{array}{c} \text{B.11} \\
 \end{array} \right) \]

\[ = \left(\rho_{c}^{2} - 1 \right) \left( \begin{array}{c} \text{B.9} \\
 \end{array} \right) \left[ \left( \begin{array}{c} \text{B.10} \\
 \end{array} \right) \right] \left( \begin{array}{c} \text{B.11} \\
 \end{array} \right) \]

\[ = \left(\rho_{c}^{2} - 1 \right) \left( \begin{array}{c} \text{B.9} \\
 \end{array} \right) \left[ \left( \begin{array}{c} \text{B.10} \\
 \end{array} \right) \right] \left( \begin{array}{c} \text{B.11} \\
 \end{array} \right) \]

\[ = \left(\rho_{c}^{2} - 1 \right) \left( \begin{array}{c} \text{B.9} \\
 \end{array} \right) \left[ \left( \begin{array}{c} \text{B.10} \\
 \end{array} \right) \right] \left( \begin{array}{c} \text{B.11} \\
 \end{array} \right) \]

\[ = \left(\rho_{c}^{2} - 1 \right) \left( \begin{array}{c} \text{B.9} \\
 \end{array} \right) \left[ \left( \begin{array}{c} \text{B.10} \\
 \end{array} \right) \right] \left( \begin{array}{c} \text{B.11} \\
 \end{array} \right) \]

\[ = \left(\rho_{c}^{2} - 1 \right) \left( \begin{array}{c} \text{B.9} \\
 \end{array} \right) \left[ \left( \begin{array}{c} \text{B.10} \\
 \end{array} \right) \right] \left( \begin{array}{c} \text{B.11} \\
 \end{array} \right) \]

\[ = \left(\rho_{c}^{2} - 1 \right) \left( \begin{array}{c} \text{B.9} \\
 \end{array} \right) \left[ \left( \begin{array}{c} \text{B.10} \\
 \end{array} \right) \right] \left( \begin{array}{c} \text{B.11} \\
 \end{array} \right) \]

\[ = \left(\rho_{c}^{2} - 1 \right) \left( \begin{array}{c} \text{B.9} \\
 \end{array} \right) \left[ \left( \begin{array}{c} \text{B.10} \\
 \end{array} \right) \right] \left( \begin{array}{c} \text{B.11} \\
 \end{array} \right) \]
The innovations in consumption and dividend, in home and foreign countries are defined as:

\[
\begin{align*}
K_{22}^d &= \left[ \sigma_{d,f} \left( (\rho_d^2 - 1) \sigma_{d,f} \sigma_{d,h}^2 \left( Pf_{f,t-1}^d + \rho_c \sigma_{c,h} \sigma_{c,f} Pf_{f,t-1}^d + Pf_{f,t-1}^d \left( (\sigma_{c,h}^2 + Pf_{h,t-1}^d) \right) \right) \right) \\
&\quad - \left( Pf_{f,t-1}^d \sigma_{d,f} \sigma_{d,h} \sigma_{c,h} \sigma_{c,f} Pf_{h,t-1}^d \right) \right] \\
&\quad + \left[ (\rho_d^2 - 1) Pf_{f,t-1}^d \sigma_{d,f} \sigma_{d,h} \sigma_{c,h} \sigma_{c,f} Pf_{h,t-1}^d \right] \right]
\end{align*}
\]

\[
\begin{align*}
K_{23}^d &= \left[ \sigma_{d,f} \sigma_{c,f} \left( Pf_{f,t-1}^d - Pf_{f,t-1}^d \right) \sigma_{d,f} \sigma_{d,h} \sigma_{c,h} \\
&\quad + \left( (\rho_d^2 - 1) Pf_{f,t-1}^d \sigma_{d,f} \sigma_{d,h} \sigma_{c,h} \sigma_{c,f} Pf_{h,t-1}^d \right) \right] \\
&\quad + \left[ (\rho_d^2 - 1) Pf_{f,t-1}^d \sigma_{d,f} \sigma_{d,h} \sigma_{c,h} \sigma_{c,f} Pf_{h,t-1}^d \right] \right]
\end{align*}
\]

\[
\begin{align*}
K_{24}^d &= \left[ \sigma_{c,f} \lambda \left( Pf_{f,t-1}^d \sigma_{c,f} Pf_{h,t-1}^d \right) \sigma_{c,h} \sigma_{c,f} Pf_{h,t-1}^d \right] \\
&\quad + \left[ (\rho_d^2 - 1) Pf_{f,t-1}^d \sigma_{d,f} \sigma_{d,h} \sigma_{c,h} \sigma_{c,f} Pf_{h,t-1}^d \right] \right]
\end{align*}
\]

Or, in the innovation presentation, let \( \nu_{LC,D,t} = \begin{pmatrix} \nu_{c,t}^d \\ \nu_{c,t}^p \\ \nu_{d,t}^d \\ \nu_{d,t}^p \end{pmatrix} \)

The innovations in consumption and dividend, in home and foreign countries are defined as:
\[
\begin{align*}
\nu_{c,t} &= \sigma_{c,h} \cdot \varepsilon_{c,t} + z_t - \tilde{z}_t \\
\nu^*_{c,t} &= \sigma_{c,f} \cdot \varepsilon^*_{c,t} + z^*_t - \tilde{z}^*_t \\
\nu_{d,t} &= \sigma_{d,h} \cdot \varepsilon_{d,t} + \lambda \cdot (z_t - \tilde{z}_t) \\
\nu^*_{d,t} &= \sigma_{d,f} \cdot \varepsilon^*_{d,t} + \lambda \cdot (z^*_t - \tilde{z}^*_t)
\end{align*}
\] (B.15)

We have,

\[
\begin{align*}
\Delta c_t &= \mu + \tilde{z}_t + \nu_{c,t} \\
\Delta c^*_t &= \mu + \tilde{z}^*_t + \nu^*_{c,t} \\
\Delta d_t &= \mu_d + \lambda \cdot \tilde{z}_t + \nu_{d,t} \\
\Delta d^*_t &= \mu_d + \lambda \cdot \tilde{z}^*_t + \nu^*_{d,t}
\end{align*}
\] (B.16-19)

And the one-step-ahead state evolution equations for the filtered home and foreign long-run persistent components are:

\[
\begin{align*}
\tilde{z}_t &= \rho \cdot \tilde{z}_{t-1} + K^d_{11} \cdot \nu_{c,t} + K^d_{12} \cdot \nu^*_{c,t} + K^d_{13} \cdot \nu_{d,t} + K^d_{14} \cdot \nu^*_{d,t} \\
\tilde{z}^*_t &= \rho \cdot \tilde{z}^*_{t-1} + K^d_{21} \cdot \nu_{c,t} + K^d_{22} \cdot \nu^*_{c,t} + K^d_{23} \cdot \nu_{d,t} + K^d_{24} \cdot \nu^*_{d,t}
\end{align*}
\] (B.20-21)

The steady state Kalman filter is the solution to the following Discrete Algebraic Riccati Equation:

\[
A^d \cdot P_{ss}^d \cdot A^d^T - A^d \cdot P_{ss}^d \cdot H^d^T \cdot \left[ H^d \cdot P_{ss}^d \cdot H^d^T + R^d \right]^{-1} \cdot H^d \cdot P_{ss}^d \cdot A^d^T + Q^d
\] (B.22)
Appendix C. Derivation of Pricing Kernel

In this section I elaborate the asset pricing results obtained from log-linearization of the Epstein-Zin Utility. This exercise was performed to provide some intuition through analytical solutions. I used numerical third order approximations in the simulations and plots and did not use the analytical first order approximation results presented in this section.

C1. Cash Flow Model

The baseline linear state space representation is

\[
\begin{align*}
\Delta c_t &= \mu + z_t + \eta_{c,t} \\
\Delta c^{*}_t &= \mu + z^{*}_t + \eta^{*}_{c,t} \\
z_t &= \rho \cdot z_{t-1} + \eta_{z,t} \\
z^{*}_t &= \rho \cdot z^{*}_{t-1} + \eta^{*}_{z,t} \\
\Delta d_t &= \mu_d + \lambda \cdot z_t + \eta_{d,t} \\
\Delta d^{*}_t &= \mu_d + \lambda \cdot z^{*}_t + \eta^{*}_{d,t}
\end{align*}
\]

(C.1)

where

\[
\eta_t = \begin{pmatrix} 
\eta_{c,t} \\
\eta^{*}_{c,t} \\
\eta_{z,t} \\
\eta^{*}_{z,t} \\
\eta_{d,t} \\
\eta^{*}_{d,t}
\end{pmatrix} \sim N.i.i.d.(0, S)
\]

(C.2)

C2. Log Linearization

I follow (Croce et al., 2012). This implementation involves two-countries and the results are similar to the standard long-run risk literature, up to a difference in the timing convention in the long-run persistent component. Define the price dividend ratio of an asset that pays a consumption stream \(C_t\) at end of period \(t\) as \(W_{c,t} = P^{C}_t/C_t\), \(R_{c,t+1} = (P^{C}_{t+1} + C_{t+1})/P^{C}_t\), then the Campbell-Shiller log linearization yields:

\[
\begin{align*}
\log w_{c,t} &= \log w_{c} + \sum_{i=0}^{\infty} \kappa_c E_t[\Delta c_{t+1+i}] - \sum_{i=0}^{\infty} \kappa^i c E_t[\Delta r_{c,t+1+i}]
\end{align*}
\]

(C.3)

\[
\kappa_c = \frac{\exp(w_c)}{1 + \exp(w_c)}
\]

(C.4)

The first order condition of the Epstein-Zin Utility yields:
\[ m_{t+1} = \bar{m} - \frac{1}{\psi} z_{t+1} - \kappa_c \frac{\gamma - 1/\psi}{1 - \rho \kappa_c} \eta_{z,t+1} - \gamma \eta_{c,t+1} \]
\[ r_{c,t+1} = \bar{r}_c + \frac{1}{\psi} z_{t+1} + \kappa_c \frac{1 - 1/\psi}{1 - \rho \kappa_c} \eta_{z,t+1} + \eta_{c,t+1} \]
\[ r_{f,t} = \bar{r}_f + \frac{1}{\psi} z_{t+1} \]
\[ w_{c,t} = \bar{w}_c + \frac{1 - 1/\psi}{1 - \kappa_c \rho} z_{t+1} \]

Define the price dividend ratio of an asset that pays a consumption stream \( D_t \) at end of period \( t \) as \( W_{d,t} = P_{D_t} / D_t \), \( R_{d,t+1} = (P_{t+1}^D + D_{t+1}) / P_t^D \), then the Campbell-Shiller log linearization yields:

\[ w_{d,t} = \bar{w}_d + \frac{\lambda - 1/\psi}{1 - \kappa_d \rho} z_{t+1} \]
\[ \kappa_d = \frac{\exp(w_d)}{1 + \exp(w_d)} \]
\[ r_{d,t+1} = \bar{r}_d + \frac{1}{\psi} z_{t+1} + \kappa_d \frac{\lambda - 1/\psi}{1 - \rho \kappa_d} \eta_{z,t+1} + \eta_{d,t+1} \]

Thus, in vector form, we have that for the home country:

\[ m_{t+1} = \bar{m} - \frac{1}{\psi} z_{t+1} + \Gamma_m \eta_{t+1} \]
\[ \Gamma_m = \begin{bmatrix} -\gamma & 0 & -\kappa_c \frac{\gamma - 1/\psi}{1 - \kappa_c \rho} & 0 & 0 & 0 \end{bmatrix} \]
\[ r_{c,t+1} = \bar{r}_c + \frac{1}{\psi} z_{t+1} + \Gamma_c \eta_{t+1} \]
\[ \Gamma_c = \begin{bmatrix} 1 & 0 & \kappa_c \frac{1 - 1/\psi}{1 - \kappa_c \rho} & 0 & 0 & 0 \end{bmatrix} \]
\[ r_{d,t+1} = \bar{r}_d + \frac{1}{\psi} z_{t+1} + \Gamma_d \eta_{t+1} \]
\[ \Gamma_d = \begin{bmatrix} 0 & 0 & \kappa_d \frac{\lambda - 1/\psi}{1 - \kappa_d \rho} & 0 & 1 & 0 \end{bmatrix} \]
Similarly, for the foreign country:

\[ m_{t+1}^* = \bar{m}^* - \frac{1}{\psi^*} \eta_{t+1}^* + \Gamma_m^* \eta_{t+1} \]  
\[ \Gamma_m^* = \begin{bmatrix} 0 & -\gamma & 0 & -\kappa_c \frac{1}{1-\kappa_c \rho} & 0 & 0 \end{bmatrix} \]  
\[ r_{c,t+1}^* = \bar{r}_c^* + \frac{1}{\psi} z_{c,t+1}^* + \Gamma_c^* \eta_{t+1} \]  
\[ \Gamma_c^* = \begin{bmatrix} 0 & 1 & 0 & \kappa_c \frac{1}{1-\kappa_c \rho} & 0 & 0 \end{bmatrix} \]  
\[ r_{d,t+1}^* = \bar{r}_d^* + \frac{1}{\psi} z_{d,t+1}^* + \Gamma_d^* \eta_{t+1} \]  
\[ \Gamma_d^* = \begin{bmatrix} 0 & 0 & 0 & \kappa_d \lambda \frac{1}{1-\kappa_d \rho} & 0 & 1 \end{bmatrix} \]  

Since

\[ E_t[r_{c,t+1}^{ex}] = -\text{cov} (m_{t+1} - E_t[m_{t+1}], r_{c,t+1} - E_t[r_{c,t+1}]) - \frac{1}{2} \text{Var} (r_{c,t+1} - E_t[r_{c,t+1}]) \]  
\[ E_t[r_{d,t+1}^{ex}] = -\text{cov} (m_{t+1} - E_t[m_{t+1}], r_{d,t+1} - E_t[r_{d,t+1}]) - \frac{1}{2} \text{Var} (r_{d,t+1} - E_t[r_{d,t+1}]) \]

We have

\[ E_t[r_{c,t+1}^{ex}] = -\Gamma_m^* \Gamma_c' - \frac{1}{2} \Gamma_c^* \Gamma_c' \]  
\[ E_t[r_{c,t+1}^{ex*}] = -\Gamma_m^* \Gamma_c' - \frac{1}{2} \Gamma_c^* \Gamma_c' \]  
\[ E_t[r_{d,t+1}^{ex}] = -\Gamma_m^* \Gamma_d' - \frac{1}{2} \Gamma_d^* \Gamma_d' \]  
\[ E_t[r_{d,t+1}^{ex*}] = -\Gamma_m^* \Gamma_d' - \frac{1}{2} \Gamma_d^* \Gamma_d' \]  

C3. Coefficients

By definition of the pricing kernel:

\[ E[r_f] = -\log \delta + \frac{1}{\psi} \mu + \frac{1 - \theta}{\theta} \left( -\Gamma_m^* \Gamma_c' - \frac{1}{2} \Gamma_c^* \Gamma_c' \right) - \frac{1}{2 \theta} \Gamma_m^* \Gamma_m' \]

where \( \theta = \frac{1 - \gamma}{1 - 1/\psi} \)

Thus, the intercept can be shown to be

\[ \bar{m} = \theta \log \delta - \frac{\theta}{\psi} \mu + (\theta - 1)(E[r_c^{ex}] + E[r_f]) \]
Since the Euler Equations holds for all values of the long-run persistent component, plug-in the case $z_{t+1} = 0$ can pin down expressions for $\kappa_c$ and $\kappa_d$

\begin{align*}
\kappa_c &= \delta e \left(1 - \frac{1}{\psi}\right) \left(\mu - \frac{1}{2} \gamma - 1\right) \text{Var} \left[ \eta_{c,t+1} + \frac{\kappa_c \eta_{z,t+1}}{1 - \kappa_c p} \right] \\
\kappa_d &= e^{\rho + \frac{1}{2}} \text{Var} \left[ (\Gamma_m + \Gamma_d) \eta_{t+1} \right]
\end{align*}
Appendix D. Mapping of Information Structure

The asset pricing results obtained from log-linearization of Epstein-Zin preference shown in Appendix C can be readily applied to different information structures, as long as they are expressed in terms of the baseline linear state space representation of equation (C.1). To achieve this, I derive the mappings from innovation space representation into the baseline linear state space representation of consumption, long-run component, and dividend shocks.

D1. Full Information

It is trivial to transform the full information case into the baseline linear state space representation

\[
\begin{pmatrix}
\eta_{c,t} \\
\eta^*_{c,t} \\
\eta_{z,t} \\
\eta^*_{z,t} \\
\eta_{d,t} \\
\eta^*_{d,t}
\end{pmatrix} =
\begin{pmatrix}
\sigma_{c,h} & 0 & 0 & 0 & 0 & 0 \\
0 & \sigma_{c,f} & 0 & 0 & 0 & 0 \\
0 & 0 & \sigma_{z,h} & 0 & 0 & 0 \\
0 & 0 & 0 & \sigma_{z,f} & 0 & 0 \\
0 & 0 & 0 & 0 & \sigma_{d,h} & 0 \\
0 & 0 & 0 & 0 & 0 & \sigma_{d,f}
\end{pmatrix}
\begin{pmatrix}
\varepsilon_{c,t} \\
\varepsilon^*_{c,t} \\
\varepsilon_{z,t} \\
\varepsilon^*_{z,t} \\
\varepsilon_{d,t} \\
\varepsilon^*_{d,t}
\end{pmatrix}
\]

In other words

\[
\eta_t = \Sigma_{FI}\varepsilon_t
\]

(D.2)

\[
S_{FI} = \Omega
\]

(D.3)

D2. Learning from Consumption Stream

The following mapping allows the transformation of the innovation space representation for the case of learning from consumption stream into the baseline linear state space representation.

\[
\begin{pmatrix}
\eta_{c,t} \\
\eta^*_{c,t} \\
\eta_{z,t} \\
\eta^*_{z,t} \\
\eta_{d,t} \\
\eta^*_{d,t}
\end{pmatrix} =
\begin{pmatrix}
1 & 0 \\
0 & 1 \\
K_{11}^c & K_{12}^c \\
K_{21}^c & K_{22}^c \\
\lambda & 0 \\
0 & \lambda
\end{pmatrix}
\begin{pmatrix}
\nu_{c,t} \\
\nu^*_{c,t}
\end{pmatrix}
\]

In other words

\[
\eta_t = \Sigma_{LC}\nu_{LC,t}
\]

(D.5)

Thus

\[
S_{LC} = \Sigma_{LC}P_{LC}\Sigma_{LC}^\prime
\]

(D.6)

where \(K^c_{ij}\) are the elements Kalman Gain matrix for the case of learning from consumption stream. \(P_{LC} = E[\nu_{LC}\nu_{LC}^\prime]\) is a non-linear transformation of the \(\Omega\) matrix obtained by solving the steady state Kalman filtering problem.
D3. Learning from Consumption Stream and Dividend Stream

The following mapping allows the transformation of the innovation space representation for the case of learning from consumption stream and dividend stream into the baseline linear state space representation.

\[
\begin{pmatrix}
\eta_{c,t} \\
\eta^*_{c,t} \\
\eta_{z,t} \\
\eta^*_z,t \\
\eta_{d,t} \\
\eta^*_d,t
\end{pmatrix}
= \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
K_{d11} & K_{d12} & K_{d13} & K_{d14} \\
K_{d21} & K_{d22} & K_{d23} & K_{d24} \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
\nu_{c,t} \\
\nu^*_{c,t} \\
\nu_{d,t} \\
\nu^*_{d,t}
\end{pmatrix}
\]

In other words

\[
\eta_t = \Sigma_{LCD}\nu_{LCD,t}
\]

Thus

\[
S_{LCD} = \Sigma_{LCD}P_{LCD}\Sigma'_{LCD}
\]

where \(K_{ij}\) are the elements Kalman Gain matrix for the case of learning from consumption and dividend stream. \(P_{LCD} = E[\nu_{LCD}\nu'_{LCD}]\) is a non-linear transformation of the \(\Omega\) matrix obtained by solving the steady state Kalman filtering problem.
Figures

Figure 1. Impulse responses of various variables to temporary home and foreign shocks implied in the information structure of Full Information. Each shock is given one standard deviation of positive impulse. The first column is short-run consumption shock. The second column is shock to the long-run persistent component. The third column is dividend shock. $z_t$ is the long-run persistent component; $\Delta \ln C_t$ is log consumption growth; $\Delta \ln D_t$ is log dividend growth; $m$ is pricing kernel; $\Delta e_t$ is log exchange rate growth; $r_{f,t}$ is log risk free rate; $r_{d,t}$ is log return on dividend stream; $r_{ex,t}$ is equity risk premium; $\ln P_t C_t$ is the log price consumption ratio; $r_{c,t}$ is the log return on the asset which pays the consumption stream.
Figure 2. Impulse responses of various variables to temporary home and foreign shocks implied in the information structure of Learning from Consumption Stream. Each shock is given one standard deviation of positive impulse. The first column is short-run consumption shock. The second column is shock to the long-run persistent component. \( z_t \) is the long-run persistent component; \( \Delta \ln C_t \) is log consumption growth; \( \Delta \ln D_t \) is log dividend growth; \( m \) is pricing kernel; \( \Delta e_t \) is log exchange rate growth; \( rf,t \) is log risk free rate; \( rd,t \) is log return on dividend stream; \( rex,t \) is equity risk premium; \( \ln \frac{P}{C} \) is the log price consumption ratio; \( rc,t \) is the log return on the asset which pays the consumption stream.
Figure 3. Impulse responses of various variables to temporary home and foreign shocks implied in the information structure of Learning from Consumption and Dividend Stream. Each shock is given one standard deviation of positive impulse. The first column is short-run consumption shock. The second column is shock to the long-run persistent component. The third column is dividend shock. $z_t$ is the long-run persistent component; $\Delta \ln C_t$ is log consumption growth; $\Delta \ln D_t$ is log dividend growth; $m$ is pricing kernel; $\Delta e_t$ is log exchange rate growth; $r_{f,t}$ is log risk free rate; $r_{d,t}$ is log return on dividend stream; $r_{ex,t}$ is equity risk premium; $\ln \frac{P_t}{C_t}$ is the log price consumption ratio; $r_{c,t}$ is the log return on the asset which pays the consumption stream.
Figure 4. Impulse responses of various variables to temporary home and foreign shocks implied in the information structure of Learning from Consumption Stream. Each shock is given one standard deviation of positive impulse. The shock is consumption innovation shock. $z_t$ is the long-run persistent component; $\Delta \ln C_t$ is log consumption growth; $\Delta \ln D_t$ is log dividend growth; $m$ is pricing kernel; $\Delta \epsilon_t$ is log exchange rate growth; $r_{f,t}$ is log risk free rate; $r_{d,t}$ is log return on dividend stream; $r_{ex,t}$ is equity risk premium; $\ln \frac{P_t}{C_t}$ is the log price consumption ratio; $r_{c,t}$ is the log return on the asset which pays the consumption stream.
Figure 5. Impulse responses of various variables to temporary home and foreign shocks implied in the information structure of Learning from Consumption and Dividend Stream. Each shock is given one standard deviation of positive impulse. The first column is consumption innovation shock. The second column is dividend innovation shock. $z_t$ is the long-run persistent component; $\Delta \ln C_t$ is log consumption growth; $\Delta \ln D_t$ is log dividend growth; $m_t$ is pricing kernel; $\Delta e_t$ is log exchange rate growth; $r_{f,t}$ is log risk free rate; $r_{d,t}$ is log return on dividend stream; $r_{ext,t}$ is equity risk premium; $\ln \frac{P_t}{C_t}$ is the log price consumption ratio; $r_{ct,t}$ is the log return on the asset which pays the consumption stream.
### Tables

#### Table 1—Parameters used for Calibration

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persistence of long run shock $\rho$</td>
<td>0.98</td>
</tr>
<tr>
<td>Subjective discount factor $\delta$</td>
<td>0.9985</td>
</tr>
<tr>
<td>Co-integration parameter $\tau$</td>
<td>0.0005</td>
</tr>
<tr>
<td>Long run mean of Consumption Growth $\mu$</td>
<td>0.0015</td>
</tr>
<tr>
<td>Long run mean of Dividend Growth $\mu_d$</td>
<td>0.0007</td>
</tr>
<tr>
<td>Risk aversion $\gamma$</td>
<td>10</td>
</tr>
<tr>
<td>Intertemporal elasticity of substitution $\psi$</td>
<td>2</td>
</tr>
<tr>
<td>Leverage. Dividend process LongRun component multiple $\lambda$</td>
<td>5.5</td>
</tr>
<tr>
<td>Composite Parameter $\theta$</td>
<td>-0.055556</td>
</tr>
<tr>
<td>Ratio of Dividend shock and Short-Run shock volatilities $\varphi_d$</td>
<td>5</td>
</tr>
<tr>
<td>Correlation of Home/Foreign Short-Run shock $\rho(\varepsilon_{a,t}, \varepsilon^*_{a,t})$</td>
<td>0.35</td>
</tr>
<tr>
<td>Correlation of Home/Foreign Long-Run shock $\rho(\varepsilon_{z,t}, \varepsilon^*_{z,t})$</td>
<td>0.93</td>
</tr>
<tr>
<td>Correlation of Home/Foreign Dividend shock $\rho(\varepsilon_{d,t}, \varepsilon^*_{d,t})$</td>
<td>-0.1</td>
</tr>
<tr>
<td>Variance of Home Short-Run shock $\sigma^2_{a,t}$</td>
<td>8.28e-006</td>
</tr>
<tr>
<td>Covariance of Home/Foreign Short-Run shock $\sigma_{a,a^*,t}$</td>
<td>3.381e-006</td>
</tr>
<tr>
<td>Variance of Foreign Short-Run shock $\sigma^2_{a^*,t}$</td>
<td>1.127e-005</td>
</tr>
<tr>
<td>Variance of Home Long-Run shock $\sigma^2_{z,t}$</td>
<td>4.6656e-008</td>
</tr>
<tr>
<td>Covariance of Home/Foreign Long-Run shock $\sigma_{z,z^*,t}$</td>
<td>5.0622e-008</td>
</tr>
<tr>
<td>Variance of Foreign Long-Run shock $\sigma^2_{z^*,t}$</td>
<td>6.3504e-008</td>
</tr>
<tr>
<td>Variance of Home Dividend shock $\sigma^2_{d,t}$</td>
<td>0.000207</td>
</tr>
<tr>
<td>Covariance of Home/Foreign Dividend shock $\sigma_{d,d^*,t}$</td>
<td>-2.415e-005</td>
</tr>
<tr>
<td>Variance of Foreign Dividend shock $\sigma^2_{d^*,t}$</td>
<td>0.00028175</td>
</tr>
</tbody>
</table>
Table 2—Variance-Covariance Matrix for the Full Information Case

<table>
<thead>
<tr>
<th></th>
<th>$\varepsilon_{a,t}$</th>
<th>$\varepsilon_{a,t}^*$</th>
<th>$\varepsilon_{z,t}$</th>
<th>$\varepsilon_{z,t}^*$</th>
<th>$\varepsilon_{d,t}$</th>
<th>$\varepsilon_{d,t}^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon_{a,t}$</td>
<td>8.28e-006</td>
<td>3.381e-006</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\varepsilon_{a,t}^*$</td>
<td>3.381e-006</td>
<td>1.127e-005</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\varepsilon_{z,t}$</td>
<td>0</td>
<td>0</td>
<td>4.6656e-008</td>
<td>5.0622e-008</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\varepsilon_{z,t}^*$</td>
<td>0</td>
<td>0</td>
<td>5.0622e-008</td>
<td>6.3504e-008</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\varepsilon_{d,t}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.000207</td>
<td>-2.415e-005</td>
</tr>
<tr>
<td>$\varepsilon_{d,t}^*$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.00028175</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3—True Correlation Matrix for the Full Information Case

<table>
<thead>
<tr>
<th></th>
<th>$\varepsilon_{a,t}$</th>
<th>$\varepsilon_{a,t}^*$</th>
<th>$\varepsilon_{z,t}$</th>
<th>$\varepsilon_{z,t}^*$</th>
<th>$\varepsilon_{d,t}$</th>
<th>$\varepsilon_{d,t}^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon_{a,t}$</td>
<td>1</td>
<td>0.35</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\varepsilon_{a,t}^*$</td>
<td>0.35</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\varepsilon_{z,t}$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.93</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\varepsilon_{z,t}^*$</td>
<td>0</td>
<td>0</td>
<td>0.93</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\varepsilon_{d,t}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-0.1</td>
</tr>
<tr>
<td>$\varepsilon_{d,t}^*$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-0.1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4—Correlation Matrix for Mapped Shocks under Learning from Consumption Stream

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\varepsilon}_{a,t}$</th>
<th>$\hat{\varepsilon}_{a,t}^*$</th>
<th>$\hat{\varepsilon}_{z,t}$</th>
<th>$\hat{\varepsilon}_{z,t}^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\varepsilon}_{a,t}$</td>
<td>1</td>
<td>0.37639</td>
<td>0.88709</td>
<td>0.76154</td>
</tr>
<tr>
<td>$\hat{\varepsilon}_{a,t}^*$</td>
<td>0.37639</td>
<td>1</td>
<td>0.88709</td>
<td>0.76154</td>
</tr>
<tr>
<td>$\hat{\varepsilon}_{z,t}$</td>
<td>0.88709</td>
<td>0.76154</td>
<td>1</td>
<td>0.97473</td>
</tr>
<tr>
<td>$\hat{\varepsilon}_{z,t}^*$</td>
<td>0.76154</td>
<td>0.88709</td>
<td>0.97473</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 5—Correlation Matrix for Mapped Shocks under Learning from Consumption and Dividend Stream

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\varepsilon}_{a,t}$</th>
<th>$\hat{\varepsilon}_{a,t}^*$</th>
<th>$\hat{\varepsilon}_{z,t}$</th>
<th>$\hat{\varepsilon}_{z,t}^*$</th>
<th>$\hat{\varepsilon}_{d,t}$</th>
<th>$\hat{\varepsilon}_{d,t}^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\varepsilon}_{a,t}$</td>
<td>1</td>
<td>0.36687</td>
<td>0.56191</td>
<td>0.45826</td>
<td>0.039716</td>
<td>0.03239</td>
</tr>
<tr>
<td>$\hat{\varepsilon}_{a,t}^*$</td>
<td>0.36687</td>
<td>1</td>
<td>0.45826</td>
<td>0.56191</td>
<td>0.03239</td>
<td>0.039716</td>
</tr>
<tr>
<td>$\hat{\varepsilon}_{z,t}$</td>
<td>0.56191</td>
<td>0.45826</td>
<td>1</td>
<td>0.97112</td>
<td>0.61576</td>
<td>0.50218</td>
</tr>
<tr>
<td>$\hat{\varepsilon}_{z,t}^*$</td>
<td>0.45826</td>
<td>0.56191</td>
<td>0.97112</td>
<td>1</td>
<td>0.50218</td>
<td>0.61576</td>
</tr>
<tr>
<td>$\hat{\varepsilon}_{d,t}$</td>
<td>0.039716</td>
<td>0.03239</td>
<td>0.61576</td>
<td>0.50218</td>
<td>1</td>
<td>-0.060154</td>
</tr>
<tr>
<td>$\hat{\varepsilon}_{d,t}^*$</td>
<td>0.03239</td>
<td>0.039716</td>
<td>0.61576</td>
<td>0.50218</td>
<td>-0.060154</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 6—Numerical value of steady-state Covariance matrix of the filtering errors for Learning from Consumption Stream.

<table>
<thead>
<tr>
<th></th>
<th>$\tilde{\epsilon}_{h,t}$</th>
<th>$\tilde{\epsilon}_{f,t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tilde{\epsilon}_{h,t}$</td>
<td>4.5324e-007</td>
<td>4.5394e-007</td>
</tr>
<tr>
<td>$\tilde{\epsilon}_{f,t}$</td>
<td>4.5394e-007</td>
<td>6.1691e-007</td>
</tr>
</tbody>
</table>

Table 7—Numerical value of steady-state Covariance matrix of the filtering errors for Learning from Consumption and Dividend Stream.

<table>
<thead>
<tr>
<th></th>
<th>$\tilde{\epsilon}_{h,t}$</th>
<th>$\tilde{\epsilon}_{f,t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tilde{\epsilon}_{h,t}$</td>
<td>3.1137e-007</td>
<td>2.9626e-007</td>
</tr>
<tr>
<td>$\tilde{\epsilon}_{f,t}$</td>
<td>2.9626e-007</td>
<td>4.2381e-007</td>
</tr>
</tbody>
</table>

Table 8—Numerical Simulation Results: Annualized Mean under different information structures.

*FI* is full information benchmark; *LI*$_{cs}$ is learning from consumption stream using steady-state Kalman filter; *LI*$_{cd}$ is learning from consumption and dividend stream using steady-state Kalman filter. $z_t$ is the long-run persistent component; $\Delta \ln C_t$ is log consumption growth; $\Delta \ln D_t$ is log dividend growth; $m$ is pricing kernel; $\Delta \epsilon_t$ is log exchange rate growth; $r_{f,t}$ is log risk free rate; $r_{d,t}$ is log return on dividend stream; $r_{ex,t}$ is equity risk premium; $\ln \frac{P_t}{C_t}$ is the log price consumption ratio; $r_{c,t}$ is the log return on the asset which pays the consumption stream. Simulation is done with third order numerical approximation for 25 chains each with 700 periods in monthly frequency but reported values are annualized and multiplied by 100.

<table>
<thead>
<tr>
<th></th>
<th>$FI$</th>
<th>$LI_{cs}$</th>
<th>$LI_{cd}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z_t$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\Delta \ln C_t$</td>
<td>1.8</td>
<td>1.8</td>
<td>1.8</td>
</tr>
<tr>
<td>$\Delta \ln D_t$</td>
<td>0.84</td>
<td>0.84</td>
<td>0.84</td>
</tr>
<tr>
<td>$m_t$</td>
<td>-8.4586</td>
<td>-8.762</td>
<td>-8.7359</td>
</tr>
<tr>
<td>$\Delta \epsilon_t$</td>
<td>-2.0553</td>
<td>-2.1795</td>
<td>-2.1791</td>
</tr>
<tr>
<td>$r_{f,t}$</td>
<td>2.3328</td>
<td>2.1909</td>
<td>2.231</td>
</tr>
<tr>
<td>$r_{d,t}$</td>
<td>6.4177</td>
<td>7.2046</td>
<td>6.6513</td>
</tr>
<tr>
<td>$r_{ex,t}$</td>
<td>4.0849</td>
<td>5.0137</td>
<td>4.4204</td>
</tr>
<tr>
<td>$\ln \frac{P_t}{C_t}$</td>
<td>79.2296</td>
<td>77.8268</td>
<td>77.9568</td>
</tr>
<tr>
<td>$r_{c,t}$</td>
<td>3.0044</td>
<td>3.0205</td>
<td>3.019</td>
</tr>
</tbody>
</table>
Table 9—Numerical Simulation Results: Annualized Volatility under different information structures. \(FI\) is full information benchmark; \(LI_{cs}^c\) is learning from consumption stream using steady-state Kalman filter; \(LI_{cd}^c\) is learning from consumption and dividend stream using steady-state Kalman filter. \(z_t\) is the long-run persistent component; \(\Delta \ln C_t\) is log consumption growth; \(\Delta \ln D_t\) is log dividend growth; \(m\) is pricing kernel; \(\Delta e_t\) is log exchange rate growth; \(r_{f,t}\) is log risk free rate; \(r_{d,t}\) is log return on dividend stream; \(r_{ex,t}\) is equity risk premium; \(\ln P_t/C_t\) is the log price consumption ratio; \(r_{c,t}\) is the log return on the asset which pays the consumption stream. Simulation is done with third order numerical approximation for 25 chains each with 700 periods in monthly frequency but reported values are annualized and multiplied by 100.

<table>
<thead>
<tr>
<th></th>
<th>(FI)</th>
<th>(LI_{cs}^c)</th>
<th>(LI_{cd}^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(z_t)</td>
<td>0.3218</td>
<td>0.2695</td>
<td>0.2999</td>
</tr>
<tr>
<td>(\Delta \ln C_t)</td>
<td>1.0444</td>
<td>1.0696</td>
<td>1.0589</td>
</tr>
<tr>
<td>(\Delta \ln D_t)</td>
<td>5.1658</td>
<td>5.8829</td>
<td>5.2499</td>
</tr>
<tr>
<td>(m_t)</td>
<td>34.85</td>
<td>36.4356</td>
<td>35.6694</td>
</tr>
<tr>
<td>(\Delta e_t)</td>
<td>19.2524</td>
<td>19.9782</td>
<td>20.2438</td>
</tr>
<tr>
<td>(r_{f,t})</td>
<td>0.1576</td>
<td>0.132</td>
<td>0.1469</td>
</tr>
<tr>
<td>(r_{d,t})</td>
<td>15.4034</td>
<td>16.6043</td>
<td>16.1085</td>
</tr>
<tr>
<td>(r_{ex,t})</td>
<td>15.4068</td>
<td>16.6067</td>
<td>16.1112</td>
</tr>
<tr>
<td>(\ln P_t/C_t)</td>
<td>7.5098</td>
<td>6.284</td>
<td>6.9969</td>
</tr>
<tr>
<td>(r_{c,t})</td>
<td>2.012</td>
<td>2.3885</td>
<td>2.2648</td>
</tr>
</tbody>
</table>

Table 10—Numerical Simulation Results: Autocorrelation-ACF(1) under different information structures. \(FI\) is full information benchmark; \(LI_{cs}^c\) is learning from consumption stream using steady-state Kalman filter; \(LI_{cd}^c\) is learning from consumption and dividend stream using steady-state Kalman filter. \(z_t\) is the long-run persistent component; \(\Delta \ln C_t\) is log consumption growth; \(\Delta \ln D_t\) is log dividend growth; \(m\) is pricing kernel; \(\Delta e_t\) is log exchange rate growth; \(r_{f,t}\) is log risk free rate; \(r_{d,t}\) is log return on dividend stream; \(r_{ex,t}\) is equity risk premium; \(\ln P_t/C_t\) is the log price consumption ratio; \(r_{c,t}\) is the log return on the asset which pays the consumption stream. Simulation is done with third order numerical approximation for 25 chains each with 700 periods in monthly frequency but reported values are annualized and multiplied by 100.

<table>
<thead>
<tr>
<th></th>
<th>(FI)</th>
<th>(LI_{cs}^c)</th>
<th>(LI_{cd}^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(z_t)</td>
<td>96.9764</td>
<td>97.0699</td>
<td>97.5444</td>
</tr>
<tr>
<td>(\Delta \ln C_t)</td>
<td>10.3912</td>
<td>8.1219</td>
<td>11.2292</td>
</tr>
<tr>
<td>(\Delta \ln D_t)</td>
<td>13.1464</td>
<td>8.1219</td>
<td>10.7614</td>
</tr>
<tr>
<td>(m_t)</td>
<td>1.8545</td>
<td>-1.4132</td>
<td>0.7303</td>
</tr>
<tr>
<td>(\Delta e_t)</td>
<td>-3.7041</td>
<td>0.2235</td>
<td>0.9167</td>
</tr>
<tr>
<td>(r_{f,t})</td>
<td>96.9764</td>
<td>97.0709</td>
<td>97.5447</td>
</tr>
<tr>
<td>(r_{d,t})</td>
<td>4.4295</td>
<td>-1.3378</td>
<td>0.7313</td>
</tr>
<tr>
<td>(r_{ex,t})</td>
<td>4.2478</td>
<td>-1.4892</td>
<td>0.5728</td>
</tr>
<tr>
<td>(\ln P_t/C_t)</td>
<td>96.9759</td>
<td>97.07</td>
<td>97.5442</td>
</tr>
<tr>
<td>(r_{c,t})</td>
<td>2.4978</td>
<td>-0.2555</td>
<td>2.1265</td>
</tr>
</tbody>
</table>
Table 11—Numerical Simulation Results: Correlation of Home/Foreign Counterparts, under different information structures. \( FI \) is full information benchmark; \( LI_{cs} \) is learning from consumption stream using steady-state Kalman filter; \( LI_{cd} \) is learning from consumption and dividend stream using steady-state Kalman filter. \( z_t \) is the long-run persistent component; \( \Delta \ln C_t \) is log consumption growth; \( \Delta \ln D_t \) is log dividend growth; \( m_t \) is pricing kernel; \( \Delta \epsilon_t \) is log exchange rate growth; \( r_{f,t} \) is log risk free rate; \( r_{d,t} \) is log return on dividend stream; \( r_{ex,t} \) is equity risk premium; \( \ln \frac{P_t}{C_t} \) is the log price consumption ratio ; \( r_{c,t} \) is the log return on the asset which pays the consumption stream. Simulation is done with third order numerical approximation for 25 chains each with 700 periods in monthly frequency but reported values are annualized and multiplied by 100.

<table>
<thead>
<tr>
<th></th>
<th>( FI )</th>
<th>( LI_{cs} )</th>
<th>( LI_{cd} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( z_t )</td>
<td>92.1642</td>
<td>97.6228</td>
<td>96.0112</td>
</tr>
<tr>
<td>( \Delta \ln C_t )</td>
<td>39.5981</td>
<td>40.9778</td>
<td>39.2543</td>
</tr>
<tr>
<td>( \Delta \ln D_t )</td>
<td>1.9399</td>
<td>40.9778</td>
<td>0.7725</td>
</tr>
<tr>
<td>( m_t )</td>
<td>88.3361</td>
<td>88.0758</td>
<td>87.1885</td>
</tr>
<tr>
<td>( \Delta \epsilon_t )</td>
<td>-100</td>
<td>-100</td>
<td>-100</td>
</tr>
<tr>
<td>( r_{f,t} )</td>
<td>92.1937</td>
<td>97.6037</td>
<td>96.0116</td>
</tr>
<tr>
<td>( r_{d,t} )</td>
<td>82.1113</td>
<td>84.8666</td>
<td>80.5397</td>
</tr>
<tr>
<td>( r_{ex,t} )</td>
<td>82.1154</td>
<td>84.8671</td>
<td>80.5463</td>
</tr>
<tr>
<td>( \ln \frac{P_t}{C_t} )</td>
<td>79.262</td>
<td>93.2036</td>
<td>87.8564</td>
</tr>
<tr>
<td>( r_{c,t} )</td>
<td>71.2044</td>
<td>75.3846</td>
<td>72.3027</td>
</tr>
</tbody>
</table>

Table 12—Numerical Simulation Results: Theoretical \( \beta_{UIP} \) regression coefficients of the equation \( \Delta \epsilon_{t+1} = \alpha + \beta_{UIP} \cdot (r_{f,t} - r_{*f,t}) + \epsilon_t \). \( FI \) is full information benchmark; \( LI_{hh}^{cs} \) is learning from consumption stream using a not yet converged Kalman filter from a prior in which variance of estimation error of home’s latent variable was believed to be 150% the steady-state Kalman filter value; \( LI_{hh}^{cd} \) is learning from consumption and dividend stream using a not yet converged Kalman filter from a prior in which variance of estimation error of home’s latent variable was believed to be 150% the steady-state Kalman filter value. Simulation is done with third order numerical approximation for 25 chains each with 100 periods in monthly frequency.

<table>
<thead>
<tr>
<th>( \beta_{UIP_{small sample}} )</th>
<th>( FI )</th>
<th>( LI_{hh}^{cs} )</th>
<th>( LI_{hh}^{cd} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_{UIP_{small sample}} )</td>
<td>1.0000</td>
<td>-2.8378</td>
<td>-0.7408</td>
</tr>
</tbody>
</table>
References


