Real Exchange Rates with Heterogeneous Beliefs

Shixiang Xia∗

November 25, 2019
Click here for the most recent version

Abstract

This paper proposes a parsimonious two-country, two-good, and complete-market model featuring heterogeneous beliefs to address the Backus-Smith, volatility, and forward premium puzzles in international finance. The presence of the time-varying difference in beliefs has direct and indirect effects on equilibrium exchange rates. The direct effect appears as a wedge in the pricing kernels while the endogenous indirect effect operates through the dynamic reallocation of equilibrium consumption shares in response to the belief difference. With a general setup of the belief difference, the direct and indirect effects jointly help to qualitatively address these puzzles. In calibration, the model reconciles highly correlated pricing kernels with moderately correlated consumption growth rates. Moreover, the model generates a sizable currency risk premium and a disconnect between exchange rate changes and consumption growth differentials.

Keywords: Heterogeneous beliefs, real exchange rate, Backus-Smith puzzle, volatility puzzle, forward premium puzzle

∗Kellogg School of Management, Northwestern University. E-mail: s-xia@kellogg.northwestern.edu. I am very grateful to Konstantin Milbradt, Ian Dew-Becker, Zhengyang Jiang and Sergio Rebelo for advice and support. I also thank Nicolas Crouzet, Anthony DeFusco, Martin Eichenbaum, Apoorv Gupta, Qiushi Huang, Howard Kung, Brittany Lewis, Robert McDonald, Timothy McQuade, Jacopo Ponticelli, Marius Ring, Yujie Ruan, Tobias Sichert, Costis Skiadas, Kandarp Srinivasan, Viktor Todorov, Ngoc-Khanh Tran, Raj Vasu, Adrien Verdelhan, and seminar participants at Northwestern Kellogg, Trans-Atlantic Doctoral Conference, and SoFiE Financial Econometrics Summer School for helpful discussions and comments.
1 Introduction

This paper offers a novel approach featuring cross-country heterogeneous beliefs to explain several phenomena in international finance, namely the Backus-Smith puzzle, the exchange rate volatility puzzle, and the forward premium puzzle. These puzzles are long-standing, and it is important to resolve these perplexing empirical features in order to gain a further understanding of the link between exchange rate dynamics and equilibrium consumption. This paper provides a unified framework to reconcile these puzzles simultaneously.

The Backus and Smith (1993) condition posits that under the standard assumptions of complete markets and time-separable preferences, the change in exchange rates must be perfectly correlated with the consumption growth differentials across the two countries. However, in the data, the correlation between exchange rate changes and cross-country consumption growth differentials has been persistently low and even negative. Brandt, Cochrane and Santa-Clara (2006) shows that a high degree of international risk sharing is needed to match the moderate level of exchange rate volatility in the data. Under the standard assumptions, this implies that the consumption growths between the two countries also have to be highly correlated. Empirically, however, the average correlation is around 0.4, too small to accommodate the level of exchange rate changes volatility. Fama (1984) documents the forward premium puzzle that in the data, the high interest rate currencies tend to appreciate while uncovered interest rate parity suggests the interest rate differentials should be fully offset by the expected changes in the exchange rate. Thus, borrowing low interest rate currencies and investing in the higher ones leads to profitable carry trades on average.

This paper abstracts away from either incomplete/segmented market setting or rich preferences but instead provides a complete-market model and focuses on how heterogeneous beliefs can overcome the perfect relation between exchange rates and consumption. This is the first paper that uses heterogeneous beliefs across countries in a simple, complete-market setting to resolve the three puzzles simultaneously. The model features two countries, each having a representative household and a country-specific endowment. The households are assumed to obtain different estimates of unobserved economic fundamentals from each other and thus form different beliefs regarding the economy. The heterogeneous beliefs have rich implications in equilibrium. Importantly, differences in beliefs have both direct and indirect effects on the equilibrium exchange rate.

---

1See, for example, Bakshi, Cerrato and Crosby (2017) for an incomplete market setting and Colacito and Croce (2011) for a long-run risk framework.
Under complete markets, the real exchange rate is simply the ratio of the two pricing kernels. The direct effect of heterogeneous beliefs on the equilibrium exchange rate appears as a time-varying wedge to the marginal utilities that adjusts the relative probability measures between households. As a result, the perfect relation between exchange rate changes and the cross-country consumption growth differentials is broken down. Lustig and Verdelhan (2019) (henceforth LV19) introduce a stochastic wedge in pricing kernels under incomplete spanning. They show that the incomplete spanning itself is inadequate to simultaneously resolve the three puzzles. The key in my paper to reconciling these puzzles is the endogenous indirect effect that operates through the dynamic reallocation of consumption shares in response to different beliefs. The optimal consumption allocations become endogenous to the belief difference across the two countries. For example, the households will allocate fewer resources to a particular state if they think this state has relatively lower probability compared to the other household. In contrast, a model with homogeneous beliefs would imply constant shares of endowments at all times.

The wedge in my paper seems close to the one in LV19. However, there are several important distinctions. First, the origins of the wedges are different. In LV19, the wedge is exogenously chosen and contains risks that are outside the space of traded assets. Thus, it does not affect asset prices. In contrast, the wedge or the direct effect in my model reflects relative beliefs between the two countries and contains risks that correlate with asset returns due to the indirect effect. Second, in my paper, the key to resolving the three exchange rate puzzles simultaneously is the interaction between the direct and indirect effects. LV19 is essentially analogous to a setting with only the direct effect and the indirect effect is absent since they impose an exogenous relationship between consumption and the wedge. They have shown that the direct effect alone cannot explain all three puzzles. Furthermore, the formulation of the no-arbitrage constraint in LV19 also differs from my paper. In LV19, as the wedge only consists of orthogonal risks to the traded assets, the pricing cannot be affected by the inclusion of the wedge. In my paper, since the wedge correlates with asset returns, the underlying probability measure must be appropriately adjusted when pricing assets with the wedge to ensure no arbitrage.

The interplay between the direct and indirect effects of the belief difference creates a negative correlation between the exchange rate changes and the consumption growth differentials and thus resolves the Backus-Smith puzzle. Denote the two countries as $A$ and $B$. Suppose country $A$ perceives a particular event with a higher probability of occurring than what $B$ thinks. On one hand, this implies that $A$ will allocate more resources to this
particular event relative to $B$, as in $A$’s mind this event is relatively more likely to happen. Therefore, country $A$’s consumption growth for such an event is higher than $B$’s. On the other hand, due to home bias, the consumption of the domestic good by country $A$ increases more than the foreign good. As a result, the price for country $A$’s output relatively increases and $A$’s currency appreciates. Thus, the exchange rate changes move oppositely from the consumption growth differentials.

A high correlation between pricing kernels is required to generate a modest level of exchange rate changes volatility. In addition to marginal utilities, because of the direct effect, the pricing kernel of each country contains a wedge that reflects different probability measures between the two countries. The correlation between pricing kernels thus involves the interactions between the wedges and marginal utilities. Suppose country $A$ perceives an event with a higher probability, i.e. an increase in their wedge. Due to the indirect effect, $B$’s consumption decreases and marginal utilities increase. As a result, the two pricing kernels are highly correlated because of the interactions between the wedge from the direct effect and the consumption reallocation that endogenously responds to the belief difference. In this way, we can achieve highly correlated pricing kernels while maintaining a moderate correlation between consumption growths.

The presence of heterogeneous beliefs introduces an additional premium to currency carry trade on top of endowment risks. The risk premium includes compensation for the endowment risks because countries have home bias and the endowment shocks are not perfectly correlated. The investors are also compensated for the risks that arise from disagreement. It is not only their own belief that matters for pricing assets, but investors also need to take into account different probability measures in the two countries, reflecting higher orders of beliefs. The additional risk generates a time-varying currency risk premium, thus explaining the carry trade.

To quantitatively examine the performance of the model, I provide a simple example of learning processes based on Xiong and Yan (2009). I calibrate the model and simulate 5000 paths of the economy for 50 years. The model generates a mild and negative correlation between exchange rate changes and cross-country consumption growth differentials. Furthermore, the model reconciles a relatively smooth level of exchange rate changes volatility with moderately correlated consumption growths. The model also yields a plausible size of risk premium. Other important macroeconomic and asset pricing moments are maintained in line with data. Moreover, the level of disagreement in the model is plausible as the households have small differences in estimates and their subjective beliefs are barely distinguishable.
Many papers have been devoted to resolving these three puzzles. Bakshi, Cerrato and Crosby (2017), Favilukis, Garlappi and Neamati (2015), Maurer and Tran (2016), Burnside and Graveline (2019), Sandulescu, Trojani and Vedolin (2019), and Lustig and Verdelhan (2019) employ market incompleteness or segmentation to address these issues. Papers featuring habits include Verdelhan (2010), Heyerdahl-Larsen (2014), and Stathopoulos (2016). Colacito and Croce (2011), Bansal and Shaliastovich (2012), and Yu (2013) adopt long-run risk framework, and Burnside, Eichenbaum, Kleshchelski and Rebelo (2010) and Farhi and Gabaix (2016) study how rare disasters can help to explain these empirical phenomena.

There are also a number of papers that use heterogeneous beliefs to explain other asset pricing anomalies. These papers include Scheinkman and Xiong (2003), Dumas, Kurshev and Uppal (2009), and Xiong and Yan (2009), and Dumas, Lewis and Osambela (2017). This paper employs these tools to quantitatively explain the phenomena in exchange rate dynamics.

The remaining of the paper is organized as follows. Section 2 introduces the model setup, defines the problem, and presents heterogeneous beliefs. Section 3 solves for the equilibrium consumption allocations, stochastic discount factors, real exchange rates, and interest rates. Section 4 relates heterogeneous beliefs to the exchange rate puzzles and shows how these puzzles can be qualitatively resolved. Section 5 quantitatively assesses the model using a specific example of learning processes in simulations and Section 6 concludes.

2 The Model

This section presents a parsimonious continuous-time two-country model based on Lucas (1978). A representative household within each country trades fruits from their endowed tree in a frictionless market. For some unspecified reasons, the households disagree on the values of unobserved fundamentals and form different beliefs regarding the economy. In what follows, I use the words “household” and “country” interchangeably, and the sub-/superscript $i$ refers to a generic country $i$.

2.1 Environment and preferences

Two countries $A$ and $B$ are each populated by a representative household and each endowed with a tradable good. The endowment $X_{it}$ follows the process
\[
\frac{dX_{it}}{X_{it}} = \mu_{it} dt + \sigma_{i} dZ_{it},
\]

where \( \mu_{it} \) is the time-varying expected growth rate, and \( \sigma_{i} \) denotes the volatility of country \( i \)'s output. The standard Brownian motions \( dZ_{At} \) and \( dZ_{Bt} \) are correlated with coefficient \( \rho \) to reflect some degree of integrity of outputs. Financial markets are complete, and households have access to a complete set of Arrow-Debreu securities. The trading of endowment goods and financial assets is frictionless.

Each household has log utility over a consumption basket consisting of traded goods from each country.\(^2\) The consumption basket for country \( i \) is defined to be

\[
C_{it} = \left( C_{it}^{i} \right)^{\alpha} \left( C_{jt}^{i} \right)^{1-\alpha},
\]

where \( C_{it}^{i} \) and \( C_{jt}^{i} \) represents the household \( i \)'s consumption for the traded goods produced by country \( i \) and \( j \) respectively. The weight \( \alpha \) on the domestic good is assumed to be \( \frac{1}{2} < \alpha < 1 \) to address home bias.

The optimization problem for country \( i \) is

\[
\max \{ C_{At}, C_{Bt} \} \mathbb{E}_0 \left[ \int_0^\infty e^{-\beta t} \log C_{it} dt \right]
\]

subject to the standard life-time budget constraint, where \( \beta > 0 \) is the time discount. Notice that the expectation is taken over household \( i \)'s own probability measure at time zero.

Finally, in this endowment economy, market clearing immediately implies that consumption of each good equals its output: \( C_{At}^{A} + C_{Bt}^{B} = X_{At} \) and \( C_{Bt}^{A} + C_{Bt}^{B} = X_{Bt} \).

### 2.2 Heterogeneous beliefs

Households in this model have limited knowledge about the underlying economy. While the two households observe the outputs levels \( X_{At} \) and \( X_{Bt} \), the expected growth rates \( \mu_{At} \) and \( \mu_{Bt} \) as well as the output shocks \( Z_{At} \) and \( Z_{Bt} \), there are some time-varying economic fundamentals \( \theta_t \) that are not directly observed by them. To allow for generality, the model does not specify what the unobserved quantities are — they could be anything related to the economy such as long-run growth rates and persistence of business cycles.\(^3\) The model

\(^2\)Log utility gives clean analytical results. See Appendix A for the extension to a general CRRA utility model with risk aversion parameter \( \gamma \neq 1 \) that is used in model calibration.

\(^3\)The quantitative exercise in Section 5 will give an example of unobserved fundamentals for calibration.
further assumes that the expected growth rate \( \mu_{it} \) can be expressed as

\[
d\mu_{it} = u(\mu_{it}, \theta_t) \, dt + v(\mu_{it}) \, dZ_{\mu_{it}},
\]

where the volatility \( v(\mu_{it}) \) could be a constant. The Brownian motions \( Z_{\mu t} = (Z_{\mu_{At}}, Z_{\mu_{Bt}}) \) are not directly observed and are referred to as unobserved fundamental shocks.

**Assumption 1.** Brownian motions \( Z_{\mu t} \) are mutually independent of each other and of output shocks \( Z_{At} \) and \( Z_{Bt} \).

As the dynamics of expected growth rates depend on \( \theta_t \), in order to allocate consumption and price assets, households will have to infer the values of these unobservables. For some unspecified reasons, households obtain different estimates at each point in time.\(^4\) In addition to estimating values of the unknowns, households also form their own beliefs regarding the underlying processes. Since the two countries’ estimates of the unobservables always differ, they each will have a unique probability measure. In particular, each country will have their own subjective Brownian motions with regards to unobserved fundamental shocks \( Z_{\mu t} \), which I denote \( W_{\mu t}^i \) as the perceived fundamental shocks under the subjective probability measure of country \( i \).\(^5\)

In order to draw comparisons and to ease the computation of expectations under two probability measures, the model uses a common probability measure to serve as a bridge connecting the two countries. In particular, I choose the common measure to reflect the average belief of the two countries. In the remainder of the paper, every asset can be exclusively priced under this chosen measure.

**Definition 1.** The probability measure \( \mathcal{E} \) represents the average belief of the two countries and is referred to as the benchmark (probability) measure.

Denote the perceived fundamental shocks under the benchmark measure as \( dW_{\mu t}^\mathcal{E} \). I define the process \( \xi_{it} \) to relate \( dW_{\mu t}^\mathcal{E} \) to country \( i \)'s Brownian motions \( dW_{\mu t}^i \) and impose restrictions on \( \xi_{it} \) as follows.

**Definition 2.** Let \( \xi_{it} \) be a stochastic process adapted to filtrations generated by Brownian motions under the benchmark measure \( \mathcal{E} \). It is given by

\[
\xi_{it} \, dt \equiv dW_{\mu t}^i - dW_{\mu t}^\mathcal{E}.
\]  

\(^4\)A microfoundation of estimating \( \theta_t \) is provided in Section 5.

\(^5\)Recall that countries observe \( Z_{At} \) and \( Z_{Bt} \). Thus, they both agree on the distributions of these shocks.
Assumption 2. The process for $\xi_{it}$ satisfies that $\exp \left( \xi_{it} - \frac{1}{2} \langle \xi_{it} \rangle \right)$ is a strictly positive martingale under the benchmark measure, where $\langle \xi_{it} \rangle$ is the quadratic variation of $\xi_{it}$.

The following theorem establishes the link between the benchmark and individual countries’ probability measures.

**Theorem 1** (Change of measure). The random variable $\eta_{it}^i$ defined to be

$$
\frac{d\eta_{it}^i}{\eta_{it}^i} = -\xi_{it}^T dW_{\mu t}^\mathcal{E}
$$

is the change-of-measure that translates the benchmark measure into country $i$’s. Moreover, Definition 1 implies that $\xi_{A t} + \xi_{B t} = 0$

The random variable $\eta_t \equiv (\eta_t^A, \eta_t^B)$ provides a connection between individual country’s probability measure to the benchmark measure. The discrete-time counterpart is simply the ratio of the probabilities between country $i$ and the benchmark. When the two countries have the same belief, $dW_{\mu t}^\mathcal{E} = dW_{\mu t}^A = dW_{\mu t}^B$. Thus, $\eta_t^A = \eta_t^B$ and $\xi_{A t} = \xi_{B t} = 0$. As soon as the countries have a disagreement, $\eta_t^A$ and $\eta_t^B$ are no longer equal to each other. Using this technique, country $i$’s problem can be expressed completely in terms of the benchmark measure. In particular, the expectation $E^i[Y]$ under $i$’s measure for some random variable $Y$ is equivalent to $E^\mathcal{E}[Y\eta_t]$. From now on, $\eta_t$ is referred to as the belief difference.

3 Equilibrium

3.1 Consumption allocations

Under the assumption that markets are complete, the marginal utilities for each good are equalized across the two countries. The equilibrium consumption baskets are given in the following proposition.\(^6\)

**Proposition 1.** Denote $\lambda_i$ as a parameter corresponding to the initial wealth in country $i$. The optimal consumption basket for country $A$ is

$$
C_{A t} = \left[ \frac{\lambda_A \alpha \eta_t^A}{\lambda_A \alpha \eta_t^A + \lambda_B (1 - \alpha) \eta_t^B} X_{A t} \right]^\alpha \left[ \frac{\lambda_A (1 - \alpha) \eta_t^A}{\lambda_A (1 - \alpha) \eta_t^A + \lambda_B \alpha \eta_t^B} X_{B t} \right]^{1 - \alpha},
$$

\(^6\)See Appendix B for the derivation of the competitive equilibrium.
and for country \( B \) is

\[
C_{Bt} = \left[ \frac{\lambda_B (1 - \alpha) \eta^B_t}{\lambda_A \alpha \eta^A_t + \lambda_B (1 - \alpha) \eta^B_t} X_{At} \right]^{1-\alpha} \left[ \frac{\lambda_B \alpha \eta^B_t}{\lambda_A (1 - \alpha) \eta^A_t + \lambda_B \alpha \eta^B_t} X_{Bt} \right]^\alpha. \tag{6}
\]

Moreover, the process for the equilibrium consumption basket of country \( i \) under the benchmark measure follows

\[
\frac{dC_{it}}{C_{it}} = \mu_{C,i} dt + \sigma_{C,i}^\top dW_t^\varepsilon = \mu_{C,i} dt + \sigma_{A,i}^C dZ_{At} + \sigma_{B,i}^C dZ_{Bt} + \sigma_{\mu,t}^C dW_{\mu,t}, \tag{7}
\]

where \( W_t^\varepsilon \equiv (Z_{At}, Z_{Bt}, W_{\mu,t})^\top \), and the drift and volatilities are given in Appendix F.

The change-of-measure variable \( \eta_t \) in equations (5) and (6) plays an important role in determining the equilibrium consumption allocations. Suppose the Pareto weights are the same (\( \lambda_A = \lambda_B \)) in a symmetric setup. In an economy with homogeneous beliefs, this will imply that each country shares a constant fraction of each endowment.\(^7\) This will not be the case when countries have heterogeneous beliefs. For simplicity, consider a discrete-time case with a state \( \omega \) at time \( t \). If country \( A \) associates a higher probability of \( \omega \) occurring at \( t \) relative to \( B \) (i.e. \( \eta^A_t > \eta^B_t \)), country \( A \) will allocate a relatively larger fraction of endowments to the state \( \omega \).

Equation (7) shows that the consumption processes are subject to both output shocks \((dZ)\) and fundamental shocks \((dW_{\mu,t}^\varepsilon)\) under the benchmark measure. The loadings on the output shocks are familiar and constant. The novel shocks \( dW_{\mu,t}^\varepsilon \) arise endogenously through heterogeneous beliefs across countries. The households now need to anticipate the other party’s belief when adjusting their own consumption allocations, reflecting higher orders of expectations. Indeed, when households have the same belief, the loadings on \( dW_{\mu,t}^\varepsilon \) will become identically zero, and we are back to the traditional consumption-based asset pricing framework. Moreover, unlike the output shocks, the volatilities \( \sigma_{\mu,t}^C \) are time-varying as an outcome of the stochastic differences in beliefs. They capture the instantaneous changes in equilibrium consumption shares due to fluctuations in the belief difference \( \eta_t \). The magnitudes of \( \sigma_{\mu,t}^C \) increase in (absolute values of) changes in \( \eta_t \), meaning that the volatilities also incorporate the severity of the belief difference — the higher degree of heterogeneity, the higher volatilities on perceived fundamental shocks.

\(^7\)The shares for country \( A \) are \( \frac{\lambda_A^{\alpha \lambda} \eta^A_t}{\lambda_A (1 - \alpha) \eta^A_t + \lambda_B (1 - \alpha) \eta^B_t} \) for domestic good and \( \frac{\lambda_A (1 - \alpha) \eta^A_t + \lambda_B \alpha \eta^B_t}{\lambda_A (1 - \alpha) \eta^A_t + \lambda_B \alpha \eta^B_t} \) for foreign good. Similar expressions can be obtained for country \( B \).
3.2 Stochastic discount factors

Since each country now takes into account of different probabilities that arise from differences in estimates, the stochastic discount factors (SDFs) are no longer simply the time-discounted marginal utilities. The following proposition gives country $i$’s SDF and its dynamics under the benchmark measure.$^8$

**Proposition 2.** Denote $M_{it}^\varepsilon$ as the stochastic discount factor for country $i$ under the benchmark measure. For log utility, it is given by

$$M_{it}^\varepsilon = e^{-\beta t} \frac{1}{C_{it}^\varepsilon} \eta_i^t. \quad (8)$$

The process for the above SDF under the benchmark measure follows

$$\frac{dM_{it}^\varepsilon}{M_{it}^\varepsilon} = \mu_{Mt} dt - \sigma_{\varepsilon Mt}^T dW_t^\varepsilon = \mu_{Mt} dt - \sigma_{\varepsilon A}^M dZ_{At} - \sigma_{\varepsilon B}^M dZ_{Bt} - \sigma_{\varepsilon \mu t}^M dW_{\varepsilon \mu t}. \quad (9)$$

The expressions for drift and the market price of risk are reported in Appendix F.

The extra term $\eta_i^t$ in the SDF (8) is necessary because the price must be consistent under all probability measures. It translates between the pricing measures of the households and the benchmark. The price not only incorporates how much the households’ utilities will increase with an additional unit of consumption, as captured by $\frac{1}{C_{it}^\varepsilon}$, but also takes into account how likely the events will occur in the minds of the two households, which is embedded in $\eta_i^t$.

Like consumption allocations, the pricing kernels are also subject to output shocks ($dZ$) and perceived fundamental shocks ($dW_{\varepsilon \mu t}$). The market price of risk for the output shocks is the same as the traditional literature and is constant by model assumptions. The market price of risk for $dW_{\varepsilon \mu t}$ reflects disagreements between households. The differences in estimates about the unobservables result in a time-varying risk premium on perceived fundamental shocks. It captures changes in relative beliefs and in marginal utilities due to fluctuations in the belief difference. Indeed, in the case of homogeneous beliefs, this time-varying risk premium vanishes.

---

$^8$Consider a discrete-time example. Denote, for example, $M_{At}^\varepsilon$ as country $A$’s SDF under their own measure. Let $R_{t+1}$ be an asset return at $t+1$ denominated in $A$’s currency. By definition and no arbitrage, we have $\mathbb{E}_t^A \left[ M_{At+1}^\varepsilon R_{t+1} \right] = \mathbb{E}_t^\varepsilon \left[ M_{At+1}^\varepsilon R_{t+1} \right]$. Since the equation above is true for any asset and markets are complete, we must have for any event $\omega$ that $\pi_{t+1}^A(\omega)M_{At+1}^A(\omega)R_{t+1}(\omega) = \pi_{t+1}^\varepsilon(\omega)M_{At+1}^\varepsilon(\omega)R_{t+1}(\omega)$, where $\pi_{t+1}^A(\omega)$ is the probability of $\omega$ under measure $t$. This implies $\frac{\pi_{t+1}^A(\omega)}{\pi_{t+1}^\varepsilon(\omega)}M_{At+1}^A(\omega) = M_{At+1}^\varepsilon(\omega)$. 


3.3 Interest rates

A country-specific risk-free bond purchased at time \( t \) pays a unit of the country’s currency at time \( t + dt \). Denote \( r_{it} \) as the risk-free rate and let \( \Omega_W \) be the correlation matrix of \( dW_t^C \). Then, we have in equilibrium

\[
  r_{it} = \beta + (\mu_{Ci}^t - \xi_{it}^t \sigma_{Ci}^t) - \sigma_{Ci}^t \Omega_W \sigma_{Ci}^t
\]

The term in parentheses is country \( i \)'s expected consumption growth \( \mu_{Ci}^t \) under \( i \)'s own measure. On one hand when household \( i \) perceives a high expected growth \( \mu_{Ci}^t \), they anticipate a high level of consumption in the future. Because utility is concave, households prefer smoothing consumption over time. Thus, a high expected growth induces them to consume more today by issuing bonds, driving the interest rate up. On the other hand, high consumption volatility triggers a precautionary savings motive. The risk-free bond becomes particularly valuable at times of high uncertainty, which reduces the interest rate.

3.4 Real exchange rate

Standard models with complete markets and time-separable utilities imply that the exchange rate equals the ratio of marginal utilities of the two countries because they must price all Arrow-Debreu securities denominated in either currency consistently. However, this relation no longer holds under heterogeneous beliefs. The following result gives the equilibrium exchange rate and its dynamics.

**Proposition 3.** Under the benchmark measure, the real exchange rate \( S_t \) is the ratio of two countries’ pricing kernels under homogeneous belief adjusted by the change-of-measure random variable. That is,

\[
S_t = \frac{M_{B,t}^C}{M_{A,t}^C} = \eta_h^B M_{B,t}^B = \eta_h^A M_{A,t}^A = \eta_h^A \eta_t^B C_{At} \eta_t^B C_{Bt},
\]

\[\text{The volatilities are preserved when switching to another measure.}\]

\[\text{Consider again a discrete-time case. Let } R_{t+1} \text{ be the return at time } t + 1 \text{ of a risky asset denominated in } B's \text{ currency. By no arbitrage, } E^A_t \left[ M_{A,t+1}^A R_{t+1} \frac{S_{t+1}}{S_t} \right] = E^B_t \left[ M_{B,t+1}^B R_{t+1} \right]. \text{ Complete markets imply for any event } \omega \text{ that } \pi_{t+1}^A(\omega) M_{A,t+1}(\omega) R_{t+1}(\omega) \frac{S_{t+1}(\omega)}{S_t} = \pi_{t+1}^B(\omega) M_{B,t+1}(\omega) R_{t+1}(\omega). \text{ Therefore, we must have } \frac{S_{t+1}(\omega)}{S_t} = \frac{\pi_{t+1}^B(\omega) M_{B,t+1}(\omega)}{\pi_{t+1}^A(\omega) M_{A,t+1}(\omega)}.\]

10
where $M_i^{\mu t} \equiv e^{-\beta t} \frac{1}{\mathcal{C}_{\mu t}}$ is country $i$’s pricing kernel under measure $i$.\(^{11}\) Moreover, the exchange rate follows the process

$$
\frac{dS_t}{S_t} = \mu_{St} dt - \sigma_{St}^T dW_t^\epsilon = \mu_{St} dt + \sigma_{\mu t}^{St} dZ_{At} + \sigma_{\mu t}^{St} dZ_{Bt} + \sigma_{\mu t}^{St} dW_{\mu t},
$$

where

$$
\mu_{St} = r_{At} - r_{Bt} + \sigma_{\mu t}^{St} \Omega_{W}(\sigma_{MA}^{St} - \sigma_{MB}^{St}).
$$

(12)

The explicit expressions for the drift and volatilities are reported in Appendix F.

The term $\frac{\eta_{At}}{\eta_{Bt}}$ in (11) has two key effects on the equilibrium exchange rate. The first one is the direct effect where it introduces a wedge between the exchange rate and ratio of marginal utilities. This is because households price assets under different probability measures and the equilibrium exchange rate must capture the valuation effect. The direct effect immediately implies that the wedge breaks down the perfect relation between the exchange rate and the ratio of consumption. The second effect is indirect, which operates through endogenous dynamic reallocation of consumption shares across the two countries. In optimizations, households maximize utilities under their own probability measures. The more likely a state of the economy is perceived to occur by country $i$, the more consumption will be allocated to $i$.

The equilibrium price for each good also depends on both parties’ beliefs. As a result, the real exchange rate $S_t$ moves in response to fluctuations in the belief difference $\eta_t$. For example, once $\eta_{At}$ fluctuates, the relative prices of outputs $X_{At}$ and $X_{Bt}$ also changes. Equation (39) in Appendix B shows that an increase in $\eta_{At}$ makes country $A$’s output more expensive. This is because as $\eta_{At}$ increases while keeping $\eta_{Bt}$ fixed, the relative share of goods consumed by country $A$ also increases. Since $A$ has home bias toward their domestic good, their consumption of domestic good surges even more than foreign good. Thus, given a fixed supply of $X_{At}$ and a stronger desire to consume domestic good, the price of $X_{At}$ increases, making country $A$’s per-unit consumption bundle more expensive, or equivalently country $B$’s consumption bundle becomes cheaper. Therefore, country $A$’s currency appreciates. Similarly, when $\eta_{At}$ decreases, country $A$ will cut down more consumption on domestic good, in which case country $A$’s consumption bundle becomes cheaper and their currency depreciates. Similar arguments apply to the case of $\eta_{Bt}$.

\(^{11}\)Real exchange rate in a general CRRA utility with risk aversion $\gamma$ is $S_t = \frac{\eta_{Bt}}{\eta_{At}} \left( \frac{\mathcal{C}_{At}}{\mathcal{C}_{Bt}} \right)^\gamma$.\(^{11}\)
4 Exchange Rate Puzzles

This section discusses three long-lasting puzzles in exchange rates and shows how heterogeneous beliefs can resolve all three puzzles qualitatively.

4.1 Backus-Smith puzzle

The Backus and Smith (1993) puzzle states that under the assumptions of complete markets and time-separable CRRA utilities, changes in exchange rate are perfectly correlated with consumption growth differentials. To see this, denote the log pricing kernels to be $m_{A,t}$ and $m_{B,t}$ in discrete time. The log exchange rate is $s_t = m_{B,t} - m_{A,t}$. Let $\gamma$ be the risk aversion parameter. We then have $\Delta s_t = \gamma(\Delta c_{A,t} - \Delta c_{B,t})$, which implies perfect correlation between $\Delta s_t$ and $\Delta c_{A,t} - \Delta c_{B,t}$. However, this prediction is rejected by data as the correlation is weak and even negative for most of the major currency pairs. The following shows how heterogeneous beliefs can resolve this puzzle in the special case of log utility in continuous time.

Define $\Delta C_t \equiv d\log C_{At} - d\log C_{Bt}$ as the difference in log consumption growth rates. We have

$$d\log S_t = d\log \eta^B_t - d\log \eta^A_t + \Delta C_t.$$  

We can see from the above that because of the direct effect, the belief difference breaks down the perfect correlation between exchange rate changes and differences in consumption growth rates. Importantly, consumption growths are endogenous to the belief difference due to the indirect effect. The following proposition shows heterogeneous beliefs can indeed help to resolve the puzzle.

**Proposition 4.** The covariance between $d\log S_t$ and $\Delta C_t$ is given by

$$\frac{1}{dt}\text{Cov}_t (d\log S_t, \Delta C_t) = \left(\sigma^C_{A,t} - \sigma^C_{B,t}\right)^2 + \left(\sigma^C_{B,t} - \sigma^C_{A,t}\right)^2 + 2\rho \left(\sigma^C_{A,t} - \sigma^C_{A,t}\right)\left(\sigma^C_{B,t} - \sigma^C_{B,t}\right) + \left(\xi_{At} - \xi_{Bt} + \sigma^C_{\mu,t} - \sigma^C_{\mu,t}\right)\top \left(\sigma^C_{\mu,t} - \sigma^C_{\mu,t}\right).$$  

(13)

In particular, the difference between two countries’ consumption volatilities loaded on the perceived fundamental shocks $dW^E_{\mu,t}$ can be written as

$$\sigma^C_{\mu,t} - \sigma^C_{\mu,t} = -f(\eta_t) (\xi_{At} - \xi_{Bt}).$$
where $0 < f(\eta_t) < 1$. Consequently, the presence of $\xi_{it}$ decreases the covariance between exchange rate changes and consumption growth differentials.

The above proposition implies that the second line in equation (13) is always less than zero. Moreover, the covariance between exchange rate changes and consumption growth differentials can potentially be negative, which implies a negative correlation between the two. Therefore, heterogeneous beliefs between the two countries help to lower the covariance and in turn reduce the correlation, hence resolving the Backus-Smith puzzle.

To see the intuition, consider a particular state of the economy. Suppose country $A$ perceives such a state with a higher probability of occurring than what $B$ thinks. This implies a lower wedge ($d \log \eta^B_t - d \log \eta^A_t < 0$). Since in country $A$’s mind, this event is relatively more likely to happen, this implies that $A$’s consumption growth for such state is higher than $B$’s ($\Delta C_t > 0$). At the same time, as country $A$’s consumption increases, the demand for domestic good surges due to home bias. This implies an increase in the price of $X_{At}$, leading to an appreciation in country $A$’s currency. Since the exchange rate is defined to be units of $A$’s currency per unit of $B$’s, we immediately see that exchange rate changes and consumption growth differentials move in opposite directions.

It is important to note that both the direct and indirect are necessary to resolve this puzzle. Suppose only the direct effect is present. The covariance (13) becomes the same as the homogeneous belief case. Although the correlation between exchange rate changes and consumption growth differentials is reduced, it can never become negative. If only the indirect effect is present, then again the covariance is always positive. The interaction between the two effects is the key to having a negative correlation.

### 4.2 Volatility puzzle

Since $d \log S_t = d \log M^F_{Bt} - d \log M^F_{At}$, taking unconditional variance on both sides yields

$$\text{Var}(d \log S_t) = \text{Var}(d \log M^F_{At}) + \text{Var}(d \log M^F_{Bt}) - 2\text{Cov}(d \log M^F_{At}, d \log M^F_{Bt}).$$

Brandt, Cochrane and Santa-Clara (2006) show that in this context high pricing kernel volatility implied from the Sharpe ratio in asset markets must be coupled with a high correlation between pricing kernels in order to obtain a reasonable level of exchange rate volatility ($\sim 10\%$ per year). Under homogeneous beliefs, this implies highly correlated cross-country consumption growths since $d \log M^F_{it} = -d \log C_{it}$. However, data shows that such correlation

13
is moderate, thus resulting in a much higher model-implied exchange rate changes volatility. The following shows with heterogeneous beliefs, a modest volatility of exchange rate can coexist with moderately correlated cross-country consumption growths.

Section 3.2 implies that the log pricing kernel can be expressed as

\[ d \log M^\epsilon_{it} = \text{drift}_M + d \log \eta^i_t - d \log C_{it}. \]

The correlation between pricing kernels crucially depends on the interaction between consumption and belief difference. The following proposition shows the covariance between growths in domestic consumption and negative foreign belief difference is always positive.

**Proposition 5.** The covariance between domestic consumption growth \( d \log C_{it} \) and changes in negative foreign belief difference \( -d \log \eta^j_t \), where \( i \neq j \), is

\[ \text{Cov}_t (d \log C_{it}, -d \log \eta^j_t) = \sigma_{\mu t} C_i^\top \xi_{jt} > 0. \tag{14} \]

Therefore, for a given level of pricing kernel volatility, a high correlation between SDFs is not necessarily because of highly correlated cross-country consumption growths. That is, to generate a modest volatility of \( d \log S_t \) while maintaining a low correlation between \( d \log C_{At} \) and \( d \log C_{Bt} \), it is sufficient to generate a high correlation through the interactions between consumption and the belief difference. Notice that the direct effect alone actually exacerbates the puzzle since \( \text{Cov}_t (d \log \eta^A_t, d \log \eta^B_t) < 0 \). Equation (14) shows that it is the interplay between the direct and indirect effects that generates highly correlated pricing kernels.

### 4.3 Forward premium puzzle

The uncovered interest rate parity (UIP) posits that the interest rate differentials between the two countries can be explained exactly by expected changes in the exchange rate. In other words, the risks in exchange rate must be fully reflected in the two countries’ riskless rates. However, in reality, there exist sizable currency carry trade premia. As can be seen from (12), the exchange rate premium is nonzero and time-varying. This section derives explicit expressions for the expected currency carry trade return and interest rate differentials and show that both are closely linked to heterogeneous beliefs.

Consider a trading strategy from the perspective of country A’s investors. At time \( t \), investors borrow country A’s bond (denominated in currency A), immediately convert to
currency B and then invest in country B’s bond. At time \( t + dt \), the investors convert the proceeds \( 1 + r_{Bt} dt \) from country B’s bond to currency A and pay back \( 1 + r_{At} dt \) to their lender. The realized log return \( rx_{t+dt} \), or the currency risk premium, from this trading strategy is\(^{12}\)

\[
rx_{t+dt} = d \log S_t + r_{Bt} dt - r_{At} dt.
\]

Taking conditional expectations under the benchmark measure yields the following result on the expected log carry trade return.

**Proposition 6.** The conditional expected log currency risk premium of borrowing currency A and lending B at time \( t \) is

\[
\mu_{t+dt}^{rx} \equiv \frac{1}{dt} \mathbb{E}^E_t [rx_{t+dt}] = \frac{1}{2} (\sigma_{MAt} + \sigma_{MBt})^\top \Omega_W (\sigma_{MAt} - \sigma_{MBt}),
\]

where \( \sigma_{M,i} \) is the market price of risk for country \( i \) and \( \Omega_W \) is the correlation matrix of \( dW_t^E \). Writing the expression more explicitly gives

\[
\mu_{t+dt}^{rx} = \frac{1}{2} \begin{pmatrix} \sigma_{MA}^2 + \sigma_{MB}^2 - \sigma_A^2 - \sigma_B^2 \end{pmatrix}
\begin{pmatrix}
\text{(constant) endowment shocks}
\end{pmatrix}
+ \frac{1}{2} \begin{pmatrix} \sigma_{\mu A} + \sigma_{\mu B} \end{pmatrix}
\begin{pmatrix}
\text{(stochastic) perceived fundamental shocks}
\end{pmatrix}
\begin{pmatrix}
\sigma_{\mu A} - \sigma_{\mu B}
\end{pmatrix},
\]

where \( n(\eta_t) > 0 \) is given in Appendix F.\(^{13}\)

A typical currency carry trade strategy is to borrow a low interest rate currency and invest in a high one at each point in time because high interest rate currencies tend to appreciate. The following proposition shows what determines the interest rate differentials in the model.

**Proposition 7.** The interest rate differentials \( r_{At} - r_{Bt} \) take the form

\[
r_{At} - r_{Bt} = constant + (2\alpha - 1)(\mu_{At} - \mu_{Bt}) + \frac{1}{2} n(\eta_t) (\eta_t^B - \eta_t^A),
\]

where \( n(\eta_t) \) is the same as the one in (17).

The expected return (17) consists of two components. The first is related to endowment risks as usual and is constant. It becomes zero when the two countries have the same output risks as usual and is constant. It becomes zero when the two countries have the same output risks as usual and is constant. It becomes zero when the two countries have the same output

\(^{12}\)The risk premium in levels can be found in Appendix F.

\(^{13}\)The constant part is 0 when \( \sigma_A = \sigma_B \).
volatility. The second component accounts for perceived fundamental shocks. The source of such risks stems from the disagreement of the two countries in estimating the unobserved fundamentals. As the budget constraint must hold, when an agent decides how much to consume they must anticipate the other party’s belief, which is fluctuating over time. Thus, the expected return is affected by the belief difference, and heterogeneous beliefs constitute a risk factor in addition to the traditional output risks.

The time-varying expected log return in this model depends on the relative belief differences $\eta_t^A$ and $\eta_t^B$ for the two countries. To see the intuition, consider some shock in $dW_{\mu t}^A$ that drives $\eta_t^A$ up. On one hand, by symmetry in (4), $\eta_t^B$ decreases. As $\eta_t^A$ increases, country $A$ has a stronger desire to consume and the demand for their own good surges even more because of home bias. Thus, country $A$’s currency appreciates and $B$’s depreciates. On the other hand, the difference in the changes in $\eta_t^A$ and $\eta_t^B$ depends on the current levels of belief differences. When $\eta_t^A > \eta_t^B$, the same shock makes $\eta_t^A$ react more than $\eta_t^B$. It can be shown that such shock drives up $A$’s pricing kernel, which means from the perspective of country $A$, the payoff from buying currency $B$ is low during bad times. The opposite happens when $\eta_t^A < \eta_t^B$. Therefore, the payoff from the strategy of long currency $B$ and short $A$ negatively correlates with country $A$’s SDF when $\eta_t^A > \eta_t^B$ and vice versa. The investor must be compensated in the former while paying for hedging in the latter. As a result, the expected return (17) increases with relative belief difference $\eta_t^A - \eta_t^B$ between country $A$ and $B$.

Meanwhile, equation (18) shows an opposite implication for the interest rate differentials that $r_{At} - r_{Bt}$ is decreasing in $\eta_t^A - \eta_t^B$. This is because as $\eta_t^A > \eta_t^B$, the demand for good $A$ becomes more volatile, which makes the price of good $A$ more volatile as well. Thus, country $A$’s risk-free bond, which guarantees the delivery of one unit of $A$’s consumption basket, becomes valuable and pays a lower interest rate. Combining this result with the expected risk premium, we see that borrowing a low interest rate currency and investing in a high one generates positive carry trade profits on average.

Home bias also plays an important role. It introduces some non-tradability between the two countries so that the shocks do not spread uniformly across the countries. When there is no home bias, the households are indifferent to consuming either good, and the economy is equivalent to the one with only one traded good. As a result, the exchange rate is always one in this case. Indeed, when $\alpha = 0.5$, the expected return from currency carry trade is zero.

The positive coefficient $n(\eta_t)$ in (17) and (18) incorporates both the direct and indirect effects of the belief difference. With only the direct effect, the belief difference does not appear
in the interest differentials because the interest rate must be the same under any probability measure. Since consumption in this case is not endogenous to the belief difference, the pricing of the domestic risk-free bond by domestic household will not convey any information about differences in beliefs. With only the indirect effect, it can be shown that both the risk premium and interest rate differentials move in the same direction with $\eta_t^B - \eta_t^A$, invalidating the fact that high interest rate currencies appreciate on average. Hence, the interaction between the direct and indirect effects is necessary to resolve the forward premium puzzle.

4.4 Qualitative discipline of heterogeneous beliefs

We have seen in the previous sections that the model qualitatively explains all three puzzles. One may cast doubt that with arbitrary beliefs the model is able to explain everything. However, it is important to bear in mind that even with a general process for the belief difference $\eta_t$ as in Section 2.2, the model cannot match all arbitrary facts. For example, imagine that the real-world data actually confirms the Backus-Smith condition. That is, assume counterfactually that there is a perfect correlation between exchange rate changes and consumption growth differentials. If there were truly unlimited degrees of freedom in the model, one could pick a particular process of $\eta_t$ that would match the counterfactual. However, Proposition 4 clearly proves this assertion invalid. Equation (13) shows that the correlation is reduced and can even be negative with a general specification of the belief difference. The reason for being unable to match all arbitrary facts is again due to the interaction between the direct and indirect effects of heterogeneous beliefs: the equilibrium consumption shares and $\eta_t$ comove in a particular way such that the model is capable of explaining only a few certain facts, which include the three puzzles.

4.5 Comparison to Lustig and Verdelhan (2019)

Notice from (8) that the marginal utilities are now coupled with a time-varying wedge $\eta_t$, which is the key to resolving the anomalies in real exchange rates and thus merits further discussion. LV19 show that even with an arbitrary wedge, one cannot address the three puzzles simultaneously under incomplete spanning. In their setup, a wedge $\tilde{\eta}_t$ containing payoff-irrelevant risks is introduced in pricing kernels. They find that the wedge parameters (e.g. mean and variance) cannot accommodate all three puzzles at the same time. They conclude that the incomplete spanning alone is insufficient to resolve the three puzzles. My paper circumvents these issues by introducing a new mechanism — heterogeneous beliefs.
Both equilibrium consumption allocations and the exchange rate are endogenous to the stochastic wedge generated by disagreements between households, which is the reason that the wedge in my paper can reconcile all puzzles at the same time.

The no-arbitrage constraint in LV19 for pricing risk-free bonds is

\[
\text{LV19: } \mathbb{E}_{t}^{LV} [M_{i,t+1} \tilde{\eta}_{t+1}] = \mathbb{E}_{t}^{LV} [M_{i,t+1}] \mathbb{E}_{t}^{LV} [\tilde{\eta}_{t+1}] + \text{Cov}^{LV} (M_{i,t+1} \tilde{\eta}_{t+1}) = \mathbb{E}_{t}^{LV} [M_{i,t+1}] . \tag{19}
\]

Equation (19) holds since \( \tilde{\eta}_{t} \) only contains risks outside of the traded asset space and the domestic risk-free bond must be consistently priced. In my paper, the wedge consists only of payoff-relevant risks and therefore the same condition need not hold. Once multiplying the original SDF by the wedge, we must also alter the underlying probability measure in order to maintain equality. We have for each state of the world that

\[
\text{This paper: } M_{i,t+1}^{\mathcal{E}} = M_{i,t+1}^{\mathcal{E}} \frac{\tilde{\eta}_{t+1}}{\eta_{t+1}} = \frac{\tilde{M}_{i,t+1}}{M_{i,t+1}} \frac{M_{i,t+1}^{\mathcal{E}}}{\tilde{M}_{i,t+1}} \eta_{t+1}^{\mathcal{E}} . \tag{20}
\]

Equation (20) implies that there exists an equivalent martingale measure where \( M_{i}^{\mathcal{E}} \) and \( M_{i}^{\mathcal{E}} \) are the state-price deflators under measures \( i \) and \( \mathcal{E} \), respectively. Such existence of a unique equilibrium measure guarantees no arbitrage. What differentiates this no-arbitrage condition from LV19 is again the change of probability measure when SDF is adjusted by the wedge.\(^{15}\)

An important implication of this paper is that the belief difference \( \eta_{t} \) affects the equilibrium exchange rate both directly and indirectly. Equation (20) shows that in this paper the pricing kernel for country \( i \) under the benchmark measure can be decomposed into three elements:

1. The term \( \tilde{M}_{i,t+1} \) is country \( i \)'s pricing kernel in the case of homogeneous beliefs. By

\(^{14}\)I abuse the notations for simplified exposition, where e.g. \( M_{t+1} \) and \( \eta_{t+1}^{i} \) should really be \( M_{t+1}^{i} / M_{t} \) and \( \eta_{t+1}^{i} / \eta_{t}^{i} \).

\(^{15}\)Suppose one incorrectly draws an equivalence between the wedges in my paper and LV19. Then they would immediately and falsely conclude that there is arbitrage in this setting because the expectations with and without the wedge do not agree in my paper: the reason all three anomalies can be resolved is due to the violation of the no-arbitrage constraint. However, the truth is that there is no arbitrage opportunity in this setup. If one literally thinks the two wedges are the same, then there would be what I call “fictitious arbitrage”. That is, if we totally ignore the presence of multiple probability measures and focus on measure \( i \) alone, then the no-arbitrage condition will not hold as we have \( \mathbb{E}_{t}^{i} [M_{i,t+1}] \neq \mathbb{E}_{t}^{\mathcal{E}} [M_{i,t+1}^{\mathcal{E}} \eta_{t+1}] \). But this arbitrage is not realistic and is entirely spurious. One must acknowledge that there are two agents with different beliefs and also that the wedge contains the risks that span the asset space. As mentioned above, I discipline the pricing kernels in such a way that no matter what probability measure is used, they must price assets consistently across agents with appropriately adjusted probability measures.
the assumption of CRRA utilities, the consumption shares are constant for each endowment, and $\tilde{M}_{t+1}$ is simply an exogenously given process.

(2). The **direct effect** appears as a time-varying wedge multiplied to the ratio of marginal utilities, which can be seen in (11). This may appear to resemble LV19. However, the wedge in LV19 is a result of incomplete spanning and contains only the risks that are outside the traded asset space and will not affect asset prices. In my paper, the wedge has a significantly different meaning. It is a result of heterogeneous expectations, and it tightly links to the perceived fundamental risks $dW^\varepsilon_t$. The wedge here serves, loosely speaking, as the ratio of the two probabilities in the minds of the households and the benchmark. The direct effect is an important component to resolve all three puzzles, but the wedge itself is insufficient. For example, suppose the indirect effect from (20) is removed. Since the belief difference $\eta_{t+1}$ is orthogonal to $\tilde{M}_{t+1}$ by Assumption 1, we immediately arrive at the LV19 condition (19) as $\text{Cov}_t (M_{t+1}, \eta_{t+1}) = 0$.

(3). Another key feature in the equilibrium exchange rate determination is the **indirect effect** of heterogeneous beliefs. As shown in Proposition 1, the optimal consumption shares are endogenous to the belief difference $\eta_t$ instead of being constant under homogeneous belief. The term $\frac{\tilde{M}_{t+1}}{M_{t+1}}$ in (20) captures the effect of endogenous reallocation of equilibrium consumption shares relative to a homogeneous beliefs world. Households care not only about their own beliefs but also about the other country’s. This endogenous response to heterogeneous beliefs affects the equilibrium real exchange rate via the ratio of marginal utilities. This is a stark contrast to LV19, in which there is an exogenous relationship between the consumption allocations (or pricing kernels) and the wedge. In other words, the endogenous indirect effect is completely shut down in their framework. In my paper, both the direct and indirect effects are in action, and the interplay between these two effects is the key to resolving all three puzzles simultaneously.

Like the direct effect, the indirect effect alone is not sufficient to resolve all three puzzles. For example, if the direct effect is ignored, then we are back to the standard models, in which gives

$$d \log S_t = d \log C_{At} - d \log C_{Bt}.$$ 

The interaction between the direct and indirect effects can be illustrated in the covariance between exchange rate changes and cross-country consumption growth differen-
tials. Under LV19 setting, we have

\[ \text{Cov}^{LV}_t (d \log S_t, \Delta C_t) = \text{Cov \ under \ homogeneous \ belief}. \]

In my paper, however, the covariance becomes

\[ \text{Cov}_t (d \log S_t, \Delta C_t) = \text{Cov \ under \ homogeneous \ belief} - h(\eta_t) (\xi_{At} - \xi_{Bt})^\top (\xi_{At} - \xi_{Bt}), \]

where \( h(\eta_t) > 0 \) for all \( \eta_t \). We immediately see that without the endogenous indirect effect, the second term in the equation above becomes zero and the covariance is exactly the same as the LV19 setting. When the indirect effect is present, the covariance decreases and can even become negative, thus solving the puzzle.

5 Quantitative Analysis

This section offers a simple, specific example of learning processes and then use this example for the calibration exercise with risk aversion parameter \( \gamma > 1 \). The details of the model extension are provided in Appendix A.

5.1 An example of learning processes

5.1.1 Unobservables and signals

Suppose the time-varying instant growth rate \( \mu_{it} \) in (1) follows

\[ d\mu_{it} = a(\theta_{it} - \mu_{it}) \, dt + \sigma_\mu dZ_{\mu_{it}}, \tag{21} \]

where \( a \) is the mean-reversion parameter, \( \theta_{it} \) is the long-run growth rate, and \( \sigma_\mu \) is constant volatility. The Brownian motions \( dZ_{\mu_{At}} \) and \( dZ_{\mu_{Bt}} \) are independent of each other and of output shocks \( Z_{At} \) and \( Z_{Bt} \).

The time-varying long-run growth rate \( \theta_{it} \) is given by the following OU process

\[ d\theta_{it} = b(\bar{\theta}_i - \theta_{it}) \, dt + \sigma_\theta dZ_{\theta_{it}}, \tag{22} \]

where \( b \) is the mean-reversion parameter, \( \bar{\theta}_i \) is steady-state of \( \theta_{it} \), and \( \sigma_\theta \) is constant.
volatility.\textsuperscript{16} The Brownian motions $Z_{\theta A t}$ and $Z_{\theta B t}$ are independent of each other and of $Z_{A t}, Z_{B t}, Z_{\mu A t}$ and $Z_{\mu B t}$.

As mentioned in Section 2.2, households observe outputs, instant growth rates as well as output shocks. In this example, the households are assumed to not observe the long-run growth rates $\theta_{it}$, the shocks realizations $Z_{\theta_i t}$, or the volatilities $\sigma_{\theta_i}$ for both countries. Instead, they receive public signals $s_{At}$ and $s_{Bt}$ regarding the shocks $dZ_{\theta A t}$ and $dZ_{\theta B t}$, respectively. For simplicity, let the true public signal processes be just pure noise and follow

$$ds_{it} = \sigma_s dZ_{s, it},$$

where $Z_{s, At}$ and $Z_{s, Bt}$ are independent Brownian motions.

Assume each household $i$ attaches correlations $\phi^i_A$ and $\phi^i_B$ between signals and shocks to long-run growth rates. That is, in household $i$’s mind the signal process $ds_{jt}$, where $j \in \{A, B\}$, follows

$$ds_{jt} = \sigma_s \phi^j_{s, it} dZ_{\theta_j t} + \sigma_s \sqrt{1 - (\phi^j_s)^2} dZ_{s,jt}.$$ \hspace{1cm} (23)

In addition, each household $i$ also misperceives the volatilities of $\theta_{At}$ and $\theta_{Bt}$ as $\sigma^i_{\theta A}$ and $\sigma^i_{\theta B}$.\textsuperscript{17} To ease the exhibition and algebra, I assume the following values for perceived signal correlations and long-run growth rate volatilities

$$\phi^A_A = \phi^B_B = \phi \hspace{1cm} \phi^A_B = \phi^B_A = -\phi$$

$$\sigma^A_{\theta A} = \sigma^B_{\theta B} = \frac{1}{\sqrt{1 - \phi^2}} \sigma_{\theta A} \hspace{1cm} \sigma^A_{\theta B} = \sigma^B_{\theta A} = \frac{1}{\sqrt{1 - \phi^2}} \sigma_{\theta B}.$$ \hspace{1cm} (24)

Remark 1. The above signal specification follows Xiong and Yan (2009). It is assumed that households think fundamentals are positively and negatively correlated with domestic foreign signals, respectively. However, the households do not know the exact model/signal specification in the mind of the other party, and they refuse to learn the true correlations of the signals because of their irrationality to be discussed below. Importantly, it is not necessary

\textsuperscript{16}This setting for $\theta_{it}$ is different from the long-run risk literature, in which people assume the two countries have a highly correlated long-run/persistent risk. Here, the process for the long-run mean $\mu_{it}$ is independent of each other. One can think of $\mu_{it}$ as expected quarterly GDP growth and $\theta_{it}$ as long-run GDP growth projection.

\textsuperscript{17}In the usual scenarios where the random variable $\theta_{it}$ is observable, one can in principle learn about its volatility precisely. In the current context, $\theta_{it}$ cannot be directly observed and thus the inference about its volatility will not be perfect. Instead of letting households infer the volatilities of $\theta_{it}$, I simply assume they stick to their priors, which gives more tractability to the model.
that households believe positive correlations in the domestic signals. The signal specifications (24) are made for the ease of computation and explanation, and we can see from Section 4 that the qualitative results do not rely on specific signal processes.

The above features of signal interpretations imply that the irrationality of households arises in this model. For a fully rational agent, they should adjust the correlation in the signals over time. In this specific example, while the households are assumed to rationally optimize consumption allocations, they are irrational in learning and processing the signals. In particular, they refuse to learn the true correlations due to behavioral reasons and instead impose a prior value on the correlation. The households disagree on the informativeness of the signals at each point in time, and they continue to estimate the unobservables by using their own interpretations of the signals even in the long run. In addition, what they learn about is a time-varying quantity. As a result, their estimates will not converge and will keep fluctuating over time.

5.1.2 Filtering and learning

To serve as a reference, imagine there exists a hypothetical rational benchmark agent (E) who receives the same information as the households but understands the true signal structures. Following Scheinkman and Xiong (2003), I assume the learning problem has reached its long-run steady state so that the variances of the forecasts are time-invariant. Denote \( \hat{\theta}^E \) as the mean value of the benchmark agent’s posterior distribution about \( \theta_{it} \). This quantity is referred to as the estimate of the long-run growth rate for country \( i \)’s good. Applying filtering theory from Liptser and Shiryaev (2000), the benchmark agent’s estimate for \( \theta_{it} \) follows

\[
d\hat{\theta}^E_{it} = b\left(\bar{\theta}_i - \hat{\theta}^E_{it}\right)dt + a\frac{\gamma_{\theta_i}}{\sigma_{\mu_i}^2} \left(d\mu_{it} - a\left(\hat{\theta}^E_{it} - \mu_{it}\right)dt\right) = b\left(\bar{\theta}_i - \hat{\theta}^E_{it}\right)dt + a\frac{\gamma_{\theta_i}}{\sigma_{\mu_i}} dW^E_{\mu_{it}},
\]

(26)

where \( dW^E_{\mu_{it}} \) is a standard Brownian motion under the benchmark probability measure. The steady-state variance of forecast error is given by

\[
\gamma_{\theta_i} = \frac{-b\sigma_{\mu_i}^2 + \sigma_{\mu_i}^2 \sqrt{b^2 + a^2 \frac{\sigma_{\theta_i}^2}{\sigma_{\mu_i}^2}}}{a^2}.
\]
We can see clearly from (26) that the estimate about $\theta_{it}$ in the long run will become precise if $\theta_{it}$ is deterministic ($\sigma_{\theta_i} = 0$). What prevents the estimate to converge to the true values is the shock $dZ_{\theta_{it}}$ at each point in time.

Denote $\hat{\theta}_{it}$ and $\hat{\theta}_{jt}$ as the country $i$’s estimates about $\theta_{it}$ and $\theta_{jt}$, where $i \neq j$. We have

$$d\hat{\theta}_{it} = b\left(\bar{\theta}_i - \hat{\theta}_{it}\right) dt + a \frac{\gamma_{\theta_i}}{\sigma_{\mu_i}} dW_{\mu_{i}t}^i + \frac{\phi_{\sigma_{\theta_i}}}{\sqrt{1 - \phi^2}} dW_{s_{it}}^i,$$ (27)

$$d\hat{\theta}_{jt} = b\left(\bar{\theta}_j - \hat{\theta}_{jt}\right) dt + a \frac{\gamma_{\theta_j}}{\sigma_{\mu_j}} dW_{\mu_{j}t}^j - \frac{\phi_{\sigma_{\theta_j}}}{\sqrt{1 - \phi^2}} dW_{s_{jt}}^j,$$ (28)

where $dW_{\mu_{i}t}^i$ and $dW_{\mu_{j}t}^j$ are defined similarly as $dW_{\mu_{it}}^\epsilon$ in (26). All shocks in (27) and (28) are Brownian motions under country $i$’s probability measure and are mutually independent.

The quantity $dW_{it}^\mu \equiv (dW_{\mu_{At}}^i, dW_{\mu_{Bt}}^i)$ represents perceived fundamental shocks under country $i$’s measure. A positive shock means that the estimates are too low and the household needs to adjust their estimations upward, and vice versa. The shocks $dW_{s_{At}}^i$ and $dW_{s_{Bt}}^i$ reflect the effects of signals on country $i$’s estimation. Since country $i$ believes that the domestic signal has a positive correlation with the fundamental, a positive (negative) signal shock will induce country $i$ to upwardly (downwardly) update the estimates about domestic long-run growth rates. The effects are opposite for the foreign signal. Note that by definition

$$ds_{it} = \sigma_s dW_{s_{it}}^A = \sigma_s dW_{s_{it}}^B = \sigma_s dW_{s_{it}}^\epsilon, \quad i \in \{A, B\}.$$

The only difference in the Brownian motions of the signals is that countries think domestic (foreign) signals are positively (negatively) correlated with the fundamental process.

We also observe that for a particular estimate, the benchmark agent’s estimation error coincides with the household’s. This is because on one hand, households think the signals are informative, and thus in their mind, the estimation is more precise; on the other hand, households postulate higher volatility for the unobservables, which could noise up the estimation errors. The two effects cancel out each other by the choice of values in (24) and (25), and we obtain the following result.

**Lemma 1.** The average of the two countries’ estimates about a particular long-run growth rate $\theta_{it}$ is equal to the benchmark agent’s estimate. That is, at each point in time,

$$\bar{\theta}_{it}^i - \bar{\theta}_{it}^j = \bar{\theta}_{it}^\epsilon - \bar{\theta}_{it}^j,$$
where $i \neq j$. Moreover, the perceived fundamental shocks $dW^\varepsilon_{\mu t}$ under the benchmark measure are equal to the average of the two countries

$$dW^\varepsilon_{\mu t} = \frac{1}{2} (dW^A_{\mu t} + dW^B_{\mu t}).$$

The above results state that an individual country may obtain a different estimate from the benchmark agent, the average of the two is always equal to the benchmark estimate. In other words, the benchmark agent tracks the average belief in the two countries.

Define $g^i_{At} \equiv \hat{\theta}^e_{At} - \hat{\theta}^i_{At}$ and $g^j_{Bt} \equiv \hat{\theta}^e_{Bt} - \hat{\theta}^i_{Bt}$ to be the differences in estimates between the benchmark agent and country $i$ about the two unobservables. By equations (26), (27) and (28), the processes for estimation differences satisfy

$$dg^i_{it} = -\left( b + a^2 \frac{\gamma_{\theta_i}}{\sigma^2_{\mu_i}} \right) g^i_{it} dt - \frac{\phi \sigma_{\theta_i}}{\sqrt{1 - \phi^2}} dW^e_{s,t}$$

$$dg^j_{jt} = -\left( b + a^2 \frac{\gamma_{\theta_j}}{\sigma^2_{\mu_j}} \right) g^j_{jt} dt + \frac{\phi \sigma_{\theta_j}}{\sqrt{1 - \phi^2}} dW^e_{s,t}, \quad i \neq j.$$ (29)

We can see that the differences in estimates have a long-run mean of zero. What prevent them to converge to the mean are the shocks to the signals $W^e_{s,t}$ as a result of the countries having different interpretations of the signals from the benchmark agent. The two countries disagree on the informativeness of the signals exactly in the opposite fashion. Thus, their steady-state variances of forecast errors coincide. In addition, as mentioned above, the two countries overvalue the long-run growth rate volatilities in a particular way such that the effects of estimation errors on the perceived fundamental shocks are eliminated. The signal shocks are preserved because of heterogeneous interpretations of the signals. Indeed, with homogeneous interpretation of the signals, the estimation differences are identically zero.

The differences in estimates also represent the relative attitude toward the economy by the two countries. For example, a negative $g^A_{At}$ implies that country $A$ is more optimistic about their domestic good than the benchmark agent. Since country $A$ values the shocks to signal as an informative source and believes that the domestic signal is positively correlated with $dZ_{\theta_A t}$, a positive realization of the signal will make country $A$ even more optimistic, thus reducing the value of $g^A_{At}$. As a result, we see that $dg^A_{At} < 0$ in (29) for $dW^e_{sA t} > 0$. 

24
By definition, the Brownian motions under different measures have the following relation.

\[ dW^\xi_{\mu t} = dW^A_{\mu t} - \alpha \frac{g^A_{it}}{\sigma_{\mu}}, dt = dW^B_{\mu t} - \alpha \frac{g^B_{it}}{\sigma_{\mu}}, dt. \] (31)

Let \( \xi_{At} = (\xi^A_{At}, \xi^B_{At}) \equiv \left( \frac{g^A_{it}}{\sigma_{\mu A}}, \frac{g^A_{it}}{\sigma_{\mu B}} \right) \) be country A’s scaled differences in estimates. Lemma 1 immediately implies that these quantities for country B is simply \( -\xi_{At} \). Then the change-of-measure variables \( \eta_t \) in (4) become

\[ \frac{d\eta^A_t}{\eta^A_t} = -\xi^A_{At} dW^\xi_{\mu At} - \xi^A_{Bt} dW^\xi_{\mu Bt} \] (32)
\[ \frac{d\eta^B_t}{\eta^B_t} = \xi^A_{At} dW^\xi_{\mu At} + \xi^A_{Bt} dW^\xi_{\mu Bt}. \] (33)

5.2 Data

The model is mapped to the United States for country A and United Kingdom for country B. The data are from 1980Q1 to 2018Q4 at quarterly frequency and are seasonally adjusted if applicable. All nominal quantities are deflated using either CPI or GDP deflators. As the model assumes a representative consumer in each country, for the data to match the model, the relevant empirical quantities must be converted to a per capita basis. The population data is from World Bank’s World Development Indicators (WDI) at annual level. To obtain quarterly population data, I assume linear population growth within a year for both countries. The equilibrium quantities will not change significantly if another (reasonable) form of population growth is used.

The real GDP data for US and UK are from Global Financial Database (GFD). Per capita values are obtained by dividing the real GDP by population. The real GDP per capita is mapped to the output processes \( X_{At} \) and \( X_{Bt} \) in the model. The real consumption data is from OECD.

The nominal foreign exchange rate (FX) data between US and UK is from WM/Reuters with quote currency as USD. To calculate the real FX, I multiply the nominal FX with the UK CPI and then divide it by US CPI. Finally, the nominal risk-free rates are proxied by three-month treasury bills denominated in local currencies, and the data is from GFD. The real interest rates are obtained by adjusting nominal rates with CPI-implied inflation rates.
Table 1: Calibrated parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output volatility</td>
<td>$\sigma_i$</td>
<td>0.03</td>
</tr>
<tr>
<td>Correlation of output</td>
<td>$\rho$</td>
<td>0.2</td>
</tr>
<tr>
<td>Correlation between signal and fundamental</td>
<td>$\phi$</td>
<td>0.90</td>
</tr>
<tr>
<td>Home bias</td>
<td>$\alpha$</td>
<td>0.87</td>
</tr>
<tr>
<td>Risk aversion</td>
<td>$\gamma$</td>
<td>5</td>
</tr>
<tr>
<td>Time discount</td>
<td>$\beta$</td>
<td>0.01</td>
</tr>
<tr>
<td>Pareto weight</td>
<td>$\lambda_i$</td>
<td>1</td>
</tr>
</tbody>
</table>

This table reports calibrated parameters that are either directly matched to the data, such as endowment volatilities, or taken from conventional values used in literature. All parameters are annualized where applicable.

5.3 Parameters

The perfect symmetry is imposed on the two countries, which gives 12 parameters to be determined in this model. Table 1 reports the 7 calibrated parameters. The (annualized) endowment volatilities are set to be $\sigma_A = \sigma_B = 0.03$ and the correlation between the outputs to be $\rho = 0.2$ in order to match the real GDP growth volatilities and correlation in the data. The risk aversion parameter $\gamma$ is set to 5 in order to generate reasonable SDF volatilities, and the time preference $\beta$ is 0.01. The home bias parameter is chosen to be $\alpha = 0.87$ to match the average trade-to-GDP ratio. To show the effects of heterogeneous beliefs, I set the signal correlation $\phi$ to be 0.9. The Pareto weights $\lambda_A$ and $\lambda_B$ are chosen to be 1 to indicate that countries are set to be symmetric in the model.

The remaining 5 parameters are estimated using simulated methods of moments (SMM) with details provided in Appendix C. The annualized model parameters and their standard errors are reported in Table 2. Since the number of targeted moments is larger than the number of parameters to be estimated, in order to see whether the model’s moment conditions are valid, I perform Hansen’s $J$-test for over-identifying restrictions. The $J$-statistic is 1.955 with a $p$-value of 0.855, indicating that the model is not misspecified.
Table 2: Estimated parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Volatility of expected growth rate</td>
<td>$\sigma_{\mu_i}$</td>
<td>0.011</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Volatility of long-run growth rate volatility</td>
<td>$\sigma_{\theta_i}$</td>
<td>0.002</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Steady-state long-run growth rate</td>
<td>$\bar{\theta}_i$</td>
<td>0.012</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Mean reversion parameter of expected growth</td>
<td>$a$</td>
<td>0.223</td>
<td>(0.037)</td>
</tr>
<tr>
<td>Mean reversion parameter of long-run growth</td>
<td>$b$</td>
<td>0.121</td>
<td>(0.020)</td>
</tr>
</tbody>
</table>

This table reports the remaining parameters used in model simulations. These parameters are estimated using simulated methods of moments outlined in Appendix C. All parameters are annualized where applicable. The standard errors are displayed in parentheses.

5.4 Simulation results

I choose quarterly time steps ($dt = 0.25$) and simulate 5000 paths of the economy for 60 years. In each sample path, the first ten years of data are dropped to avoid dependency on initial conditions. The results of average values of the key moments from simulations are reported in Table 3. For comparison, I also report a benchmark case in which countries have homogeneous beliefs ($\phi = 0$).

The Backus-Smith puzzle suggests that in complete markets with time-separable utility, the change in exchange rate is perfectly correlated with the cross-country differences in consumption growths. In that case, the correlation coefficient between $\Delta \log S$ and $\Delta \log C_B - \Delta \log C_A$ is 1, as can be seen in the homogeneous belief case. However, in the current model with heterogeneous beliefs, the correlation is weak and negative with a value of only $-0.23$, much closer to the empirical value.

I now move on to the exchange rate volatility puzzle. Brandt, Cochrane and Santa-Clara (2006) show that the modest volatility of exchange rate changes implies a high correlation between countries’ SDFs. With homogeneous beliefs, we would expect highly correlated consumption growths, which is counterfactual as the average consumption growths correlation between the two countries is around 0.3 to 0.4. As we see in the homogeneous beliefs...
Table 3: Simulation results — key moments

<table>
<thead>
<tr>
<th></th>
<th>Homogeneous</th>
<th>Model</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>$corr(\Delta \log S, \Delta \log C_B - \Delta \log C_A)$</td>
<td>1.00</td>
<td>-0.23</td>
<td>-0.17</td>
</tr>
<tr>
<td>$\sigma(\Delta \log S)$</td>
<td>4.78%</td>
<td>8.01%</td>
<td>10.7%</td>
</tr>
<tr>
<td>$corr(\Delta \log C_A, \Delta \log C_B)$</td>
<td>0.96</td>
<td>0.36</td>
<td>0.44</td>
</tr>
<tr>
<td>SDF correlation</td>
<td>0.96</td>
<td>0.92</td>
<td>–</td>
</tr>
<tr>
<td>Log carry trade returns</td>
<td>0.00%</td>
<td>4.60%</td>
<td>3.63%</td>
</tr>
</tbody>
</table>

This table reports key macro and asset pricing moments related to the puzzles from the simulations as well as their empirical values. Homogeneous refers to the case with homogeneous beliefs ($\phi = 0$) and Model refers to the one featuring heterogeneous beliefs. The carry trade refers to borrowing the low interest currency and investing in the high one at each point in time. The empirical moments are calculated using UK and US data from 1980Q1 to 2018Q4 at quarterly frequency. The simulated moments are the averages of 5000 sample paths, each having 50 years of quarterly data. All moments are annualized and calculated as the average of the two countries where applicable.

In the homogeneous case, the correlation of consumption growth rates across the two countries is much higher than typical values in the data. When the model incorporates heterogeneous beliefs, the correlation reduces to 0.36. At the same time, a high correlation between the two countries’ SDFs is maintained. As mentioned previously, the high correlation between SDFs is not due to a high correlation between consumption growths, but due to the interaction between consumption and the belief difference.

The model also generates a sizable currency risk premium close to data. As discussed before, when home bias is present but the two countries have the same belief, the log risk premium vanishes when two countries are symmetric, which can be seen in the homogeneous belief case. With heterogeneous beliefs being introduced, the households require an additional premium due to the perceived fundamental shocks. The average currency risk premium, in this case, is 4.60%, well in line with data.

Furthermore, Table 4 reports simulation results for other important moments of the two countries. The consumption growth volatility is a bit higher than data because in the model consumption processes consist of both endowment and belief risks, which could drive up the volatility. A stock or equity is the claim on the country $i$’s domestic endowment stream $X_{it}$. The equity excess return volatility appears to be high compared to the data while its mean seems relatively low. This difficulty of simultaneously matching interest rate and equity moments in a CRRA framework with heterogeneous beliefs is also mentioned in Dumas,
Table 4: Simulation results — other moments

<table>
<thead>
<tr>
<th></th>
<th>Model</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endowment growth mean</td>
<td>1.27%</td>
<td>1.66%</td>
</tr>
<tr>
<td>Endowment growth volatility</td>
<td>3.10%</td>
<td>2.50%</td>
</tr>
<tr>
<td>Endowment growth correlation</td>
<td>0.20</td>
<td>0.31</td>
</tr>
<tr>
<td>Consumption growth mean</td>
<td>1.52%</td>
<td>2.04%</td>
</tr>
<tr>
<td>Consumption growth volatility</td>
<td>3.50%</td>
<td>1.46%</td>
</tr>
<tr>
<td>Equity excess return mean</td>
<td>5.95%</td>
<td>8.08%</td>
</tr>
<tr>
<td>Equity excess return volatility</td>
<td>25.2%</td>
<td>16.1%</td>
</tr>
<tr>
<td>Sharpe ratio</td>
<td>0.23</td>
<td>0.45</td>
</tr>
<tr>
<td>Real interest rate mean</td>
<td>3.84%</td>
<td>1.98%</td>
</tr>
<tr>
<td>Real interest rate volatility</td>
<td>2.30%</td>
<td>1.94%</td>
</tr>
</tbody>
</table>

This table reports other macro and asset pricing moments from the simulations as well as empirical moments. The empirical moments are calculated using UK and US data from 1980Q1 to 2018Q4 at quarterly frequency. The simulated moments are the averages of 5000 sample paths, each having 50 years of quarterly data. All moments are annualized and calculated as the average of the two countries where applicable.

Kurshev and Uppal (2009). Models with more complex preferences such as habit formation may circumvent this issue.

5.5 Quantitative discipline of heterogeneous beliefs

We see from Table 3 that the model explains all three puzzles. One may wonder the reason that the model is successful is because of unlimited degrees of freedom in disagreements and unrealistic processes of belief differences. Section 4.4 has discussed the qualitative discipline of the model. This section shows that the belief difference is also quantitatively disciplined.

The simulation results show that the model matches well with not only moments related to the puzzles, but also other important asset pricing and macroeconomic moments in Table 4. Under the umbrella of the belief difference processes that are able to solve all three puzzles, not all can match the other moments in the data. In other words, keeping these additional model-implied moments in line with data puts a discipline to the process for $\eta_t$.

In addition, even if a process $\eta_t$ can match all moments in Table 3 and Table 4, one may still suspect that the specification is unrealistic in the sense that difference in beliefs may be
Table 5: Belief difference moments

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d \log \eta_t^i$ mean</td>
<td>-0.007</td>
</tr>
<tr>
<td>$d \log \eta_t^B - d \log \eta_t^A$ volatility</td>
<td>0.117</td>
</tr>
<tr>
<td>$</td>
<td>\hat{\theta}<em>{it} - \hat{\theta}</em>{jt}</td>
</tr>
</tbody>
</table>

This table reports the moments related to the belief difference in model simulations. The moments are calculated as the average across the two countries. All moments are annualized.

huge and very volatile. To address this concern, I calculate the average mean of the changes in the belief difference, the volatility of $d \log \eta_t^B - d \log \eta_t^A$ as well as the average mean of $|\hat{\theta}_{it} - \hat{\theta}_{jt}|$ across two countries in the simulations. The results are reported in Table 5. From the table, we observe that the estimation difference and the belief difference in the model are quite modest with the mean growth in the log belief difference process being $-0.7\%$ and mean absolute differences in estimates being $0.6\%$. In LV19, to generate plausible exchange rate volatility, they require the volatility of the wedge to be around $50\%$. In contrast, the growth in belief difference between the two countries has a moderate volatility of $11.7\%$ in my paper.

Table 6: Probability of belief detection

<table>
<thead>
<tr>
<th>Model</th>
<th>75 Years</th>
<th>100 Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark agent $\mathcal{E}$</td>
<td>1.04%</td>
<td>2.30%</td>
</tr>
<tr>
<td>Country B</td>
<td>6.77%</td>
<td>9.13%</td>
</tr>
</tbody>
</table>

This table reports the probability that household $A$ detects belief difference and rejects their own model in favor of an alternative. The model is simulated 10,000 times at quarterly frequency for 75 and 100 years. Household $A$ rejects their model if the test statistics exceeds the critical value at 5% significance. The results from household $B$’s perspective are similar and omitted.

Furthermore, the signal specifications and the implied beliefs should be thought plausible in the minds of the two households. That is, the households should not be able to differentiate each other’s model by observing the data. To test this, I back out the implied shocks $dW_{\mu t}^A$ and $dW_{\mu t}^B$ and ask whether they can be distinguished from each other and from the rational
benchmark agent’s shocks $dW^c_{\mu t}$ by conducting a likelihood-ratio test. Table 6 reports the probabilities that household $A$ is able to detect different beliefs by observing 75 and 100 years of data. We see that after having observed a long time-series of data, household $A$ still cannot easily distinguish each other’s model — less than 10% of the time with 100 years of data. Therefore, it is entirely plausible that the households have different interpretations of signals and form different beliefs. As a result, the belief difference process is well disciplined both qualitatively in Section 4.4 and quantitatively here.

6 Conclusion

This paper builds a parsimonious two-country, two-good general equilibrium model that incorporates cross-country heterogeneous beliefs to explain the Backus-Smith, exchange rate volatility, and forward premium puzzles. Households hold different beliefs regarding some unobserved fundamentals. Differences in beliefs break down the perfect relation between exchange rate changes and the ratio of marginal utilities. A key feature of the model is the interaction between the direct and indirect effects of the belief difference. The direct effect appears as a wedge in the pricing kernels, and the indirect effect operates through endogenous consumption shares in response to the fluctuations in beliefs. With a general setup of heterogeneous beliefs, the model is able to qualitatively address all three puzzles. To show the model works quantitatively, the paper provides a specific example of learning processes. In calibration, the model can successfully account for the puzzling empirical regularities in exchange rates by generating a disconnect between exchange rate changes and cross-country consumption growth differentials, highly correlated pricing kernels, a moderate correlation in cross-country consumption growths, a modest level of exchange rate volatility as well as a sizable carry trade return.
References


32


Favilukis, J. Y., Garlappi, L. and Neamati, S. (2015), ‘The carry trade and uncovered interest parity when markets are incomplete’.


Øksendal, B. (1997), ‘An introduction to malliavin calculus with applications to economics’.


Appendices

A Model Extension with a General CRRA Utility

This section derives the equilibrium under a general CRRA utility, i.e. \( U(C) = \frac{C^{1-\gamma}}{1-\gamma} \). This will alleviate some of the issues in the log utility case, such as low SDF volatilities. To better work with data, the simulation employs this general CRRA utility. The caveat is that now closed-form solutions cannot be obtained. I will go as far as possible to achieve analytical results.

A.1 Planner’s problem

Since we are still in complete markets, the competitive equilibrium allocations can be derived in a social planner’s optimization. The planner solves

\[
\max_{\{C_A^t, C_B^t, C_{At}^A, C_{At}^B, C_{Bt}^A, C_{Bt}^B\}} \mathbb{E}_0 \left[ \int_0^\infty e^{-\beta t} \left( \lambda_A \eta^A_t U(C_A^t) + \lambda_B \eta^B_t U(C_B^t) \right) dt \right],
\]

where

\[
U(C) = \frac{C^{1-\gamma}}{1-\gamma},
\]

subject to the resource constraint on the tradable good

\[
C_A^t + C_A^B = X_A^t
\]

\[
C_B^t + C_B^B = X_B^t.
\]

A.2 Equilibrium allocations

Denoting \( \nu_t \) as the Lagrange multiplier on the resource constraint, we obtain the first order conditions for country A

\[
\lambda_A \eta^A_t (1 - \alpha) C_A^{1-\gamma} \frac{1}{C_A^{1-\gamma}} = \nu_A^t
\]

\[
\lambda_A \eta^A_t (1 - \alpha) C_A^{1-\gamma} \frac{1}{C_A^{1-\gamma}} = \nu_B^t,
\]

and for country B

\[
\lambda_B \eta^B_t (1 - \alpha) C_B^{1-\gamma} \frac{1}{C_B^{1-\gamma}} = \nu_A^t
\]

\[
\lambda_B \eta^B_t (1 - \alpha) C_B^{1-\gamma} \frac{1}{C_B^{1-\gamma}} = \nu_B^t.
\]
We immediately see that the two countries equalize the marginal utilities on the same good due to frictionless trade, which implies
\[ \frac{\lambda_A \alpha C_A^{1-\gamma} C_B^{1-\gamma}}{\lambda_B (1 - \alpha) C_B^{1-\gamma} C_A^{1-\gamma}} \frac{\eta^A_t}{\eta^B_t} = 1 \]
\[ \frac{\lambda_A (1 - \alpha) C_A^{1-\gamma} C_B^{1-\gamma}}{\lambda_B \alpha C_B^{1-\gamma} C_A^{1-\gamma}} \frac{\eta^A_t}{\eta^B_t} = 1. \]

Re-arranging the two equations above, we obtain the following
\[ k_t \equiv \frac{\alpha C_B^A}{(1 - \alpha) C_A^A} = \frac{(1 - \alpha) C_B^A}{\alpha C_A^A}. \]

Define \( \omega_t \equiv \frac{C_A^A}{C_B^A} \) as the (equilibrium) ratio of two countries’ consumption of good A. In other words, the equilibrium consumption allocations for good A are
\[ C_A^A = \frac{\omega_t}{1 + \omega_t} X_A^t \]
\[ C_B^A = \frac{1}{1 + \omega_t} X_A^t \]

The first order conditions then imply
\[ 1 = \frac{\lambda_A \alpha}{\lambda_B (1 - \alpha)} \left( \frac{1}{\alpha} \right)^{1-\gamma} \left( \frac{C_A^A}{C_B^A} \right)^{(1-\gamma)k_t} \eta^A_t \]
\[ = \frac{\lambda_A}{\lambda_B} \left( \frac{1 - \alpha}{\alpha} \right)^{\omega_t - \gamma} k_t \frac{(1-2\alpha)(1-\gamma) \eta^A_t}{\eta^B_t}. \tag{35} \]

The market clearing conditions for good B gives
\[ \left( \frac{\alpha}{1 - \alpha} C_B^A + \frac{1 - \alpha}{\alpha} C_A^A \right) k_t = X_B^t. \]

Substituting \( C_A^A \) and \( C_B^A \) into the above equation yields
\[ k_t = \frac{X_B^t}{X_A^t} \left[ \frac{\alpha(1 - \alpha)(1 + \omega_t)}{\alpha^2 + (1 - \alpha)^2 \omega_t} \right]. \tag{36} \]

I can numerically solve for \( \omega_t \) using (35) and (36).
Proposition 8. The consumption basket for country A is

\[
C_{At} = \left( \frac{\omega_t X_{At}}{1 + \omega_t} \right)^{\alpha} \left( \frac{1 - \alpha}{\alpha} C_{At}^{A} k_t \right)^{1-\alpha}
\]

\[
= \left( \frac{\omega_t}{1 + \omega_t} X_{At} \right)^{\alpha} \left( \frac{(1 - \alpha)^2 \omega_t}{\alpha^2 + (1 - \alpha)^2 \omega_t} X_{Bt} \right)^{1-\alpha},
\]

and for country B is

\[
C_{Bt} = \left( \frac{\alpha}{1 - \alpha} C_{Bt}^{A} k_t \right)^{\alpha} \left( \frac{1}{1 + \omega_t} X_{At} \right)^{1-\alpha}
\]

\[
= \left( \frac{\alpha^2}{\alpha^2 + (1 - \alpha)^2 \omega_t} X_{Bt} \right)^{\alpha} \left( \frac{1}{1 + \omega_t} X_{At} \right)^{1-\alpha}.
\]

A.3 Dynamics of good A consumption ratio \( \omega_t \)

In order to obtain dynamics for consumption and subsequently exchange rate, we need to first derive the dynamics of the good A consumption ratio \( \omega_t \).

Proposition 9. The process for the equilibrium consumption ratio of good A follows

\[
d\omega_t = \mu_{\omega_t} dt + \sigma_{\omega_t}^{A} dZ_{At} + \sigma_{\omega_t}^{B} dZ_{Bt} + \sigma_{\mu_t}^{\omega} dW_{\mu t}^{E},
\]

where drift and volatilities are given in Appendix F.

A.4 Consumption dynamics

Proposition 10. The process for the equilibrium consumption follows

\[
dC_{it} = \mu_{C_{it}} dt + \sigma_{C_{it}}^{A} dZ_{At} + \sigma_{C_{it}}^{B} dZ_{Bt} + \sigma_{\mu_t}^{C_{it}} dW_{\mu t}^{E}, \quad i = A, B,
\]

where drift and volatilities are given in Appendix F.

A.5 Stochastic discount factors

As discussed in the main text, the pricing kernels under the benchmark measure are

\[
M_{At}^{E} = e^{-\beta t} C_{At}^{-\gamma} \eta_{At}^{A},
\]

\[
M_{Bt}^{E} = e^{-\beta t} C_{Bt}^{-\gamma} \eta_{Bt}^{B}.
\]

Proposition 11. The process for the stochastic discount factor follows

\[
dM_{it}^{E} = \mu_{M_{it}} dt - \sigma_{M_{it}}^{A} dZ_{At} - \sigma_{M_{it}}^{B} dZ_{Bt} - \sigma_{\mu_t}^{M_{it}} dW_{\mu t}^{E}, \quad i = A, B,
\]
where drift and volatilities are given in Appendix F.

A.6 Interest rates

A country-specific risk-free bond purchased at time $t$ pays a unit of the country’s currency, a unit of domestic consumption basket in this context, at time $t + dt$. It is easy to show that the instantaneous return on the risk-free bond is the additive inverse of the conditional expected growth rate in the SDF. Denote $r_{it}$ as the risk-free rate. Then we have

$$r_{it} = \beta + \gamma \mu_{Ct} + \frac{1}{2} \xi_{it}^\top \xi_{it} - \frac{1}{2} \gamma \left( \sigma_{Ct}^2 + \sigma_{Bt}^2 + 2 \rho \sigma_{ACt} \sigma_{BCt} + \sigma_{\mu t}^\top \sigma_{\mu t} \right) - \frac{1}{2} \left( \sigma_{Mt}^2 + \sigma_{Bt}^2 + 2 \rho \sigma_{AMt} \sigma_{BMt} + \sigma_{\mu t}^\top \sigma_{\mu t} \right).$$

A.7 Exchange rate

As in the main text, the exchange rate is the ratio of pricing kernels under measure $E$, and is given by

$$S_t = \frac{M_B}{M_A},$$

$$= \frac{\eta_B}{\eta_A} \left( \frac{C_{At}}{C_{Bt}} \right)^\gamma.$$

**Proposition 12.** The process for the exchange rate follows

$$\frac{dS_t}{S_t} = \mu_{St} dt + \sigma_{St}^\top dZ_{At} + \sigma_{Bt}^\top dZ_{Bt} + \sigma_{\mu t}^\top dW_{\mu t},$$

where drift and volatilities are given in Appendix F.

B Equivalence between Competitive Equilibrium (CE) and Planner’s Problem

This section solves for the competitive equilibrium for both countries and show that they are equivalent to the planner’s problem that gives solutions in Proposition 1. I define the following primitive quantities before solving the optimization problems. Denote $q_{At}$ and $q_{Bt}$ as the price of country A’s and B’s endowment goods, respectively, in units of the world numeraire. The price index, denoted by $P_{At}$ and $P_{Bt}$ for country A and B respectively, is defined to be the minimum price in units of the world numeraire to purchase a unit of domestic consumption basket. For example, the country A’s price index is solved by the
following minimization problem:

\[ P_{At} \equiv \min_{\{C_{At}^A, C_{At}^B\}} q_{At}C_{At}^A + q_{Bt}C_{At}^B \quad \text{s.t.} \quad C_{At} \equiv (C_{At}^A)^{\alpha} (C_{At}^B)^{1-\alpha} = 1. \]

The country B’s price index is found in a similar way. The solutions of price indices are

\[ P_{At} = \left( \frac{q_{At}}{\alpha} \right)^{\alpha} \left( \frac{q_{Bt}}{1-\alpha} \right)^{1-\alpha} \]
\[ P_{Bt} = \left( \frac{q_{Bt}}{\alpha} \right)^{\alpha} \left( \frac{q_{At}}{1-\alpha} \right)^{1-\alpha}. \]

Without loss of generality, let country A’s price index be the world numeraire. That is, \( P_{At} = 1 \) for all \( t \). The country A’s problem under the benchmark measure \( \mathcal{E} \) is

\[
\max_{\{C_{At}^A, C_{At}^B\}} \mathbb{E}_0^\mathcal{E} \left[ \int_0^\infty \eta_{At} e^{-\beta t} \log C_{At} dt \right]
\]
subject to the budget constraint

\[
\mathbb{E}_0^\mathcal{E} \left[ \int_0^\infty M_{At}^\mathcal{E} (q_{At} C_{At}^A + q_{Bt} C_{Bt}^A) dt \right] = \mathbb{E}_0^\mathcal{E} \left[ \int_0^\infty M_{At}^\mathcal{E} q_{At} X_{At} dt \right],
\]

where \( M_{At}^\mathcal{E} \) denotes country A’s pricing kernel under measure \( \mathcal{E} \). Notice that country A’s optimization problem includes a change-of-measure variable \( \eta_{At} \) so that the problem is equivalent to the one under A’s own measure.

The first-order conditions to country A’s problem are

\[
\eta_{At} e^{-\beta t} \frac{1}{C_{At}^A} = \nu_A M_{At}^\mathcal{E} q_{At}
\]
\[
\eta_{At} e^{-\beta t} (1 - \alpha) \frac{1}{C_{At}^B} = \nu_A M_{At}^\mathcal{E} q_{Bt},
\]

where \( \nu_A \) is the Lagrange multiplier on A’s budget constraint. Country B’s problem is defined similarly, and the FOCs are

\[
\eta_{Bt} e^{-\beta t} (1 - \alpha) \frac{1}{C_{At}^B} = \nu_B M_{At}^\mathcal{E} q_{At}
\]
\[
\eta_{Bt} e^{-\beta t} \frac{1}{C_{Bt}^B} = \nu_B M_{At}^\mathcal{E} q_{Bt}.
\]
We immediately obtain the optimal consumption allocations

\[
C_A^A = \frac{\eta_A e^{-\beta t} \alpha}{\nu_A M^E_{At} q_{At}}
\]

\[
C_B^A = \frac{\eta_A e^{-\beta t} (1 - \alpha)}{\nu_A M^E_{At} q_{At}}
\]

\[
C_A^B = \frac{\eta_B e^{-\beta t} (1 - \alpha)}{\nu_B M^E_{At} q_{At}}
\]

\[
C_B^B = \frac{\eta_B e^{-\beta t} \alpha}{\nu_B M^E_{At} q_{At}}
\]

Given the resource constraint that \(C_A^A + C_B^B = X_A\) and \(C_A^B + C_B^B = X_B\), we obtain \(M^E_{At} q_{At}\) and \(M^E_{At} q_{Bt}\) as follows.

\[
M^E_{At} q_{At} = e^{-\beta t} \frac{X_A}{X_A} \left[ \frac{\eta_A \alpha}{\nu_A} + \frac{\eta_B (1 - \alpha)}{\nu_B} \right]
\] (37)

\[
M^E_{At} q_{Bt} = e^{-\beta t} \frac{X_B}{X_B} \left[ \frac{\eta_A (1 - \alpha)}{\nu_A} + \frac{\eta_B \alpha}{\nu_B} \right]
\] (38)

The ratio \(k_t\) of prices for endowment goods is therefore

\[
k_t \equiv \frac{q_{At}}{q_{Bt}} = X_B / X_A \left[ \frac{\eta_A \alpha / \nu_A + \eta_B (1 - \alpha) / \nu_B}{\eta_A (1 - \alpha) / \nu_A + \eta_B \alpha / \nu_B} \right].
\] (39)

Since we normalize \(P_{At}\) to be one, we get the prices \(q_{At}\) and \(q_{Bt}\) as\(^{18}\)

\[
q_{At} = \alpha^\alpha (1 - \alpha)^{1-\alpha} k_t^{1-\alpha} \]

\[
q_{Bt} = \alpha^\alpha (1 - \alpha)^{1-\alpha} k_t^{-\alpha}
\] (41)

To show the relation between country A’s pricing kernel under measure \(E\) and their consumption basket, we start from the definition of \(C_{At}\). That is, \(C_{At} = (C_{At}^A)^\alpha (C_{At}^B)^{1-\alpha}\). From the formulas above, we can see that

\[
C_{At} = (C_{At}^A)^\alpha (C_{At}^B)^{1-\alpha}
\]

\[= \frac{e^{-\beta t} \eta_A \left( \frac{\alpha}{q_{At}} \right)^\alpha \left( \frac{1 - \alpha}{q_{Bt}} \right)^{1-\alpha}}{\nu_A M^E_{At}} \]

\[= \frac{e^{-\beta t} \eta_A}{\nu_A M^E_{At}}
\]

\(^{18}\)The functional forms of \(q_{At}\) and \(q_{Bt}\) do not depend on the risk aversion parameter, though the values of \(k_t\) will do.
Thus, we obtain a familiar expression as in the main text

\[ M_{At}^\varepsilon = e^{-\beta t} \frac{\eta_{At}}{\nu_{A}C_{At}}. \]

Note that the Lagrange multiplier is constant and does not affect pricing.\(^{19}\) The expression for \( M_{Bt}^\varepsilon \) can be obtained in a similar way.

In order to derive explicit expressions for \( M_{At}^\varepsilon \) and \( C_{At} \), I combine the equations (37) and (40). After some algebra, the same expression is obtained for \( C_{At} \) as in Section 3 with the Lagrange multiplier to be the inverse of Pareto weights. Country B’s consumption basket can be derived similarly. Thus, the equilibrium allocations in the competitive equilibrium are identical to those in the planner’s problem.

For a general CRRA utility with \( \gamma \neq 1 \), the above analysis carries over with slight modifications. Because the prices \( q_A \) and \( q_B \) do not explicitly depend on \( \gamma \), we still have

\[ \left( \frac{\sigma}{q_{At}} \right)^\alpha \left( \frac{1-\alpha}{q_{Bt}} \right)^{1-\alpha} = 1. \]

Thus, we get

\[ M_{At}^\varepsilon = \frac{e^{-\beta t}}{\nu_{A}} \eta_{At} C_{At}^{-\gamma} \]
\[ M_{Bt}^\varepsilon = \frac{e^{-\beta t}}{\nu_{B}} \eta_{Bt} C_{Bt}^{-\gamma}. \]

### C Simulated Methods of Moments

The remaining parameters are estimated using simulated methods of moments (SMM). I impose symmetry across the countries for \( \sigma_{\mu A,B}, \sigma_{\theta A,B} \) and \( \bar{\theta}_{A,B} \). This symmetric restriction will greatly reduce the computing time, and the number of estimated parameters decreases to only 5 instead of 8. There are 10 targeted moments in total, namely the mean and variance of real per capita GDP growth rates, mean and variance of real per capita consumption growth rates, cross-country consumption growth correlation, variance of real foreign exchange rate growth, mean and variance of real risk-free rates, and the mean and variance of equity excess returns. I simulate 5000 data sets with each containing 60 years of quarterly data with the first 10 years of data dropped in order to reduce the effects of initial conditions. The procedure for generating the estimated parameters are standard, and the weight matrix is chosen to be the inverse of variance-covariance matrix of empirical moments as usual. The estimated parameters as well as standard errors are given in Table 2.

\(^{19}\)The price of an asset payoff \( Y_T \) at time \( t < T \) is \( E_{t}^\varepsilon \left[ \frac{M_{At}^\varepsilon Y_T}{M_{At}^\varepsilon} \right] \). We see that the Lagrange multiplier is canceled out. Indeed, with symmetric setup, we can set the Lagrange multipliers to be one and equal to the Pareto weights.
D Asset Prices

Consider a generic risky asset that has a single payoff $Y_T$ at some future time $T > t$ denominated in $A$’s currency. Suppose the process of $Y$ follows

$$dY_t = \mu_Y(t, Y_t) dt + \sigma_Y(t, Y_t) \mathbf{d}W^E_t,$$

where $\mu_Y(t, Y_t)$ and $\sigma_Y(t, Y_t)$ are possibly time-varying and could depend on $Y_t$. The price $P^Y_t$ at time $t$ of such asset then satisfies

$$P^Y_t M^E_{At} = \mathbb{E}^E_t \left[ M^E_{At} Y_T \right]. \quad (42)$$

Notice that the expectation is taken under the benchmark measure and the SDF has been adjusted accordingly.

Suppose the process of price $P^Y$ satisfies

$$dP^Y_t = \mu_{P^Y}(t, P^Y_t) dt + \sigma_{P^Y}(t, P^Y_t) \mathbf{d}W^E_t. \quad (43)$$

Again, $\mu_{P^Y}(t, P^Y_t)$ and $\sigma_{P^Y}(t, P^Y_t)$ are possibly time-varying and could depend on $P^Y_t$. We obtain the following results.

Proposition 13. Given the price process (43), the drift $\mu_{P^Y}(t, P^Y_t)$ and volatility $\sigma_{P^Y}(t, P^Y_t)$ are

$$\mu_{P^Y}(t, P^Y_t) = \rho_{At} P^Y_t + \sigma^\top_{MAt} \Omega W \sigma_{P^Y}(t, P^Y_t),$$

$$\sigma_{P^Y}(t, P^Y_t) = \sigma_{MAt} M^E_{At} + \left( \mathbb{E}^E_t \left[ M^E_{At} Y_T \right] \right)^{-1} \mathbb{E}^E_t \left[ M^E_{At} \left( \frac{Y_T}{Y_t} \sigma_Y(t, Y_t) - Y_T \sigma_{MAt} \right) \right],$$

where

$$\tilde{Y}_t = \exp \left[ \int_0^t \left( \mu_Y(s, Y_s) - \frac{1}{2} \sigma_Y(s, Y_s) \sigma_Y(s, Y_s)^\top \right) ds + \int_0^t \sigma_Y(s, Y_s)^\top \mathbf{d}W_s \right].$$

Similar expressions can be derived for assets denominated in $B$’s currency. The derivations are provided in Appendix F.

D.0.1 Stock price

A stock or equity in this model is the claim on the country $i$’s domestic endowment stream $X_{it}$. Let $E^i_t$ be the country $i$’s stock price at time $t$ in units of global numeraire (country $A$’s consumption basket). By definition, it satisfies

$$E^i_t M^E_{At} = \mathbb{E}^E_t \left[ \int_t^\infty M^E_{As} q_{is} X_{is} ds \right], \quad (44)$$

42
where recall that $q_{it}$ is the price of $i$’s endowment good at time $t$ in units of country $A$’s currency. It is easy to show that the dynamics for the stock price is

$$
\frac{dE^i_t + q_{it}X_{it}dt}{E^i_t} = \mu_{E^i_t}dt + \sigma_{E^i_t}^\top dW^\xi_t,
$$

(45)

where $\mu_{E^i_t}$ and $\sigma_{E^i_t}$ denote stock expected return and volatility, respectively. Using the results in Proposition 13, we summarize the expected and volatilities of stock price as follows.

**Proposition 14.** The country $i$’s stock price $E^i_t$, in units of the global numeraire, follows

$$
\begin{align*}
\mu_{E^i_t} &= r_A + \sigma_{MA}^\top \Omega W \sigma_{E^i_t} \\
\sigma_{E^i_t} &= \sigma_{MA} + \mathbb{E}^\xi [\int_t^\infty \tilde{D}_t \left( M_{As} q_{is} X_{is} \right) ds] / E^i_t M_{At}^\xi,
\end{align*}
$$

where $\tilde{D}_t(\cdot)$ denotes the Malliavin derivative. The derivations are provided in Appendix F.

## E Basics of Malliavin Derivative

The Malliavin derivative extends the notion of differentiation to the classical Wiener space, which is usually a collection of random variables that are not differentiable in the usual sense. Suppose a random variable $X_t$ has the following SDE

$$
\begin{align*}
dX_t &= \mu(t, X_t) dt + \sigma(t, X_t)^\top dW_t, \quad X_0 = x.
\end{align*}
$$

To calculate its Malliavin derivative $D_sX_t$, one needs to find the first variation process of $X_t$. Let $\tilde{X}_t$ be the first variation, which satisfies

$$
\begin{align*}
d\tilde{X}_t &= \mu'(t, X_t) \tilde{X}_tdt + \sigma'(t, X_t)^\top \tilde{X}_tdW_t,
\end{align*}
$$

or

$$
\begin{align*}
\tilde{X}_t = \exp \left[ \int_0^t \left( \mu'(s, X_s) - \frac{1}{2} \sigma'(s, X_s)^\top \sigma'(s, X_s) \right) ds + \int_0^t \sigma'(s, X_s)^\top dW_s \right],
\end{align*}
$$

where $\mu'(s, X_s)$ and $\sigma'(s, X_s)$ denote derivatives with respect to $X_s$ in the usual sense. Then we have

$$
D_sX_t = \frac{\tilde{X}_t}{X_s} \sigma(t, X_t) \mathbb{1}_{\{s \leq t\}}.
$$

The following theorem will be particularly useful.

**Theorem 2 (Clark-Ocone).** If $F \in \mathbb{D}^{1,2}$, where $\mathbb{D}^{1,2}$ is the closure of the space of smooth
and cylindrical random variables under the norm

\[ \| F \|_{1,2} = \left[ \mathbb{E}(F^2) + \mathbb{E} \left( \int_0^T (D_sF)^2 ds \right) \right]^{\frac{1}{2}}, \]

then we have

\[ F = \mathbb{E}(F) + \int_0^T \mathbb{E}(D_sF|\mathcal{F}_s)dW_s. \]

**F  Proofs**

**Proof of Proposition 1**

*Proof.* The welfare optimization in the planner’s problem is

\[
\max_{\{C^A_{At}, C^B_{At}, C^A_{Bt}, C^B_{Bt}\}} \mathbb{E}_0^\mathcal{E} \left[ \int_0^\infty e^{-\beta t} \left( \lambda_A \eta^A_t \log C^A_{At} + \lambda_B \eta^B_t \log C^B_{Bt} \right) dt \right]
\]

(46)

subject to the resource constraints

\[
C^A_{At} + C^B_{At} = X_{At} \quad \quad C^A_{Bt} + C^B_{Bt} = X_{Bt}.
\]

(47)

Note that the household’s original problem (2) is defined under their own probability measure. The change-of-measure variable \( \eta_t^i \) in the planner’s optimization is therefore necessary in order to achieve the same consumption allocations as those obtained in individual household’s optimization. I can alternatively express, for example, the country A’s optimization problem (2) under measure \( \mathcal{E} \) as

\[
\max_{\{C^A_{At}, C^A_{Bt}\}} \mathbb{E}_0^\mathcal{E} \left[ \int_0^\infty e^{-\beta t} \eta^A_t \log C^A_{At} dt \right].
\]

Denote \( \nu_{At} \) as the Lagrange multiplier on the resource constraint (47). The FOCs for country A in the planner’s problem (46) imply

\[
\lambda_A \eta^A_t \alpha \frac{1}{C^A_{At}} = \nu_{At}
\]

\[
\lambda_A \eta^A_t \alpha (1 - \alpha) \frac{1}{C^A_{Bt}} = \nu_{Bt},
\]

and for country B

\[
\lambda_B \eta^B_t (1 - \alpha) \frac{1}{C^B_{At}} = \nu_{At}
\]

\[
\lambda_B \eta^B_t \alpha \frac{1}{C^B_{Bt}} = \nu_{Bt}.
\]
Dividing the two equations above for each $\nu_t$ and using (47), we achieve the equilibrium consumption allocations for country $A$ as

\[
C_A^t = \frac{\lambda_A \alpha \eta_t^A}{\lambda_A \eta_t^A \alpha + \lambda_B \eta_t^B (1 - \alpha)} X_A^t
\]  
(48)

\[
C_B^t = \frac{\lambda_B (1 - \alpha) \eta_t^B}{\lambda_A \eta_t^A (1 - \alpha) + \lambda_B \eta_t^B \alpha} X_B^t,
\]  
(49)

and for country $B$ as

\[
C_A^t = \frac{\lambda_A \alpha \eta_t^A}{\lambda_A \eta_t^A \alpha + \lambda_B \eta_t^B (1 - \alpha)} X_A^t
\]  
(50)

\[
C_B^t = \frac{\lambda_B \alpha \eta_t^B}{\lambda_A \eta_t^A (1 - \alpha) + \lambda_B \eta_t^B \alpha} X_B^t.
\]  
(51)

Then we immediately obtain the optimal consumption baskets.

I now derive the equilibrium process for $C_A^t$, and the analogous argument will apply for $C_B^t$. Define

\[
x(\eta_t) \equiv \lambda_A \alpha \eta_t^A + \lambda_B (1 - \alpha) \eta_t^B
\]

\[
y(\eta_t) \equiv \lambda_A (1 - \alpha) \eta_t^A + \lambda_B \alpha \eta_t^B.
\]

Taking log on both sides of (5), we get

\[
\log C_A^t = \alpha [\log X_A^t - \log x(\eta_t)] + (1 - \alpha) [\log X_B^t - \log y(\eta_t)] + \log \eta_t^A + \text{const}.
\]

Applying Ito’s lemma, we have

\[
d \log C_A^t = \frac{dC_A^t}{C_A^t} - \frac{1}{2} \frac{1}{C_A^t} (dC_A^t)^2
\]

\[
= \mu_C^A dt + \sigma_C^A dZ_A^t + \sigma_C^A dZ_B^t + \sigma_C^A \sigma_C^B \mu_t^C dW_{\mu_t}^g
\]

\[
- \frac{1}{2} \left( \sigma_C^A \sigma_C^A + \sigma_C^B \sigma_C^B + 2 \rho \sigma_C^A \sigma_C^B + \sigma_C^A \sigma_C^B \right) dt
\]

\[
d \log X_A^t = \frac{dX_A^t}{X_A^t} - \frac{1}{2} \frac{1}{X_A^t} (dX_A^t)^2
\]

\[
= \mu_X^A dt + \sigma_X^A dZ_A^t - \frac{1}{2} \sigma_X^2 dt
\]

\[
d \log X_B^t = \frac{dX_B^t}{X_B^t} - \frac{1}{2} \frac{1}{X_B^t} (dX_B^t)^2
\]

\[
= \mu_X^B dt + \sigma_X^B dZ_B^t - \frac{1}{2} \sigma_X^2 dt
\]
\[
d d \log \eta_t^A = \frac{d \eta_t^A}{\eta_t^A} - \frac{1}{2} \left( \frac{d \eta_t^A}{\eta_t^A} \right)^2 = -\xi_{At}^T dW_{\mu t} - \frac{1}{2} \xi_{At}^T \xi_{At} dt \\
\]
\[
d \log \eta_t^B = \frac{d \eta_t^B}{\eta_t^B} - \frac{1}{2} \left( \frac{d \eta_t^B}{\eta_t^B} \right)^2 = -\xi_{Bt}^T dW_{\mu t} - \frac{1}{2} \xi_{Bt}^T \xi_{Bt} dt \\
\]
and
\[
d \log (\eta_t) = \frac{\lambda A^\alpha}{x(\eta_t)} d\eta_t^A + \frac{\lambda B (1 - \alpha)}{x(\eta_t)} d\eta_t^B - \frac{1}{2} \frac{\lambda^2 A^\alpha}{x(\eta_t)^2} \left( \frac{d \eta_t^A}{\eta_t^A} \right)^2 - \frac{1}{2} \frac{\lambda^2 B (1 - \alpha)^2}{x(\eta_t)^2} \left( \frac{d \eta_t^B}{\eta_t^B} \right)^2 = -\xi_{At}^T dW_{\mu t} - \frac{1}{2} \xi_{At}^T \xi_{At} dt \\
d \log (\eta_t^B) = \frac{\lambda A^\alpha}{y(\eta_t)} d\eta_t^A + \frac{\lambda B^\alpha}{y(\eta_t)} d\eta_t^B - \frac{1}{2} \frac{\lambda^2 A^\alpha}{y(\eta_t)^2} \left( \frac{d \eta_t^A}{\eta_t^A} \right)^2 - \frac{1}{2} \frac{\lambda^2 B^\alpha}{y(\eta_t)^2} \left( \frac{d \eta_t^B}{\eta_t^B} \right)^2 = -\xi_{Bt}^T dW_{\mu t} - \frac{1}{2} \xi_{Bt}^T \xi_{Bt} dt \\
\]
Matching coefficients, we have for $C_A$ that
\[
\mu_{Ct} = \alpha \mu_t + (1 - \alpha) \mu_{Bt} - \frac{1}{2} \sigma_{C_A}^2 = \frac{\lambda A^\alpha}{x(\eta_t)^2} \xi_{At}^T \xi_{At} + \frac{1}{2} \frac{\lambda^2 A^\alpha}{x(\eta_t)^2} \left( \frac{d \eta_t^A}{\eta_t^A} \right)^2 = -\xi_{Bt}^T dW_{\mu t} - \frac{1}{2} \xi_{Bt}^T \xi_{Bt} dt \\
\]
\[
\sigma_{C_A} = \frac{\lambda A^\alpha}{x(\eta_t)^2} \xi_{At}^T \xi_{At} + \frac{1}{2} \frac{\lambda^2 A^\alpha}{x(\eta_t)^2} \left( \frac{d \eta_t^A}{\eta_t^A} \right)^2 = -\xi_{Bt}^T dW_{\mu t} - \frac{1}{2} \xi_{Bt}^T \xi_{Bt} dt \\
\]
\[
\sigma_{C_B} = (1 - \alpha) \sigma_B \\
\]
\[
\sigma_{C_{\mu t}} = -\xi_{At}^T + \frac{\lambda A^\alpha}{x(\eta_t)^2} \xi_{At}^T \xi_{At} + \frac{1}{2} \frac{\lambda^2 A^\alpha}{x(\eta_t)^2} \left( \frac{d \eta_t^A}{\eta_t^A} \right)^2 = -\xi_{Bt}^T dW_{\mu t} - \frac{1}{2} \xi_{Bt}^T \xi_{Bt} dt \\
\]
\[
\}
\]
and for $C_B$ that

$$
\begin{align*}
\mu_{C_B t} &= \alpha \mu_{B t} + (1 - \alpha) \mu_{A t} - \frac{1}{2} \alpha \sigma_B^2 - \frac{1}{2} (1 - \alpha) \sigma_A^2 - \frac{1}{2} \xi_{B t} \xi_{B t} \\
+ &\frac{1}{2} \lambda_A^2 (1 - \alpha) (\eta_t^A)^2 \xi_{A t} \xi_{A t} + \frac{1}{2} \lambda_B^2 (1 - \alpha) (\eta_t^B)^2 \xi_{B t} \xi_{B t} \\
+ &\frac{1}{2} \lambda_A^2 (1 - \alpha)^2 (\eta_t^A)^2 \xi_{A t} \xi_{A t} + \frac{1}{2} \lambda_B^2 (1 - \alpha)^3 (\eta_t^B)^2 \xi_{B t} \xi_{B t} \\
+ &\frac{1}{2} \left( \sigma_A^2 \sigma_B^2 + 2 \rho \sigma_A \sigma_B^2 + \sigma_{\mu t}^B \sigma_{\mu t}^B \right) \\
\sigma_{A t}^{C_B} &= (1 - \alpha) \sigma_A \\
\sigma_{B t}^{C_B} &= \alpha \sigma_B \\
\sigma_{\mu t}^{C_B} &= -\xi_{B t} + \frac{\lambda_A \alpha (1 - \alpha) \eta_t^A}{\lambda_A \eta_t^A + \lambda_B \eta_t^B (1 - \alpha)} \xi_{A t} + \frac{\lambda_B (1 - \alpha)^3 \eta_t^B}{\lambda_A \eta_t^A (1 - \alpha) + \lambda_B \eta_t^B \alpha} \xi_{B t} \\
+ &\frac{\lambda_B (1 - \alpha)^3 \eta_t^B}{\lambda_A \eta_t^A (1 - \alpha) + \lambda_B \eta_t^B \alpha} \xi_{B t} + \frac{\lambda_B \alpha^2 \eta_t^B}{\lambda_A \eta_t^A (1 - \alpha) + \lambda_B \eta_t^B \alpha} \xi_{B t}.
\end{align*}
$$

Proof of Proposition 2

Proof. Taking log on both sides of (8) for $i = A$, we have $\log M_{A t}^\xi = \beta t + \log \eta_t^A - \log C_{A t}$. Applying Ito’s lemma gives

$$
\begin{align*}
d \log M_{A t}^\xi &= \frac{dM_{A t}^\xi}{M_{A t}^\xi} - \frac{1}{2} \frac{1}{M_{A t}^2} (dM_{A t})^2 \\
&= -\beta dt + d \log \eta_t^A - d \log C_{A t} \\
&= \mu_{A t} dt - \sigma_A^2 dZ_{A t} - \sigma_B^2 dZ_{B t} - \sigma_{\mu t}^A dW_{\mu t} \\
&\quad - \frac{1}{2} \left( \sigma_A^2 + \sigma_B^2 + 2 \rho \sigma_A \sigma_B + \sigma_{\mu t}^A \sigma_{\mu t}^B \right) dt.
\end{align*}
$$

Matching coefficients, this yields

$$
\begin{align*}
\mu_{A t} &= -\beta - \mu_{C A t} - \frac{1}{2} \xi_{A t} \xi_{A t} + \frac{1}{2} \left( \sigma_A^2 + \sigma_B^2 + 2 \rho \sigma_A \sigma_B + \sigma_{\mu t}^A \sigma_{\mu t}^B \right) \\
+ &\frac{1}{2} \left( \sigma_A^2 + \sigma_B^2 + 2 \rho \sigma_A \sigma_B + \sigma_{\mu t}^A \sigma_{\mu t}^B \right) \\
\sigma_A^{MA} &= \sigma_A^2 \\
\sigma_B^{MA} &= \sigma_B^2 \\
\sigma_{\mu t}^{MA} &= \xi_{A t} + \sigma_{\mu t}^A.
\end{align*}
$$
Similarly, the coefficients for $M_{Bt}$ are
\[
\mu_{M_{Bt}} = -\beta - \mu_{C_{Bt}} - \frac{1}{2} \xi_{Bt}^T \xi_{Bt} + \frac{1}{2} \left( \sigma_{A_{B}}^2 + \sigma_{B_{B}}^2 + 2 \rho \sigma_{A_{B}} \sigma_{C_{B}} + \sigma_{\mu_{t}}^T \sigma_{\mu_{t}} \right) \\
+ \frac{1}{2} \left( \sigma_{A_{B}}^2 + \sigma_{B_{B}}^2 + 2 \rho \sigma_{A_{B}} \sigma_{C_{B}} + \sigma_{\mu_{t}}^T \sigma_{\mu_{t}} \right)
\]
\[
\sigma_{A_{B}} = \sigma_{A_{B}} \\
\sigma_{B_{B}} = \sigma_{B_{B}} \\
\sigma_{\mu_{t}} = \xi_{Bt} + \sigma_{\mu_{t}}.
\]

Proof of Proposition 3

Proof. Taking log on both sides of exchange rate equation in Proposition 3, we have
\[
\log S_t = \log \eta_t^B - \log \eta_t^A + \log C_{At} - \log C_{Bt}.
\]
Applying Ito’s lemma, we get
\[
d \log S_t = \frac{dS_t}{S_t} - \frac{1}{2} \left( \frac{dS_t}{S_t} \right)^2 \\
= \log \eta_t^B - \log \eta_t^A + d \log C_{At} - d \log C_{Bt} \\
= \mu_{St}dt + \sigma_{St}^A dZ_{At} + \sigma_{St}^B dZ_{Bt} + \sigma_{\mu_t}^S dW_{\mu_t}^\xi \\
- \frac{1}{2} \left( \sigma_{St}^A + \sigma_{St}^B + 2 \rho \sigma_{St}^A \sigma_{St}^B + \sigma_{\mu_t}^S \sigma_{\mu_t}^S \right) dt.
\]
Matching coefficients, this yields
\[
\mu_{St} = \mu_{C_{At}} - \mu_{C_{Bt}} + \frac{1}{2} \xi_{At}^T \xi_{At} - \frac{1}{2} \xi_{Bt}^T \xi_{Bt} + \frac{1}{2} \left( \sigma_{St}^A + \sigma_{St}^B + 2 \rho \sigma_{St}^A \sigma_{St}^B + \sigma_{\mu_t}^S \sigma_{\mu_t}^S \right) \\
- \frac{1}{2} \left( \sigma_{C_{A}}^2 + \sigma_{C_{B}}^2 + 2 \rho \sigma_{C_{A}} \sigma_{C_{B}} + \sigma_{\mu_t}^C \sigma_{\mu_t}^C \right) + \frac{1}{2} \left( \sigma_{C_{B_{A}}}^2 + \sigma_{C_{B_{B}}}^2 + 2 \rho \sigma_{C_{B_{A}}} \sigma_{C_{B_{B}}} + \sigma_{\mu_t}^C \sigma_{\mu_t}^C \right)
\]
\[
\sigma_{A} = \sigma_{A}^C - \sigma_{A}^B \\
\sigma_{B} = \sigma_{B}^C - \sigma_{B}^B \\
\sigma_{\mu_t} = \xi_{At} - \xi_{Bt} + \sigma_{\mu_t}^C - \sigma_{\mu_t}^B.
\]
It is easy to show that $\mu_{St} = r_{At} - r_{Bt} + \sigma_{\mu_t}^\top \Omega_W (\sigma_{MA_t} - \sigma_{MB_t})$, where, for example,
\[
\Omega_W = \begin{bmatrix}
1 & \rho & 0 & 0 \\
\rho & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]
when $dW_{\mu_t}^\xi$ is two-dimensional.
Proof of Proposition 4

Proof. Taking the difference between $\sigma^C_{\mu t} A$ and $\sigma^C_{\mu t} B$, we get

$$\sigma^C_{\mu t} A - \sigma^C_{\mu t} B = \left[ \frac{\lambda A\alpha(2\alpha - 1)}{\lambda A\eta^A_t + \lambda B\eta^B_t(1 - \alpha)} + \frac{\lambda A(1 - \alpha)(1 - 2\alpha)}{\lambda A\eta^A_t(1 - \alpha) + \lambda B\eta^B_t\alpha} \right] (\xi_{At} - \xi_{Bt})\eta^A_t$$

$$- \left[ \frac{\lambda B(1 - \alpha)(2\alpha - 1)}{\lambda A\eta^A_t + \lambda B\eta^B_t(1 - \alpha)} + \frac{\lambda B\alpha(1 - 2\alpha)}{\lambda A\eta^A_t(1 - \alpha) + \lambda B\eta^B_t\alpha} \right] (\xi_{At} - \xi_{Bt})\eta^B_t$$

$$= \frac{(\lambda A\eta^A_t + \lambda B\eta^B_t(1 - \alpha))}{(2\alpha - 1)^2\lambda A\lambda B\eta^A_t\eta^B_t} \left( \xi_{At} - \xi_{Bt} \right) \left( \xi_{At} - \xi_{Bt} \right)$$

$$\equiv \frac{H_t}{K_t} (\xi_{At} - \xi_{Bt}) - (\xi_{At} - \xi_{Bt}).$$

We can see that after some algebra we can see that $K_t > H_t$. Therefore, we can express $\sigma^C_{\mu t} A - \sigma^C_{\mu t} B$ as $-f(\eta_t)(\xi_{At} - \xi_{Bt})$, where $0 < f(\eta_t) < 1$.

Thus, the last two terms in (13) is negative regardless of the sign of $\xi_{it}$ because the expressions only involve $(\xi_{At} - \xi_{Bt})^T(\xi_{At} - \xi_{Bt})$. \square

Proof of Proposition 5

Proof. Since the benchmark measure tracks the average belief and according to Definition 1 and equation (3), we have $\xi_{At} = -\xi_{Bt}$. With this relation, the result is then obtained by straightforward algebra. \square

Proof of expected currency risk premium (16)

Proof. The carry trade return in levels with a long position in bond $B$ and a short position in $A$ is

$$RX_{t+dt} = \frac{S_{t+dt}}{S_t} (1 + r_{Bt}dt) - (1 + r_{At}dt)$$

$$= \frac{S_{t+dt}(1 + r_{Bt}dt)}{S_t(1 + r_{At}dt)}, \text{ since } dt \text{ is small.}$$

Taking log on both sides and applying Ito’s lemma, we get

$$r x_{t+dt} = d \log S_t + r_{Bt} - r_{At}.$$

Using Proposition 3 and expressions for risk-free rates, we obtain (16).

Now I derive explicit expression in (17). The component for belief shocks can be written as

$$(\sigma^M_{\mu t} A + \sigma^M_{\mu t} B)^T (\sigma^M_{\mu t} A - \sigma^M_{\mu t} B) = (\sigma^C_{\mu t} A + \sigma^C_{\mu t} B)^T (\xi_{At} - \xi_{Bt} + \sigma^C_{\mu t} A - \sigma^C_{\mu t} B).$$

20This is because $\alpha^2 + (1 - \alpha)^2 - (2\alpha - 1)^2 = 2\alpha - 2\alpha^2 > 0$ as $\alpha < 1$. 49
Using expressions for $\sigma_{\mu_t}^A$ and $\sigma_{\mu_t}^B$, it is straightforward to show that with Definition 1, the RHS above equals

$$\frac{\alpha(1 - \alpha)\lambda_A\lambda_B(2\alpha - 1)^2}{x(\eta_t)y(\eta_t)}\eta_t^A\eta_t^B(\xi_{At} - \xi_{Bt})^\top(\xi_{At} - \xi_{Bt})[(\eta_t^A)^2 - (\eta_t^B)^2].$$

Therefore, we have

$$n(\eta_t) \equiv \frac{\alpha(1 - \alpha)\lambda_A\lambda_B(2\alpha - 1)^2}{2x(\eta_t)y(\eta_t)}\eta_t^A\eta_t^B(\eta_t^A + \eta_t^B)(\xi_{At} - \xi_{Bt})^\top(\xi_{At} - \xi_{Bt}).$$

The expected risk premium in levels from country A’s perspective is

$$\frac{1}{dt}\mathbb{E}_t^\varepsilon[RX_{t+dt}] = \frac{1}{dt}\mathbb{E}_t^\varepsilon\left[\left(\frac{dM_{At}^\varepsilon}{M_{At}^2}\right)^2 - \frac{dM_{At}^\varepsilon dM_{Bt}^\varepsilon}{M_{At}^2 M_{Bt}^2}\right]$$

$$= \frac{1}{dt}\text{Cov}_t\left(\frac{dM_{At}^\varepsilon}{M_{At}^2}, \frac{dM_{At}^\varepsilon}{M_{At}^2} - \frac{dM_{Bt}^\varepsilon}{M_{Bt}^2}\right)$$

$$= \sigma_{\mu_t}^M(\sigma_{\mu_t}^M - \sigma_{\mu_t}^M).$$

Writing this out explicitly gives

$$\frac{1}{dt}\mathbb{E}_t^\varepsilon[RX_{t+dt}] = \frac{\lambda_A\lambda_B(2\alpha - 1)^2}{x(\eta_t)y(\eta_t)}\eta_t^A\eta_t^B(\xi_{At} - \xi_{Bt})^\top(\xi_{At} - \xi_{Bt})$$

$$\times \left[\alpha(1 - \alpha)(\eta_t^A + \eta_t^B)(\eta_t^A - \eta_t^B) + (2\alpha - 1)^2\eta_t^A\eta_t^B\right].$$

Proof of Proposition 7

Proof. The derivation is straightforward by using (10) and Definition 1.

Proof of Proposition 9

Proof. I first apply Ito’s lemma, and obtain

$$d\log \omega_t = \frac{1}{\omega_t}d\omega_t - \frac{1}{2}\frac{1}{\omega_t^2}(d\omega_t)^2$$

$$= \frac{1}{\omega_t}\left(\mu_{\omega}dt + \sigma_{A\omega}^\omega dZ_{At} + \sigma_{B\omega}^\omega dZ_{Bt} + \sigma_{\mu\omega}^\omega dW_{\mu\omega}\right)$$

$$- \frac{1}{2}\frac{1}{\omega_t^2}\left(\sigma_{A\omega}^2 + \sigma_{B\omega}^2 + 2\rho\sigma_{A\omega}^\omega \sigma_{B\omega}^\omega + \sigma_{\mu\omega}^\omega \sigma_{\mu\omega}^\omega\right)dt.$$
Matching coefficients, we obtain the drift part as

\[ d \log \eta_t^A = -\xi_{At}^\top dW_{\mu t}^\varepsilon - \frac{1}{2} \xi_{At}^\top \xi_{At} dt \]

\[ d \log \eta_t^B = -\xi_{Bt}^\top dW_{\mu t}^\varepsilon - \frac{1}{2} \xi_{Bt}^\top \xi_{Bt} dt \]

\[ d \log X_{Bt} - d \log X_{At} = (\mu_{Bt} - \mu_{At}) dt + \sigma_B dZ_{Bt} - \sigma_A dZ_{At} - \frac{1}{2} (\sigma_B^2 - \sigma_A^2) dt \]

\[ d \log(1 + \omega_t) = \frac{1}{1 + \omega_t} \left( \mu_{\omega t} dt + \sigma_{\omega A t}^2 dZ_{At} + \sigma_{\omega B t}^2 dZ_{Bt} + \sigma_{\mu t}^\omega \sigma_{\mu t}^\omega \right) dt \]

\[ d \log(\alpha^2 + (1 - \alpha)^2 \omega_t) = \frac{1 - \alpha^2}{\alpha^2 + (1 - \alpha)^2 \omega_t} \left( \sigma_{\omega A t}^2 + \sigma_{\omega B t}^2 + 2 \rho \sigma_{\omega A t} \sigma_{\omega B t} + \sigma_{\mu t}^\omega \sigma_{\mu t}^\omega \right) dt \]

Taking logs on both sides of (35) and then differentiating both sides, we get

\[ 0 = -\gamma d \log \omega_t + d \log \eta_t^A - d \log \eta_t^B + (1 - 2\alpha)(1 - \gamma) (d \log X_{Bt} - d \log X_{At}) \]

\[ + (1 - 2\alpha)(1 - \gamma) \left[ d \log(1 + \omega_t) - d \log(\alpha^2 + (1 - \alpha)^2 \omega_t) \right]. \]

Matching coefficients, we obtain the drift part as

\[ \mu_{\omega t} = \left[ \frac{\gamma}{\omega_t} - (1 - 2\alpha)(1 - \gamma) \frac{1}{1 + \omega_t} + (1 - 2\alpha)(1 - \gamma) \frac{(1 - \alpha)^2}{\alpha^2 + (1 - \alpha)^2 \omega_t} \right]^{-1} \]

\[ \times \left[ (1 - 2\alpha)(1 - \gamma) (\mu_{Bt} - \mu_{At}) + \frac{\gamma}{\omega_t^2} \left( \sigma_{\omega A t}^2 + \sigma_{\omega B t}^2 + 2 \rho \sigma_{\omega A t} \sigma_{\omega B t} + \sigma_{\mu t}^\omega \sigma_{\mu t}^\omega \right) \right. \]

\[ - \frac{1}{2} (1 - 2\alpha)(1 - \gamma) (\sigma_B^2 - \sigma_A^2) - \frac{1}{2} \frac{(1 - 2\alpha)(1 - \gamma)}{(1 + \omega_t)^2} \left( \sigma_{\omega A t}^2 + \sigma_{\omega B t}^2 + 2 \rho \sigma_{\omega A t} \sigma_{\omega B t} + \sigma_{\mu t}^\omega \sigma_{\mu t}^\omega \right) \]

\[ + \frac{1}{2} (1 - 2\alpha)(1 - \gamma) \left( \sigma_{\omega A t}^2 + \sigma_{\omega B t}^2 + 2 \rho \sigma_{\omega A t} \sigma_{\omega B t} + \sigma_{\mu t}^\omega \sigma_{\mu t}^\omega \right)^{-1} \left( \sigma_{\omega A t}^2 + \sigma_{\omega B t}^2 + 2 \rho \sigma_{\omega A t} \sigma_{\omega B t} + \sigma_{\mu t}^\omega \sigma_{\mu t}^\omega \right) \]

and the volatilities as

\[ \sigma_{\omega A t}^2 = \left[ \frac{\gamma}{\omega_t} + (1 - 2\alpha)(1 - \gamma) \frac{1}{1 + \omega_t} - (1 - 2\alpha)(1 - \gamma) \frac{(1 - \alpha)^2}{\alpha^2 + (1 - \alpha)^2 \omega_t} \right]^{-1} (1 - 2\alpha)(1 - \gamma) \sigma_A \]

\[ \sigma_{\omega B t}^2 = \left[ \frac{\gamma}{\omega_t} + (1 - 2\alpha)(1 - \gamma) \frac{1}{1 + \omega_t} - (1 - 2\alpha)(1 - \gamma) \frac{(1 - \alpha)^2}{\alpha^2 + (1 - \alpha)^2 \omega_t} \right]^{-1} (1 - 2\alpha)(\gamma - 1) \sigma_B \]

\[ \sigma_{\mu t}^2 = \left[ \frac{\gamma}{\omega_t} + (1 - 2\alpha)(1 - \gamma) \frac{1}{1 + \omega_t} - (1 - 2\alpha)(1 - \gamma) \frac{(1 - \alpha)^2}{\alpha^2 + (1 - \alpha)^2 \omega_t} \right]^{-1} (\xi_{At} - \xi_{Bt}). \]
Proof of Proposition 10

Proof. I first take logs and differentiate both side of the consumption baskets given in Proposition 8. Then, using the same techniques by applying Ito’s lemma and matching coefficients, we obtain the drift parts for $C_{At}$ as

$$
\mu_{C_{At}} = \frac{\mu_{\omega t}}{\omega_t} - \frac{1}{2\omega_t^2} \left( \sigma_{At}^2 + \sigma_{\omega t}^2 + 2\rho \sigma_{At} \sigma_{\omega t} + \sigma_{\mu t}^\top \sigma_{\omega t} \right) + \alpha \mu_{At} - \frac{1}{2} \alpha \sigma_A^2 + (1 - \alpha) \mu_{Bt} - \frac{1}{2} (1 - \alpha) \sigma_B^2
$$

$$
- \frac{1}{1 + \omega_t} \alpha \mu_{\omega t} + \frac{1}{2} \alpha \left( \frac{1}{1 + \omega_t} \right)^2 \left( \sigma_{At}^2 + \sigma_{Bt}^2 + 2\rho \sigma_{At} \sigma_{Bt} + \sigma_{\mu t}^\top \sigma_{\mu t} \right)
$$

$$
- \frac{1 - \alpha}{\omega_t} \mu_{\omega t} + \frac{1}{2} \frac{1 - \alpha}{(1 - \alpha)^2 \omega_t} \left( \sigma_{At}^2 + \sigma_{Bt}^2 + 2\rho \sigma_{At} \sigma_{Bt} + \sigma_{\mu t}^\top \sigma_{\mu t} \right)
$$

$$
+ \frac{1}{2} \left( \sigma_{Ct}^A + \sigma_{Ct}^B + 2\rho \sigma_{Ct}^A \sigma_{Ct}^B + \sigma_{\mu t}^\top \sigma_{\mu t} \right).
$$

and for $C_{Bt}$ as

$$
\mu_{C_{Bt}} = (1 - \alpha) \mu_{At} - \frac{1}{2} (1 - \alpha) \sigma_A^2 + \alpha \mu_{Bt} - \frac{1}{2} \alpha \sigma_B^2
$$

$$
- \frac{1}{1 + \omega_t} (1 - \alpha) \mu_{\omega t} + \frac{1}{2} (1 - \alpha) \left( \frac{1}{1 + \omega_t} \right)^2 \left( \sigma_{At}^2 + \sigma_{Bt}^2 + 2\rho \sigma_{At} \sigma_{Bt} + \sigma_{\mu t}^\top \sigma_{\mu t} \right)
$$

$$
- \frac{1}{\omega_t} \alpha (1 - \alpha)^2 \mu_{\omega t} + \frac{1}{2} \frac{1}{\omega_t} \left( \frac{1}{\omega_t} \right)^2 \left( \sigma_{At}^2 + \sigma_{Bt}^2 + 2\rho \sigma_{At} \sigma_{Bt} + \sigma_{\mu t}^\top \sigma_{\mu t} \right)
$$

$$
+ \frac{1}{2} \left( \sigma_{Ct}^A + \sigma_{Ct}^B + 2\rho \sigma_{Ct}^A \sigma_{Ct}^B + \sigma_{\mu t}^\top \sigma_{\mu t} \right).
$$

The volatilities for $C_{At}$ and $C_{Bt}$ are

$$
\sigma_{Ct}^A = \alpha \sigma_A + \frac{1}{\omega_t} \sigma_{At} - \frac{1}{1 + \omega_t} \alpha \sigma_{At} - \frac{(1 - \alpha)^3}{\alpha^2 + (1 - \alpha)^2 \omega_t} \sigma_{At}
$$

$$
\sigma_{Ct}^A = (1 - \alpha) \sigma_B + \frac{1}{\omega_t} \sigma_{Bt} - \frac{1}{1 + \omega_t} \alpha \sigma_{Bt} - \frac{(1 - \alpha)^3}{\alpha^2 + (1 - \alpha)^2 \omega_t} \sigma_{Bt}
$$

$$
\sigma_{\mu t}^A = \frac{1}{\omega_t} \sigma_{\mu t} - \frac{1}{1 + \omega_t} \alpha \sigma_{\mu t} - \frac{(1 - \alpha)^3}{\alpha^2 + (1 - \alpha)^2 \omega_t} \sigma_{\mu t}
$$

$$
\sigma_{Ct}^B = (1 - \alpha) \sigma_A - \frac{1}{1 + \omega_t} \alpha \sigma_{At} - \frac{\alpha (1 - \alpha)}{\alpha^2 + (1 - \alpha)^2 \omega_t} \sigma_{At}
$$

$$
\sigma_{Ct}^B = \alpha \sigma_B - \frac{1}{1 + \omega_t} \alpha \sigma_{Bt} - \frac{\alpha (1 - \alpha)}{\alpha^2 + (1 - \alpha)^2 \omega_t} \sigma_{Bt}
$$

$$
\sigma_{\mu t}^B = \frac{1}{1 + \omega_t} (1 - \alpha) \sigma_{\mu t} - \frac{\alpha (1 - \alpha)^2}{\alpha^2 + (1 - \alpha)^2 \omega_t} \sigma_{\mu t}.
$$
Proof of Proposition 11

Proof. I first take logs and differentiate both side of pricing kernels. Then, using the same techniques by applying Ito’s lemma and matching coefficients, we obtain the drift and volatilities for \( M_A^t \) as

\[
\mu_{M_A^t} = -\beta - \gamma \mu_{C_A^t} - \frac{1}{2} \xi_{A^t}^\top \xi_{A^t} + \frac{1}{2} \gamma \left( \sigma_{A^t}^{C_A^2} + \sigma_{B}^{C_A^2} + 2 \rho \sigma_{A}^{C_A} \sigma_{B}^{C_A} + \sigma_{\mu t}^{C_A} \sigma_{\mu t}^{C_A} \right) \\
+ \frac{1}{2} \left( \sigma_{A}^{M_A^2} + \sigma_{B}^{M_A^2} + 2 \rho \sigma_{A}^{M_A} \sigma_{B}^{M_A} + \sigma_{\mu t}^{M_A} \sigma_{\mu t}^{M_A} \right)
\]

\[
\sigma_{A}^{M_A} = \gamma \sigma_{A}^{C_A} \\
\sigma_{B}^{M_A} = \gamma \sigma_{B}^{C_A} \\
\sigma_{\mu t}^{M_A} = \xi_{A^t} + \gamma \sigma_{\mu t}^{C_A}
\]

The drift and volatilities for \( M_B^t \) are

\[
\mu_{M_B^t} = -\beta - \gamma \mu_{C_B^t} - \frac{1}{2} \xi_{B^t}^\top \xi_{B^t} + \frac{1}{2} \gamma \left( \sigma_{A}^{C_B^2} + \sigma_{B}^{C_B^2} + 2 \rho \sigma_{A}^{C_B} \sigma_{B}^{C_B} + \sigma_{\mu t}^{C_B} \sigma_{\mu t}^{C_B} \right) \\
+ \frac{1}{2} \left( \sigma_{A}^{M_B^2} + \sigma_{B}^{M_B^2} + 2 \rho \sigma_{A}^{M_B} \sigma_{B}^{M_B} + \sigma_{\mu t}^{M_B} \sigma_{\mu t}^{M_B} \right)
\]

\[
\sigma_{A}^{M_B} = \gamma \sigma_{A}^{C_B} \\
\sigma_{B}^{M_B} = \gamma \sigma_{B}^{C_B} \\
\sigma_{\mu t}^{M_B} = \xi_{B^t} + \gamma \sigma_{\mu t}^{C_B}
\]

\[ \Box \]

Proof of Proposition 12

Proof. By the same techniques, we obtain the drift and volatilities as

\[
\mu_{S_t} = \gamma \mu_{C_A^t} - \gamma \mu_{C_B^t} - \frac{1}{2} \gamma \left( \sigma_{A}^{C_A^2} + \sigma_{B}^{C_A^2} + 2 \rho \sigma_{A}^{C_A} \sigma_{B}^{C_A} + \sigma_{\mu t}^{C_A} \sigma_{\mu t}^{C_A} \right) \\
+ \frac{1}{2} \gamma \left( \sigma_{A}^{C_B^2} + \sigma_{B}^{C_B^2} + 2 \rho \sigma_{A}^{C_B} \sigma_{B}^{C_B} + \sigma_{\mu t}^{C_B} \sigma_{\mu t}^{C_B} \right) + \frac{1}{2} \left( \sigma_{A}^{S^2} + \sigma_{B}^{S^2} + 2 \rho \sigma_{A}^{S} \sigma_{B}^{S} + \sigma_{\mu t}^{S} \sigma_{\mu t}^{S} \right)
\]

\[
\sigma_{A}^{S} = \gamma \sigma_{A}^{C_A} - \gamma \sigma_{A}^{C_B} \\
\sigma_{B}^{S} = \gamma \sigma_{A}^{C_A} - \gamma \sigma_{B}^{C_B} \\
\sigma_{\mu t}^{S} = \xi_{A^t} - \xi_{B^t} + \gamma \sigma_{\mu t}^{C_A} - \gamma \sigma_{\mu t}^{C_B}
\]

\[ \Box \]
Proof of Proposition 13

Proof. I first note that when $T$ is fixed $P_t^Y M_{At}^ε$ is an $\mathcal{E}$-martingale. Applying Ito’s lemma to the both sides of (42), we get the following for the LHS

$$P_t^Y M_{At}^ε \left( σ_{Py} (t, P_t^Y) - σ_{MAt} M_{At}^ε \right) ^\top dW_t^ε.$$

I then apply Theorem 2 to the RHS of (42) and obtain the RHS as

$$E_t^ε \left[ D_t (M_{ATY}^ε) \right] ^\top dW_t^ε,$$

where $D_t$ denotes the Malliavin derivative operator defined in Appendix E.

To compute $D_t (M_{ATY}^ε)$, note that

$$D_t (M_{ATY}^ε) = M_{ATY}^ε \left( \frac{1}{M_{AT}^ε} D_t M_{AT}^ε + \frac{1}{Y_T^ε} D_t Y_T^ε \right).$$

The first variation processes for $M_{AT}^ε$ and $Y_t$ are given by

$$d\tilde{M}_{AT}^ε = \mu_{MAt} M_{At}^ε dt - σ_{MAt} ^\top M_{At}^ε dW_t^ε$$

$$d\tilde{Y}_t = \mu'_Y (t, Y_t) \tilde{Y}_t dt + σ'_Y (t, \tilde{Y}_t) ^\top \tilde{Y}_t dW_t^ε.$$

Therefore, the Malliavin derivatives are

$$D_t M_{AT}^ε = -M_{AT}^ε σ_{MAt}$$

$$D_t Y_T^ε = \frac{\tilde{Y}_T^ε}{Y_t^ε} σ_{Y} (t, Y_t).$$

The volatility of price $P_t^Y$ is then

$$σ_{Py} (t, P_t^Y) = σ_{MAt} M_{At}^ε + \frac{E_t^ε \left[ D_t (M_{AT}^ε Y_T^ε) \right]}{P_t^Y M_{At}^ε}$$

$$= σ_{MAt} M_{At}^ε + \frac{E_t^ε \left[ M_{AT}^ε Y_T^ε D_t (\log M_{AT}^ε + \log Y_T^ε) \right]}{P_t^Y M_{At}^ε}$$

$$= σ_{MAt} M_{At}^ε + \frac{1}{P_t^Y M_{At}^ε} E_t^ε \left[ M_{AT}^ε Y_T^ε \left( \frac{\tilde{Y}_T^ε}{Y_t^ε σ_{Y} (t, Y_t)} - σ_{MAt} \right) \right].$$

The drift of the price process has the following usual form as a result of zero drift on LHS of (42)

$$μ_{Py} (t, P_t^Y) = r_{At} P_t^Y + σ_{MAt} ^\top \Omega_w σ_{Py} (t, P_t^Y).$$

We can derive similar expressions for a generic asset denominated in $B$’s currency. □
**Proof of Proposition 14**

*Proof.* Let’s focus on country $A$’s stock. Applying Ito’s lemma on LHS of (44), we get the diffusion part as

$$E_t^A M_{At}^ε (σ_{EAt} - σ_{MAT})^T dW_t^ε.$$  \hfill (52)

By Theorem 2, the diffusion part of RHS of (44) is

$$E_t^ε \left[ \int_t^∞ D_t (M_{As}^ε q_A X_{As})ds \right],$$  \hfill (53)

where we know from Appendix B that

$$M_{At}^ε q_A X_{At} = e^{-βt} \left( \frac{η_{At} α C_{At}^{1-γ}}{ν_A} + \frac{η_{Bt} (1-α) C_{Bt}^{1-γ}}{ν_B} \right),$$

and $ν_{A,B}$ are Lagrange multipliers and set to equal $1/λ_{A,B}$. In the special case of log utility, we have

$$M_{At}^ε q_A X_{At} = e^{-βt} \left( \frac{η_{At} α}{ν_A} + \frac{η_{Bt} (1-α)}{ν_B} \right).$$

Let $η_{It}$ follow

$$\frac{dη_{It}}{η_{It}} = σ_{ηIt}^T dW_t^ε$$

where $σ_{ηAIt} = (0, 0, -ξ_{At}^T)^T$, $σ_{ηBIt} = (0, 0, -ξ_{Bt}^T)^T$. The first variation processes for $η_{It}$ and $C_{It}$ are

$$d\tilde{η}_{It} = σ_{ηIt}^T \tilde{η}_{It} dW_t^ε$$

$$d\tilde{C}_{It} = μ_{CIt} \tilde{C}_{It} dt + σ_{CIt}^T \tilde{C}_{It} dW_t^ε,$$

or

$$\tilde{η}_{It} = \exp \left[ \int_0^t \left( -\frac{1}{2} σ_{ηIs}^T σ_{ηIs} \right) ds + \int_0^t σ_{ηAs}^T dW_s^ε \right]$$

$$\tilde{C}_{It} = \exp \left[ \int_0^t \left( μ_{CIs} - \frac{1}{2} σ_{CIs}^T σ_{CIs} \right) ds + \int_0^t σ_{CIs}^T dW_s^ε \right].$$

When $γ \neq 1$, we have for $s ≥ t$

$$D_t (η_{As} C_{As}^{1-γ}) = C_{As}^{1-γ} D_t η_{As} + (1-γ) η_{As} C_{As}^{-γ} D_t C_s,$$
where

\[ D_t \eta_{As} = \eta_{As} \sigma_{\eta At} \]
\[ D_t C_{As} = C_{As} \sigma_{C At}. \]

The expression for \( D_t (\eta_{Bs} C_{Bs}^{1-\gamma}) \) can be obtained similarly. Substituting these expressions in (53) and equating (52) and (53), we obtain desired results. In particular, we have

\[ D_t \left( M_{As}^E q_{As} X_{As} \right) = e^{-\beta s} \left[ \lambda_A \alpha \left( \eta_{As} C_{As}^{1-\gamma} \sigma_{\eta At} + (1 - \gamma) \eta_{As} C_{As}^{1-\gamma} \sigma_{C At} \right) \\
\quad + \lambda_B (1 - \alpha) \left( \eta_{Bs} C_{Bs}^{1-\gamma} \sigma_{\eta Bt} + (1 - \gamma) \eta_{Bs} C_{Bs}^{1-\gamma} \sigma_{C Bt} \right) \right]. \]

The derivations for country \( B \)'s stock return (measured in \( A \)'s currency) and volatility are entirely analogous. \( \square \)