Real Exchange Rates with Heterogeneous Beliefs

Shixiang Xia*

January 23, 2020

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, and Zhengyang Jiang 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, Sergio Rebelo, 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) show that matching the moderate level of exchange rate volatility in the data requires a high degree of international risk sharing. Under the standard assumptions, this implies that the consumption growth rates 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 focuses on how heterogeneous beliefs in a simple, complete-market setting can overcome the perfect relation between exchange rates and consumption and abstracts away from either incomplete/segmented markets or rich preferences.[1] The model features two countries, each having a representative household and a country-specific output that comes from either endowment or production. 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, a general specification of belief differences will have both direct and indirect effects on the equilibrium exchange rate.

Under complete markets, the real exchange rate is simply the ratio of the two pricing

[1] See, for example, Bakshi, Cerrato and Crosby (2017) for an incomplete market setting and Colacito and Croce (2011) for a long-run risk framework.

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 outputs 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 essence of the paper is that with a general specification of belief processes, the model qualitatively resolves all three puzzles. 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, country A's consumption of the domestic good 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.

As shown in Brandt, Cochrane and Santa-Clara (2006), we need a high correlation between pricing kernels 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 growth rates.

The presence of heterogeneous beliefs introduces an additional premium to currency carry trade on top of output risks. The risk premium includes compensation for the output risks because countries have home bias and the output shocks are not perfectly correlated. The investors are also compensated for the risks that arise from disagreement that causes consumption shares to fluctuate. As a result, asset prices need to take into account different probability measures in the two countries. The additional risk generates a time-varying currency risk premium, thus explaining the carry trade.

The qualitative results show that under mild conditions any belief processes could potentially resolve all three puzzles, whether the setting is an endowment or production economy. To quantitatively examine the performance of the model, I provide a simple example of learning processes in a pure exchange economy 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 growth. 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 cannot easily distinguish each other's beliefs.

In order to test the empirical relevance of the model, I employ a novel data set developed by BlackRock Investment Institute and Systematic Active Equity. This data set contains average forecasts of GDP growth made by economists and practitioners as well as BlackRock's own predictions using machine learning. The former one is mapped to households' estimates and the latter one to the econometrician's. The time-varying wedge can then be backed out from this data with calibrated model parameters. To my best knowledge, this is the first paper that draws comparisons among countries' beliefs using this data set. The regression results on how the wedges affect consumptions, exchange rates, interest rates, and currency risk premium are in line with model predictions.

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. Section 6 provides empirical evidence in support of the model and Section 7 concludes.

2 The Model

This section presents a parsimonious continuous-time two-country model. A representative household within each country trades output from their country in a frictionless market.[2] 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 having a tradable output. The output X_{it} follows the process

$$\frac{dX_{it}}{X_{it}} = \mu_{it}dt + \sigma_i dZ_{it}, \tag{1}$$

where μ_{it} is the time-varying expected growth rate, and σ_i denotes the volatility of country i's output. The standard Brownian motions dZ_{At} and dZ_{Bt} are correlated with coefficient ρ 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 output goods and financial assets is frictionless.

Each household has log utility over a consumption basket consisting of traded goods from each country.[3] 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 α on the domestic good is assumed to be $\frac{1}{2} < \alpha < 1$ to address home bias.

The optimization problem for country i is

$$\max_{\{C_{At}^i, C_{Bt}^i\}} \mathbb{E}_0^i\left[\int_0^{\infty} e^{-\beta t}\log C_{it}dt\right] \tag{2}$$

[2] There is no restriction on the sources — outputs can come from either endowments or production.

[3] 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.

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 economy, market clearing immediately implies that consumption of each good equals its output: $C^A_{At} + C^B_{At} = X_{At}$ and $C^A_{Bt} + C^B_{Bt} = X_{Bt}$.

2.2 Heterogeneous beliefs

Households in this model have limited knowledge about the underlying economy. While the two households observe the output levels X_{At} and X_{Bt}, the expected growth rates μ_{At} and μ_{Bt} as well as the output shocks Z_{At} and Z_{Bt}, there are some *time-varying* economic fundamentals $\boldsymbol{\theta_t}$ 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, persistence of business cycles or productivity growth.[4] The model further assumes that the expected growth rate μ_{it} can be expressed as

$$d\mu_{it} = u\left(\mu_{it}, \boldsymbol{\theta_t}\right) dt + v_{\mu_i t} dZ_{\mu_i t},$$

where the volatility $v_{\mu_i t}$ can be a constant. The Brownian motions $\mathbf{Z}_{\boldsymbol{\mu}\mathbf{t}} = (Z_{\mu_A t}, Z_{\mu_B t})$ are not directly observed and are referred to as unobserved fundamental shocks.[5]

Assumption 1. *Brownian motions $\mathbf{Z}_{\boldsymbol{\mu}\mathbf{t}}$ are mutually independent of each other, if they are multi-dimensional, and of output shocks Z_{At} and Z_{Bt}.*

As the dynamics of expected growth rates depend on $\boldsymbol{\theta_t}$, 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.[6] 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 $\mathbf{Z}_{\boldsymbol{\mu}\mathbf{t}}$, which I denote $\mathbf{W}^{\mathbf{i}}_{\boldsymbol{\mu}\mathbf{t}}$ as the perceived fundamental shocks under the subjective probability measure of country i.[7]

[4] The quantitative exercise in Section 5 will give an example of unobserved fundamentals for calibration.

[5] The two countries may have the same expected growth rate. In this case, the unobserved fundamental shock is one-dimensional.

[6] A microfoundation of estimating $\boldsymbol{\theta_t}$ is provided in Section 5.

[7] Recall that countries observe Z_{At} and Z_{Bt}. Thus, they both agree on the distributions of these shocks.

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 $\mathbf{dW}^{\mathcal{E}}_{\boldsymbol{\mu}\mathbf{t}}$. I define the process $\boldsymbol{\xi_{it}}$ to relate $\mathbf{dW}^{\mathcal{E}}_{\boldsymbol{\mu}\mathbf{t}}$ to country i's Brownian motions $\mathbf{dW}^{\mathbf{i}}_{\boldsymbol{\mu}\mathbf{t}}$ and impose restrictions on $\boldsymbol{\xi_{it}}$ as follows.

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

$$\boldsymbol{\xi_{it}} dt \equiv \mathbf{dW}^{\mathbf{i}}_{\boldsymbol{\mu}\mathbf{t}} - \mathbf{dW}^{\mathcal{E}}_{\boldsymbol{\mu}\mathbf{t}}. \tag{3}$$

Assumption 2. *The process for $\boldsymbol{\xi_{it}}$ satisfies that* $\exp\left(\boldsymbol{\xi_{it}} - \frac{1}{2}\langle\boldsymbol{\xi_{it}}\rangle\right)$ *is a strictly positive martingale under the benchmark measure, where $\langle\boldsymbol{\xi_{it}}\rangle$ is the quadratic variation of $\boldsymbol{\xi_{it}}$.*

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

Theorem 1 (Change of measure)**.** *The random variable η^i_t defined to be*

$$\frac{d\eta^i_t}{\eta^i_t} = -\boldsymbol{\xi}^{\top}_{\boldsymbol{it}} \mathbf{dW}^{\mathcal{E}}_{\boldsymbol{\mu}\mathbf{t}} \tag{4}$$

is the change-of-measure that translates the benchmark measure into country i's. Moreover, Definition 1 implies that $\boldsymbol{\xi_{At}} + \boldsymbol{\xi_{Bt}} = \mathbf{0}$.

The random variable $\boldsymbol{\eta_t} \equiv \left(\eta^A_t, \eta^B_t\right)$ 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, $\mathbf{dW}^{\mathcal{E}}_{\boldsymbol{\mu}\mathbf{t}} = \mathbf{dW}^{\mathcal{A}}_{\boldsymbol{\mu}\mathbf{t}} = \mathbf{dW}^{\mathcal{B}}_{\boldsymbol{\mu}\mathbf{t}}$. Thus, $\eta^A_t = \eta^B_t$ and $\boldsymbol{\xi_{At}} = \boldsymbol{\xi_{Bt}} = \mathbf{0}$. As soon as the countries have a disagreement, η^A_t and η^B_t 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 $\mathbb{E}^i[Y]$ under i's measure for some random variable Y is equivalent to $\mathbb{E}^{\mathcal{E}}[Y\eta^i]$. From now on, $\boldsymbol{\eta_t}$ 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.[8]

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

$$C_{At} = \left[\frac{\lambda_A \alpha \eta_t^A}{\lambda_A \alpha \eta_t^A + \lambda_B (1-\alpha)\eta_t^B} X_{At}\right]^{\alpha} \left[\frac{\lambda_A (1-\alpha) \eta_t^A}{\lambda_A (1-\alpha) \eta_t^A + \lambda_B \alpha \eta_t^B} X_{Bt}\right]^{1-\alpha}, \tag{5}$$

and for country B is

$$C_{Bt} = \left[\frac{\lambda_B (1-\alpha) \eta_t^B}{\lambda_A \alpha \eta_t^A + \lambda_B (1-\alpha)\eta_t^B} X_{At}\right]^{1-\alpha} \left[\frac{\lambda_B \alpha \eta_t^B}{\lambda_A (1-\alpha) \eta_t^A + \lambda_B \alpha \eta_t^B} X_{Bt}\right]^{\alpha}. \tag{6}$$

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

$$\frac{dC_{it}}{C_{it}} = \mu_{C_i t} dt + \boldsymbol{\sigma}_{\boldsymbol{C_i t}}^{\top} \mathbf{dW_t}^{\mathcal{E}} = \mu_{C_i t} dt + \sigma_A^{C_i} dZ_{At} + \sigma_B^{C_i} dZ_{Bt} + \boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{C_i}\top} \mathbf{dW}_{\boldsymbol{\mu}\mathbf{t}}^{\mathcal{E}}, \tag{7}$$

where $\mathbf{W}_t^{\mathcal{E}} \equiv \left(Z_{At}, Z_{Bt}, \mathbf{W}_{\boldsymbol{\mu}\mathbf{t}}^{\mathcal{E}}\right)^{\top}$, and the drift and volatilities are given in Appendix H.

The change-of-measure variable $\boldsymbol{\eta_t}$ 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 output.[9] This will not be the case when countries have heterogeneous beliefs. For simplicity, consider a discrete-time case with a state ω at time t. If country A associates a higher probability of ω occurring at t *relative to* B (i.e. $\eta_t^A > \eta_t^B$), country A will allocate a *relatively* larger fraction of outputs to the state ω.

Equation (7) shows that the consumption processes are subject to both output shocks (dZ) and fundamental shocks ($\mathbf{dW}_{\boldsymbol{\mu}\mathbf{t}}^{\mathcal{E}}$) under the benchmark measure. The loadings on the

[8] See Appendix B for the derivation of the competitive equilibrium.

[9] The shares for country A are $\frac{\lambda_A \alpha}{\lambda_A \alpha + \lambda_B (1-\alpha)}$ for domestic good and $\frac{\lambda_A (1-\alpha)}{\lambda_A (1-\alpha) + \lambda_B \alpha}$ for foreign good. Similar expressions can be obtained for country B.

output shocks are familiar and constant. The novel shocks $\mathbf{dW}^{\mathcal{E}}_{\boldsymbol{\mu t}}$ arise endogenously through heterogeneous beliefs across countries since the optimal consumption allocations crucially depend on the relative belief difference between the two countries. Indeed, when households have the same belief, the loadings on $\mathbf{dW}^{\mathcal{E}}_{\boldsymbol{\mu t}}$ will become identically zero, and we are back to the traditional consumption-based asset pricing framework. Moreover, unlike the output shocks, the volatilities $\boldsymbol{\sigma}^{\boldsymbol{C_i}}_{\boldsymbol{\mu t}}$ 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 $\boldsymbol{\eta_t}$. The magnitudes of $\boldsymbol{\sigma}^{\boldsymbol{C_i}}_{\boldsymbol{\mu t}}$ increase in (absolute values of) changes in $\boldsymbol{\eta_t}$, meaning that the volatilities also incorporate the severity of the belief difference — the higher degree of heterogeneity, the higher volatilities on perceived fundamental shocks.

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.[10]

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

$$M^{\mathcal{E}}_{it} = e^{-\beta t}\frac{1}{C_{it}}\eta^i_t. \tag{8}$$

The process for the above SDF under the benchmark measure follows

$$\frac{dM^{\mathcal{E}}_{it}}{M^{\mathcal{E}}_{it}} = \mu_{M_it}dt - \boldsymbol{\sigma}^{\top}_{\boldsymbol{M_it}}\mathbf{dW}^{\mathcal{E}}_{\mathbf{t}} = \mu_{M_it}dt - \sigma^{M_i}_A dZ_{At} - \sigma^{M_i}_B dZ_{Bt} - \boldsymbol{\sigma}^{\boldsymbol{M_i}\top}_{\boldsymbol{\mu t}}\mathbf{dW}^{\mathcal{E}}_{\boldsymbol{\mu t}}. \tag{9}$$

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

The extra term η^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

[10] Consider a discrete-time example. Denote, for example, $M^{\mathcal{A}}_{At}$ 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 $\mathbb{E}^{\mathcal{A}}_t\left[M^{\mathcal{A}}_{A,t+1}R_{t+1}\right] = \mathbb{E}^{\mathcal{E}}_t\left[M^{\mathcal{E}}_{A,t+1}R_{t+1}\right]$. Since the equation above is true for any asset and markets are complete, we must have for any event ω that $\pi^{\mathcal{A}}_{t+1}(\omega)M^{\mathcal{A}}_{A,t+1}(\omega)R_{t+1}(\omega) = \pi^{\mathcal{E}}_{t+1}(\omega)M^{\mathcal{E}}_{A,t+1}(\omega)R_{t+1}(\omega)$, where $\pi^i_{t+1}(\omega)$ is the probability of ω under measure i. This implies $\frac{\pi^{\mathcal{A}}_{t+1}(\omega)}{\pi^{\mathcal{E}}_{t+1}(\omega)}M^{\mathcal{A}}_{A,t+1}(\omega) = M^{\mathcal{E}}_{A,t+1}(\omega)$.

and the benchmark. The price not only incorporates how much the households' utilities will increase with an additional unit of consumption, as captured by $\frac{1}{C_{it}}$, but also takes into account how likely the events will occur in the minds of the two households, which is embedded in η_t^i.

Like consumption allocations, the pricing kernels are also subject to output shocks (dZ) and perceived fundamental shocks ($\mathbf{dW}^{\mathcal{E}}_{\boldsymbol{\mu}\mathbf{t}}$). 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 $\mathbf{dW}^{\mathcal{E}}_{\boldsymbol{\mu}\mathbf{t}}$ 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.

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 Ω_W be the correlation matrix of $\mathbf{dW}^{\mathcal{E}}_{\mathbf{t}}$. Then, we have in equilibrium

$$r_{it} = \beta + \left(\mu_{C_i t} - \boldsymbol{\xi_{it}}^\top \boldsymbol{\sigma}^{\boldsymbol{C_i}}_{\boldsymbol{\mu t}}\right) - \boldsymbol{\sigma}^\top_{\boldsymbol{C_i t}} \Omega_W \boldsymbol{\sigma}_{\boldsymbol{C_i t}} \tag{10}$$

The term in parentheses is country i's expected consumption growth $\mu^i_{C_i t}$ under i's own measure.[11] On one hand when household i perceives a high expected growth $\mu^i_{C_i t}$, 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 rela-

[11] The volatilities are preserved when switching to another measure.

tion no longer holds under heterogeneous beliefs. The following result gives the equilibrium exchange rate and its dynamics.[12]

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,*

$$S_t = \frac{M^{\mathcal{E}}_{Bt}}{M^{\mathcal{E}}_{At}} = \frac{\eta^B_t}{\eta^A_t}\frac{M^{\mathcal{B}}_{Bt}}{M^{\mathcal{A}}_{At}} = \frac{\eta^B_t}{\eta^A_t}\frac{C_{At}}{C_{Bt}}, \tag{11}$$

where $M^i_{it} \equiv e^{-\beta t}\frac{1}{C_{it}}$ is country i's pricing kernel under measure i.[13] *Moreover, the exchange rate follows the process*

$$\frac{dS_t}{S_t} = \mu_{St}dt - \boldsymbol{\sigma}^{\top}_{\boldsymbol{St}}\mathbf{dW}^{\mathcal{E}}_{\mathbf{t}} = \mu_{St}dt + \sigma^S_A dZ_{At} + \sigma^S_B dZ_{Bt} + \boldsymbol{\sigma}^{\boldsymbol{S}}_{\boldsymbol{\mu t}}{}^{\top}\mathbf{dW}^{\mathcal{E}}_{\boldsymbol{\mu}\mathbf{t}},$$

where

$$\mu_{St} = r_{At} - r_{Bt} + \boldsymbol{\sigma}^{\top}_{\boldsymbol{M_A t}}\Omega_W(\boldsymbol{\sigma}_{\boldsymbol{M_A t}} - \boldsymbol{\sigma}_{\boldsymbol{M_B t}}). \tag{12}$$

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

The term η^B_t/η^A_t 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,

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

[13] Real exchange rate in a general CRRA utility with risk aversion γ is $S_t = \frac{\eta^B_t}{\eta^A_t}\left(\frac{C_{At}}{C_{Bt}}\right)^{\gamma}$.

the real exchange rate S_t moves in response to fluctuations in the belief difference $\boldsymbol{\eta_t}$. For example, once η_{At} fluctuates, the relative prices of outputs X_{At} and X_{Bt} also changes. Equation (44) in Appendix B shows that an increase in η_{At} makes country A's output more expensive. This is because as η_{At} increases while keeping η_{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 η_{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 η_t^B.

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 with a general specification of belief processes.

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 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 γ 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 Δs_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

$$d\log S_t = d\log \eta_t^B - d\log \eta_t^A + \Delta C_t.$$

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 growth is 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* ΔC_t *is given by*

$$\frac{1}{dt}Cov_t\left(d\log S_t, \Delta C_t\right) = \left(\sigma_A^{C_A} - \sigma_A^{C_B}\right)^2 + \left(\sigma_B^{C_A} - \sigma_B^{C_B}\right)^2 + 2\rho\left(\sigma_A^{C_A} - \sigma_A^{C_B}\right)\left(\sigma_B^{C_A} - \sigma_B^{C_B}\right)$$
$$+ \left(\boldsymbol{\xi_{At}} - \boldsymbol{\xi_{Bt}} + \boldsymbol{\sigma_{\mu t}^{C_A}} - \boldsymbol{\sigma_{\mu t}^{C_B}}\right)^\top \left(\boldsymbol{\sigma_{\mu t}^{C_A}} - \boldsymbol{\sigma_{\mu t}^{C_B}}\right). \tag{13}$$

In particular, the difference between two countries' consumption volatilities loaded on the perceived fundamental shocks $\mathbf{dW}_{\boldsymbol{\mu}\mathbf{t}}^{\mathcal{E}}$ *can be written as*

$$\boldsymbol{\sigma_{\mu t}^{C_A}} - \boldsymbol{\sigma_{\mu t}^{C_B}} = -f(\boldsymbol{\eta_t})\left(\boldsymbol{\xi_{At}} - \boldsymbol{\xi_{Bt}}\right),$$

where $0 < f(\boldsymbol{\eta_t}) < 1$*. Consequently, the presence of* $\boldsymbol{\xi_{it}}$ *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 ($d\log\eta_t^B - d\log\eta_t^A < 0$). 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 $d\log S_t = d\log M^{\mathcal{E}}_{Bt} - d\log M^{\mathcal{E}}_{At}$, taking unconditional variance on both sides yields

$$\mathrm{Var}(d\log S_t) = \mathrm{Var}(d\log M^{\mathcal{E}}_{At}) + \mathrm{Var}(d\log M^{\mathcal{E}}_{Bt}) - 2\mathrm{Cov}(d\log M^{\mathcal{E}}_{At}, d\log M^{\mathcal{E}}_{Bt}).$$

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 growth since $d\log M^{\mathcal{E}}_{it} = -d\log C_{it}$. However, data shows that such correlation 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 growth.

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

$$d\log M^{\mathcal{E}}_{it} = \mathrm{drift}_M + d\log \eta^i_t - d\log C_{it}.$$

The correlation between pricing kernels crucially depends on the interaction between consumption and belief difference. The following proposition shows the covariance between growth 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*

$$Cov_t\left(d\log C_{it}, -d\log \eta^j_t\right) = \boldsymbol{\sigma}^{\boldsymbol{C_i}\,\top}_{\boldsymbol{\mu t}} \boldsymbol{\xi_{jt}} > 0. \tag{14}$$

Therefore, for a given level of pricing kernel volatility, a high correlation between SDFs is not necessarily because of highly correlated cross-country consumption growth. 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 $\text{Cov}_t\left(d\log\eta_t^A, d\log\eta_t^B\right) < 0$. 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[14]

$$rx_{t+dt} = d\log S_t + r_{Bt}dt - r_{At}dt. \tag{15}$$

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*

$$\mu_{t+dt}^{rx} \equiv \frac{1}{dt}\mathbb{E}_t^{\mathcal{E}}\left[rx_{t+dt}\right] = \frac{1}{2}(\boldsymbol{\sigma_{M_At}} + \boldsymbol{\sigma_{M_Bt}})^{\top}\Omega_W(\boldsymbol{\sigma_{M_At}} - \boldsymbol{\sigma_{M_Bt}}), \tag{16}$$

where $\boldsymbol{\sigma_{M_it}}$ is the market price of risk for country i and Ω_W is the correlation matrix of $\mathbf{dW}_t^{\mathcal{E}}$. Writing the expression more explicitly gives

$$\begin{aligned}\mu_{t+dt}^{rx} &= \frac{1}{2}\underbrace{\left({\sigma_A^{M_A}}^2 + {\sigma_B^{M_A}}^2 - {\sigma_A^{M_B}}^2 - {\sigma_B^{M_B}}^2\right)}_{\textit{(constant) output shocks}} + \frac{1}{2}\underbrace{\left(\boldsymbol{\sigma_{\mu t}^{M_A}} + \boldsymbol{\sigma_{\mu t}^{M_B}}\right)^{\top}\left(\boldsymbol{\sigma_{\mu t}^{M_A}} - \boldsymbol{\sigma_{\mu t}^{M_B}}\right)}_{\textit{(stochastic) perceived fundamental shocks}} \\ &= \textit{constant} + n\left(\boldsymbol{\eta_t}\right)\left(\eta_t^A - \eta_t^B\right),\end{aligned} \tag{17}$$

[14] The risk premium in levels can be found in Appendix H.

where $n(\boldsymbol{\eta_t}) > 0$ *is given in Appendix H.*[15]

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*

$$r_{At} - r_{Bt} = constant\ + (2\alpha - 1)(\mu_{At} - \mu_{Bt}) + \frac{1}{2} n(\boldsymbol{\eta_t}) \left(\eta_t^B - \eta_t^A\right), \tag{18}$$

where $n(\boldsymbol{\eta_t})$ *is the same as the one in* (17).

The expected return (17) consists of two components. The first is related to output risks as usual and is constant. It becomes zero when the two countries have the same output 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. The optimal consumption allocations depend crucially on the relative belief difference between the two countries, which is fluctuating over time. Thus, the expected return is time-varying, and heterogeneous beliefs constitute a source of risk in addition to the traditional output risks.

The time-varying expected log return in this model depends on the relative belief differences η_t^A and η_t^B for the two countries. To see the intuition, consider some shock in $\mathbf{dW}^{\mathcal{E}}_{\boldsymbol{\mu}\mathbf{t}}$ that drives η_t^A up. On one hand, by symmetry in (4), η_t^B decreases. As η_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 η_t^A and η_t^B depends on the current levels of belief differences. When $\eta^A > \eta_t^B$, the same shock makes η_t^A react more than η_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.

[15] The constant part is 0 when $\sigma_A = \sigma_B$.

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 $n(\boldsymbol{\eta_t})$ 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 with a general specification of belief processes, whether the setting is an endowment or production economy. 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 $\boldsymbol{\eta_t}$ 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 $\boldsymbol{\eta_t}$ 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 $\boldsymbol{\eta_t}$ 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 $\boldsymbol{\eta_t}$, 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 $\widetilde{\eta}_t$ 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

$$\text{LV19:}\qquad \mathbb{E}_t^{LV}\left[M_{it+1}\widetilde{\eta}_{t+1}\right] = \mathbb{E}_t^{LV}\left[M_{it+1}\right]\mathbb{E}_t^{LV}\left[\widetilde{\eta}_{t+1}\right] + \mathrm{Cov}_t^{LV}\left(M_{it+1}, \widetilde{\eta}_{t+1}\right) = \mathbb{E}_t^{LV}\left[M_{it+1}\right]. \tag{19}$$

Equation (19) holds since $\widetilde{\eta}_t$ 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[16]

$$\text{This paper:}\qquad M_{it+1}^{\mathcal{E}} \equiv \underbrace{M_{it+1}^{i}\eta_{t+1}^{i}}_{\text{SDF under measure }\mathcal{E}} = \underbrace{\widetilde{M}_{it+1}}_{\substack{\text{SDF under}\\ \text{homoge. beliefs}}} \underbrace{\frac{M_{it+1}^{i}}{\widetilde{M}_{it+1}}}_{\substack{\text{endogenous}\\ \text{indirect effect}}} \underbrace{\eta_{t+1}^{i}}_{\text{direct effect}}. \tag{20}$$

[16] I abuse the notations for simplified exposition, where e.g. M_{t+1} and η_{t+1}^i should really be M_{t+1}/M_t and η_{t+1}^i/η_t^i.

Equation (20) implies that there exists an equivalent martingale measure where M_i^i 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.[17]

An important implication of this paper is that the belief difference $\boldsymbol{\eta_t}$ 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 $\widetilde{M}_{it+1}$ is country i's pricing kernel in the case of homogeneous beliefs. By the assumption of CRRA utilities, the consumption shares are constant for each output, and $\widetilde{M}_{it+1}$ 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 $\mathbf{dW}^{\mathcal{E}}_{\boldsymbol{\mu}\mathbf{t}}$. 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 η^i_{t+1} is orthogonal to $\widetilde{M}_{it+1}$ by Assumption 1, we immediately arrive at the LV19 condition (19) as $\text{Cov}_t\left(M_{t+1}, \eta^i_{t+1}\right) = 0$.

(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

[17] 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 $\mathbb{E}^i_t\left[M^i_{it+1}\right] \neq \mathbb{E}^{\mathcal{E}}_t\left[M^i_{it+1}\eta^i_{t+1}\right]$. 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.

shares are endogenous to the belief difference $\boldsymbol{\eta_t}$ instead of being constant under homogeneous belief. The term $\frac{M^i_{it+1}}{\widetilde{M}_{it+1}}$ 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 differentials. Under LV19 setting, we have

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

In my paper, however, the covariance becomes

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

where $h(\boldsymbol{\eta_t}) > 0$ for all $\boldsymbol{\eta_t}$. 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 in an endowment economy[18] and then use this example for the calibration exercise with risk aversion parameter $\gamma > 1$ to complement the qualitative results in the preceding section. 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 each country is endowed with outputs X_{it} and the time-varying instant output growth rate μ_{it} in (1) follows

$$d\mu_{it} = a\left(\theta_{it} - \mu_{it}\right)dt + \sigma_{\mu_i} dZ_{\mu_i t}, \tag{21}$$

where a is the mean-reversion parameter, θ_{it} is the long-run growth rate, and σ_{μ_i} is constant volatility. The Brownian motions $dZ_{\mu_A t}$ and $dZ_{\mu_B t}$ are independent of each other and of output shocks Z_{At} and Z_{Bt}.

The time-varying long-run growth rate θ_{it} is given by the following OU process

$$d\theta_{it} = b\left(\bar{\theta}_i - \theta_{it}\right)dt + \sigma_{\theta_i} dZ_{\theta_i t}, \tag{22}$$

where b is the mean-reversion parameter, $\bar{\theta}_i$ is steady-state of θ_{it}, and σ_{θ_i} is constant volatility.[19] The Brownian motions $Z_{\theta_A t}$ and $Z_{\theta_B t}$ are independent of each other and of $Z_{At}, Z_{Bt}, Z_{\mu_A t}$ 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 θ_{it}, the shocks realizations $Z_{\theta_i t}$, or the volatilities σ_{θ_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.

[18] Appendix D offers two additional examples, one with unobserved long-run productivity growth and the other one with unobserved persistence level of business cycle.

[19] This setting for θ_{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 θ_t is independent of each other. One can think of μ_{it} as expected quarterly GDP growth and θ_{it} as long-run GDP growth projection.

For simplicity, let the true public signal processes be just pure noise and follow

$$ds_{it} = \sigma_s dZ_{s_i t},$$

where $Z_{s_A t}$ and $Z_{s_B t}$ are independent Brownian motions.

Assume each household i attaches correlations ϕ_A^i and ϕ_B^i 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

$$ds_{jt} = \sigma_s \phi_j^i dZ_{\theta_j t} + \sigma_s \sqrt{1 - (\phi_j^i)^2} dZ_{s_j t}. \tag{23}$$

In addition, each household i also misperceives the volatilities of θ_{At} and θ_{Bt} as $\sigma_{\theta_A}^i$ and $\sigma_{\theta_B}^i$.[20] To ease the exhibition and algebra, I assume the following values for perceived signal correlations and long-run growth rate volatilities

$$\phi_A^A = \phi_B^B = \phi \qquad\qquad \phi_B^A = \phi_A^B = -\phi \tag{24}$$

$$\sigma_{\theta_A}^A = \sigma_{\theta_A}^B = \frac{1}{\sqrt{1-\phi^2}} \sigma_{\theta_A} \qquad\qquad \sigma_{\theta_B}^A = \sigma_{\theta_B}^B = \frac{1}{\sqrt{1-\phi^2}} \sigma_{\theta_B}. \tag{25}$$

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 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

[20] In the usual scenarios where the random variable θ_{it} is observable, one can in principle learn about its volatility precisely. In the current context, θ_{it} cannot be directly observed and thus the inference about its volatility will not be perfect. Instead of letting households infer the volatilities of θ_{it}, I simply assume they stick to their priors, which gives more tractability to the model.

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 ($\mathcal{E}$) 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 $\widehat{\theta}_{it}^{\mathcal{E}}$ as the mean value of the benchmark agent's posterior distribution about θ_{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 θ_{it} follows

$$\begin{aligned} d\widehat{\theta}_{it}^{\mathcal{E}} &= b\left(\bar{\theta}_i - \widehat{\theta}_{it}^{\mathcal{E}}\right) dt + a\frac{\gamma_{\theta_i}}{\sigma_{\mu_i}^2}\left(d\mu_{it} - a\left(\widehat{\theta}_{it}^{\mathcal{E}} - \mu_{it}\right) dt\right) \\ &\equiv b\left(\bar{\theta}_i - \widehat{\theta}_{it}^{\mathcal{E}}\right) dt + a\frac{\gamma_{\theta_i}}{\sigma_{\mu_i}} dW_{\mu_i t}^{\mathcal{E}}, \end{aligned} \tag{26}$$

where $dW_{\mu_i t}^{\mathcal{E}}$ 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 θ_{it} in the long run will become precise if θ_{it} is deterministic ($\sigma_{\theta_i} = 0$). What prevents the estimate to converge to the true values is the shock $dZ_{\theta_i t}$ at each point in time.

Denote $\widehat{\theta}_{it}^{i}$ and $\widehat{\theta}_{jt}^{i}$ as the country i's estimates about θ_{it} and θ_{jt}, where $i \neq j$. We have

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

$$d\widehat{\theta}_{jt}^{i} = b\left(\bar{\theta}_j - \widehat{\theta}_{jt}^{i}\right) dt + a\frac{\gamma_{\theta_j}}{\sigma_{\mu_j}} dW_{\mu_j t}^{i} - \frac{\phi\sigma_{\theta_j}}{\sqrt{1-\phi^2}} dW_{s_j t}^{i}, \tag{28}$$

where $dW^i_{\mu_i t}$ and $dW^i_{\mu_j t}$ are defined similarly as $dW^{\mathcal{E}}_{\mu_i t}$ in (26). All shocks in (27) and (28) are Brownian motions under country i's probability measure and are mutually independent.

The quantity $\mathbf{dW^i_{\mu t}} \equiv \left(dW^i_{\mu_A t}, dW^i_{\mu_B t}\right)$ 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 $dW^i_{s_A t}$ and $dW^i_{s_B t}$ 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

$$ds_{it} = \sigma_s dW^{\mathcal{A}}_{s_i t} = \sigma_s dW^{\mathcal{B}}_{s_i t} = \sigma_s dW^{\mathcal{E}}_{s_i t}, \quad i \in \{A, B\}.$$

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 θ_{it} is equal to the benchmark agent's estimate. That is, at each point in time,*

$$\widehat{\theta}^i_{it} - \widehat{\theta}^{\mathcal{E}}_{it} = \widehat{\theta}^{\mathcal{E}}_{it} - \widehat{\theta}^j_{it},$$

where $i \neq j$. Moreover, the perceived fundamental shocks $\mathbf{dW^{\mathcal{E}}_{\mu t}}$ under the benchmark measure are equal to the average of the two countries

$$\mathbf{dW^{\mathcal{E}}_{\mu t}} = \frac{1}{2}\left(\mathbf{dW^{\mathcal{A}}_{\mu t}} + \mathbf{dW^{\mathcal{B}}_{\mu t}}\right).$$

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. It helps us to analyze an economy, in which the average belief is rational but individuals may have irrational expectations.

Define $g^i_{At} \equiv \widehat{\theta}^{\mathcal{E}}_{At} - \widehat{\theta}^i_{At}$ and $g^i_{Bt} \equiv \widehat{\theta}^{\mathcal{E}}_{Bt} - \widehat{\theta}^i_{Bt}$ 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

$$dg^i_{it} = -\left(b + a^2 \frac{\gamma_{\theta_i}}{\sigma^2_{\mu_i}}\right) g^i_{it} dt - \frac{\phi \sigma_{\theta_i}}{\sqrt{1-\phi^2}} dW^{\mathcal{E}}_{s_i t} \tag{29}$$

$$dg^i_{jt} = -\left(b + a^2 \frac{\gamma_{\theta_j}}{\sigma^2_{\mu_j}}\right) g^i_{jt} dt + \frac{\phi \sigma_{\theta_j}}{\sqrt{1-\phi^2}} dW^{\mathcal{E}}_{s_j t}, \quad i \neq j. \tag{30}$$

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 $W^{\mathcal{E}}_{s_i t}$ 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 $dW^{\mathcal{E}}_{s_A t} > 0$.

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

$$dW^{\mathcal{E}}_{\mu_i t} = dW^{\mathcal{A}}_{\mu_i t} - a \frac{g^A_{it}}{\sigma_{\mu_i}} dt = dW^{\mathcal{B}}_{\mu_i t} - a \frac{g^B_{it}}{\sigma_{\mu_i}} dt, \tag{31}$$

and we can express the expected output growth rate μ_{it} as

$$d\mu_{it} = a\left(\widehat{\theta}^{\mathcal{E}}_{it} - \mu_{it}\right) dt + \sigma_{\mu_i} dW^{\mathcal{E}}_{\mu_i t}. \tag{32}$$

Let $\boldsymbol{\xi}_{\boldsymbol{At}} = (\xi^A_{At}, \xi^A_{Bt}) \equiv \left(a\frac{g^A_{At}}{\sigma_{\mu_A}}, a\frac{g^A_{Bt}}{\sigma_{\mu_B}}\right)$ be country A's scaled differences in estimates. Lemma 1 immediately implies that these quantities for country B is simply $-\boldsymbol{\xi}_{\boldsymbol{At}}$. Then the change-of-measure variables $\boldsymbol{\eta_t}$ in (4) become

$$\frac{d\eta_t^A}{\eta_t^A} = -\xi_{At}^A dW_{\mu_A t}^{\mathcal{E}} - \xi_{Bt}^A dW_{\mu_B t}^{\mathcal{E}} \tag{33}$$

$$\frac{d\eta_t^B}{\eta_t^B} = \xi_{At}^A dW_{\mu_A t}^{\mathcal{E}} + \xi_{Bt}^A dW_{\mu_B t}^{\mathcal{E}}. \tag{34}$$

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.

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 γ is set to 5 in order to generate reasonable SDF volatilities, and the time preference β is 0.01. The home bias parameter is chosen to be $\alpha = 0.87$ to

Table 1: Calibrated parameters

Parameter	Symbol	Value
Output volatility	σ_i	0.03
Correlation of output	ρ	0.2
Correlation between signal and fundamental	ϕ	0.90
Home bias	α	0.87
Risk aversion	γ	5
Time discount	β	0.01
Pareto weight	λ_i	1

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.

match the average trade-to-GDP ratio. To show the effects of heterogeneous beliefs, I set the signal correlation ϕ to be 0.9. The Pareto weights λ_A and λ_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.

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 growth. In that case, the correlation coefficient between $\Delta \log S$ and $\Delta \log C_B -$

Table 2: Estimated parameters

Parameter	Symbol	Value
Volatility of expected growth rate	σ_{μ_i}	0.011 (0.006)
Volatility of long-run growth rate volatility	σ_{θ_i}	0.002 (0.000)
Steady-state long-run growth rate	$\bar{\theta}_i$	0.012 (0.003)
Mean reversion parameter of expected growth	a	0.223 (0.037)
Mean reversion parameter of long-run growth	b	0.121 (0.020)

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.

$\Delta \log C_A$ 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 growth, which is counterfactual as the average consumption growth correlation between the two countries is around 0.3 to 0.4. As we see in the homogeneous beliefs 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 growth, 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

Table 3: Simulation results — key moments

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

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.

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, 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

Table 4: Simulation results — other moments

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

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.

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 $\boldsymbol{\eta_t}$.

In addition, even if a process $\boldsymbol{\eta_t}$ 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 huge and very volatile. To address this concern, I calculate the average mean and volatility of the changes in the belief difference as well as the average mean of $|\widehat{\theta}^i_{it} - \widehat{\theta}^j_{it}|$ 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 movement in belief difference for each country has a moderate volatility of 5.7% in my paper.

Furthermore, the signal specifications and the implied beliefs should be thought plausible

Table 5: Belief difference moments

	Value
$d \log \eta_t^i$ mean	-0.007
$d \log \eta_t^i$ volatility	0.057
$\lvert\widehat{\theta}_{it}^i - \widehat{\theta}_{it}^j\rvert$ mean	0.006

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.

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 $\mathbf{dW}_{\boldsymbol{\mu}\mathbf{t}}^{\mathcal{A}}$ and $\mathbf{dW}_{\boldsymbol{\mu}\mathbf{t}}^{\mathcal{B}}$ and ask whether they can be distinguished from each other and from the rational benchmark agent's shocks $\mathbf{dW}_{\boldsymbol{\mu}\mathbf{t}}^{\mathcal{E}}$ 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.

Table 6: Probability of belief detection

Model	75 Years	100 Years
Benchmark agent $\mathcal{E}$	1.04%	2.30%
Country B	6.77%	9.13%

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.

6 Empirical Evidence

A key object in the model is the belief difference variable $\boldsymbol{\eta_t}$. In the previous sections, I have demonstrated $\boldsymbol{\eta_t}$ can qualitatively and quantitatively resolve the three puzzles. In this section, I first construct belief difference $\boldsymbol{\eta_t}$ from data and then provide various empirical tests for the model's main predictions relating to this object. Due to data limitations to be discussed below, the tests only involve six major currencies and cover periods from 2010Q1 to 2018Q4 at quarterly frequency. In the following empirical exercise, each country takes turns to serve as the home country instead of focusing on US as the home country/currency.

6.1 Data description

The main data set used to proxy for disagreements and to construct $\boldsymbol{\eta_t}$ is the BlackRock Macroeconomic Dashboard. This data set contains 12-month ahead real GDP growth forecasts from both Consensus Economics and BlackRock Nowcast developed by BlackRock Investment Institute and Systematic Active Equity. Consensus Economics polls hundreds of economists and practitioners around the world every month and collects their opinions about many macroeconomic variables, and the average forecasts are reported in this data set. The surveying processes are quite similar to the Survey of Professional Forecasters. The BlackRock Nowcast instead uses computer algorithms to predict GDP growth by exploiting the information from various macroeconomic indicators. This data series started in early 2010 and updates weekly in the early stage and now at daily frequency. In the following empirical analysis, I use the forecast data for Australia, Canada, EMU4,[21] Japan, UK, and US while the data set also includes China, France, Germany, Italy, and Spain.[22] Detailed descriptions of other data sets such as GDP and consumptions can be found in Appendix E.

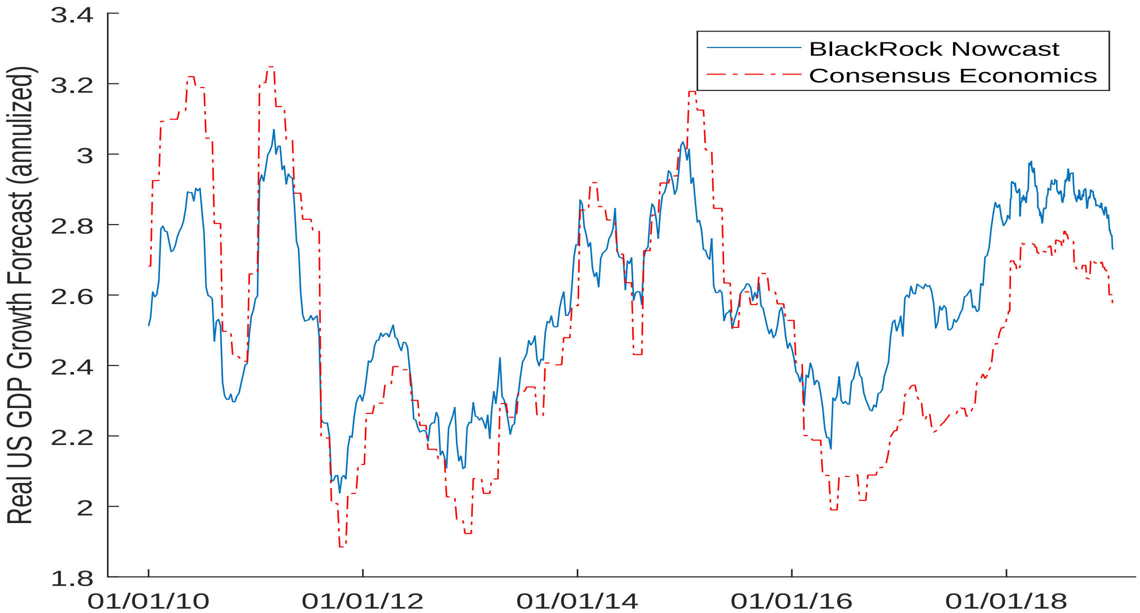
6.2 Construction of belief differece

To test the main predictions from the model, it is necessary to empirically construct the belief difference $\boldsymbol{\eta_t}$. Recall the definitions (33) and (34) that each country i's belief difference η_t^i depends on two components: (i) instant forecast differences with the econometrician about i's own country as well as about foreign countries and (ii) the econometrician's belief shocks.

[21] EMU4 consists of France, Germany, Italy, and Spain.

[22] The data for China is the forecast of Purchasing Managers' Index (PMI), not real GDP growth per se. It also started 4 years later than the rest of the countries. Since France, Germany, Italy, and Spain have the same currency, I treat them as a single entity and focus instead on EMU4.

Figure 1: Comparison between BlackRock Nowcast and Consensus Economics



This plot shows the time series of BlackRock Nowcast and average forecast in Consensus Economics about 12-month ahead real US GDP growth. The series starts on 01/01/2010 and ends on 12/31/2018. BlackRock Nowcast updates at weekly frequency in the beginning of the sample and daily in recent years. Consensus Economics reports data at monthly frequency throughout the sample.
Source: BlackRock Macro Dashboard.

In order to produce the first component, it is sufficient to map the agents' estimates in the model to data. The econometrician in the model is rational in interpreting information. Thus, the estimates $\hat{\theta}^{\mathcal{E}}$ are mapped to BlackRock Nowcast as the predictions are produced by computers, which eliminates irrational human behavior. The households in the model, on the other hand, subjectively choose how to interpret the information presented to them. This is consistent with the responses from the special surveys conducted by both Philadelphia Fed SPF and ECB SPF, in which more than 90% of respondents indicate that their forecasts are either purely experience-based or model-based with subjective adjustments.[23] Therefore, the estimates $\hat{\theta}^i_i$ by households about domestic countries are mapped to the forecast data in Consensus Economics. Since data for estimates about foreign countries is missing, I shut down this channel in the construction.[24] In the model, there is a persistent difference in estimates by the econometrician and households. Indeed, we can see, for example in Figure 1, that there are a lot of discrepancies between the two forecasts for US GDP growth,

[23]See the SPF survey results at this link.
[24]Results are qualitatively the same in an alternative construction where the foreign forecasts are backed out using Lemma 1.

reflecting different interpretations of information.

Table 7: Consumption growth differential regression

	(1) $d\log C_{Ht} - d\log C_{it}$	(2) $d\log C_{Ht} - d\log C_{it}$	(3) $d\log C_{Ht} - d\log C_{it}$
$d\log\eta_t^H - d\log\eta_t^i$	0.040**	0.040**	0.044***
	(0.017)	(0.017)	(0.017)
$d\log X_{Ht} - d\log X_{it}$			0.380***
			(0.064)
Currency pair FE	No	Yes	Yes
N	1080	1080	1080
adj. R^2	0.004	0.004	0.047

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

This table reports regression coefficients in (35). The panel includes Australia, Canada, EMU4 (France, Germany, Italy, and Spain), Japan, UK, and US from 2010Q1 to 2018Q4 at quarterly frequency. Each country takes turns to serve as the home country so that there are 30 home-foreign pairs. Real consumption and GDP are seasonally adjusted. Belief difference variables are constructed following methods outlined in Section 6.2. The regression includes currency pair fixed effects to mitigate concerns about time-invariant unobservables. Robust standard errors with two-way clustering by home country × time and foreign country × time are displayed in parentheses.

The second component, econometrician's belief shocks, is simply backed out using (32), where the model parameters are taken from Section 5.3 and the expected growth rates μ_{it} are proxied by quarter-on-quarter real GDP growth rates. Finally, the belief difference η_t^i for country i is obtained via numerical Ito integration.[25]

6.3 Results

Let's first consider the relationship between consumption growth differential and changes in belief difference. Recall that Proposition 1 states there is a positive correlation between these two objects. In what follows, I designate country A as home country (H) and B as

[25]The explicit formula of Ito integral is $\eta_t^i = \exp\left[-\sum_j \left(\frac{1}{2}\int_0^t (\xi_{jt}^i)^2 dt + \int_0^t \xi_{jt}^i dW_{\mu_j t}^{\mathcal{E}}\right)\right]$, where $\xi_{jt}^i = ag_{jt}^i/\sigma_{\mu_j}$ and the summation is over all countries.

foreign. To test this relationship, I run the following panel regression:

$$\Delta(\log C_{Ht} - \log C_{it}) = \alpha_{iH} + \beta_1 \Delta(\log \eta_t^H - \log \eta_t^i) + \beta_2 \Delta(\log X_{Ht} - \log X_{it}) + \varepsilon_{iH,t+1}, \quad (35)$$

where i refers to foreign countries and α_{iH} denotes currency-pair fixed effects. I also control for GDP growth differentials. Table 7 reports the regression results. The main coefficient of interest is the one associated with differences of changes in belief difference, i.e. β_1. We see

Table 8: FX growth regression

	(1) $d\log FX_t$	(2) $d\log FX_t$	(3) $d\log FX_t$	(4) $d\log FX_t$	(5) $d\log FX_t$
$d\log \eta_t^H - d\log \eta_t^i$	-0.217*	-0.244**	-0.232**	-0.242**	-0.310***
	(0.111)	(0.113)	(0.114)	(0.112)	(0.118)
$d\log C_{Ht} - d\log C_{it}$			-0.315*	-0.237	-0.147
			(0.173)	(0.177)	(0.174)
$d\log X_{Ht} - d\log X_{it}$				-0.597*	-0.454
				(0.313)	(0.322)
$\sigma_{X_{Ht}} - \sigma_{X_{it}}$					0.172
					(1.545)
Currency pair FE	No	Yes	Yes	Yes	Yes
N	1080	1080	1080	1080	1050
adj. R^2	0.004	0.005	0.008	0.011	0.011

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

This table reports regression coefficients in (36). The panel includes Australia, Canada, EMU4 (France, Germany, Italy, and Spain), Japan, UK, and US from 2010Q1 to 2018Q4 at quarterly frequency. Each country takes turns to serve as the home country so that there are 30 home-foreign pairs. Real consumption and GDP are seasonally adjusted. The annualized real GDP volatility is constructed using trailing moving standard deviation with a window size of 4 (one year). When there are fewer elements than the window size in the beginning of the sample, I take the standard deviation over the elements that are available. Belief difference variables are constructed following methods outlined in Section 6.2. The regression includes currency pair fixed effects to mitigate concerns about time-invariant unobservables. Robust standard errors with two-way clustering by home country × time and foreign country × time are displayed in parentheses.

that β_1 is always positive and significant whether or not fixed effects or control variables are included. This result suggests that countries' consumption growth differentials respond to the relative changes in belief difference. We can also see in Proposition 1 that consumption growth differential is increasing in differences in endowment growth. The coefficient β_2 on GDP growth differential confirms this prediction.

The second result relates the movements in real exchange rates to belief difference. Proposition 4 implies that the negative correlation between $d \log C_{At} - d \log C_{Bt}$ and $d \log \eta_t^B - d \log \eta_t^A$ helps mitigate the Backus-Smith puzzle.[26] In other words, after controlling for consumption growth difference, exchange rate movement should be *negatively* correlated with $d \log \eta_t^A - d \log \eta_t^B$. Table 8 reports the results of the following regression:

$$\Delta \log FX_{iH,t} = \alpha_{iH} + \beta_1 \Delta(\log \eta_t^H - \log \eta_t^i) + \boldsymbol{\beta_2^\top Control_{iH,t}} + \varepsilon_{iH,t+1}, \tag{36}$$

where $FX_{iH,t}$ denotes the bilateral real exchange rate between country i and home country H, and the control variables include differences in consumption growth, GDP growth and GDP volatility. We can see that the coefficient β_1 on belief difference is always negative and significant whether or not we include controls. This strong relationship suggests that belief difference is indeed a key determinant of exchange rate movement, and the disconnection between FX and macro fundamentals is at least partially due to investors' (or households') different beliefs regarding the underlying economies.

The last set of results pertains to interest rates and currency risk premium. As shown in Proposition 7, the interest rate differential negatively correlates with gap in belief difference after controlling for GDP growth differential. At the same time, (17) and (18) imply that expected currency risk premium positively correlates gap in with belief *and* the effect of belief difference is stronger for risk premium than for interest rate differential, which helps to achieve positive carry trade returns on average. In order to test these predictions I run the following two regressions:

$$r_{Ht} - r_{it} = \alpha_{iH}^r + \beta_1^r(\eta_t^H - \eta_t^i) + \boldsymbol{\beta_2^{r\top} Control_{iH,t}^r} + \varepsilon_{iH,t+1}^r \tag{37}$$

$$rx_{iH,t+1} = \alpha_{iH}^{rx} + \beta_1^{rx}(\eta_t^H - \eta_t^i) + \beta_2^{rx}(\sigma_{X_H} - \sigma_{X_i}) + \varepsilon_{iH,t+1}^{rx}, \tag{38}$$

where the log risk premium from home country's perspective is $rx_{iH,t+1} \equiv \Delta \log FX_{iH,t} + r_{it} - r_{Ht}$ and control variable in the first regression include GDP growth and volatility differences. Results in Table 9 confirm these predictions about both direction and magnitude, suggesting that interest rate differentials and carry trade returns are closely linked to belief difference

[26] We have already seen this negative relationship between $d \log C_{At} - d \log C_{Bt}$ and $d \log \eta_t^B - d \log \eta_t^A$ in Table 7.

among countries.

Table 9: Interest rate difference and risk premium regression

	(1)	(2)	(3)	(4)	(5)
	$r_{Ht} - r_{it}$	$r_{Ht} - r_{it}$	$r_{Ht} - r_{it}$	rx_{t+1}	rx_{t+1}
$\eta_t^H - \eta_t^i$	-3.090***	-1.770***	-2.142**	8.658**	9.384***
	(0.800)	(0.761)	(0.839)	(3.081)	(3.079)
$d\log X_{Ht} - d\log X_{it}$		0.233***	0.253***		
		(0.045)	(0.046)		
$\sigma_{X_{Ht}} - \sigma_{X_{it}}$			-0.648***		-1.075
			(0.220)		(1.147)
Currency pair FE	Yes	Yes	Yes	Yes	Yes
N	1080	1080	1050	1080	1050
adj. R^2	0.040	0.051	0.068	0.009	0.010

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

This table reports regression coefficients in (37) for columns (1)-(3) and in (38) for (4)-(5). The panel includes Australia, Canada, EMU4 (France, Germany, Italy, and Spain), Japan, UK, and US from 2010Q1 to 2018Q4 at quarterly frequency. Each country takes turns to serve as the home country so that there are 30 home-foreign pairs. Real interest rates are constructed by subtracting CPI-implied inflation from 3-month treasury bills of each country. Log currency risk premium from home country's perspective is one-period ahead real FX growth and the difference between foreign and home real interest rates. Real GDP data are seasonally adjusted. The real GDP volatility is constructed using trailing moving standard deviation with a window size of 4 (one year). When there are fewer elements than the window size in the beginning of the sample, I take the standard deviation over the elements that are available. Belief difference variables are constructed following methods outlined in Section 6.2. All variables are annualized. The regression includes currency pair fixed effects to mitigate concerns about time-invariant unobservables. Robust standard errors with two-way clustering by home country $\times$ time and foreign country $\times$ time are displayed in parentheses.

7 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 an endowment economy. 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 growth, a modest level of exchange rate volatility as well as a sizable carry trade return. Finally, this paper provides empirical evidence to support the main predictions of the model.

References

Andrei, D., Hasler, M. and Jeanneret, A. (2018), 'Asset pricing with persistence risk', *The Review of Financial Studies* **32**(7), 2809–2849.

Bacchetta, P. and Van Wincoop, E. (2006), 'Can information heterogeneity explain the exchange rate determination puzzle?', *American Economic Review* **96**(3), 552–576.

Backus, D. K. and Smith, G. W. (1993), 'Consumption and real exchange rates in dynamic economies with non-traded goods', *Journal of International Economics* **35**(3-4), 297–316.

Bakshi, G., Cerrato, M. and Crosby, J. (2017), 'Implications of incomplete markets for international economies', *The Review of Financial Studies* **31**(10), 4017–4062.

Bansal, R. and Shaliastovich, I. (2012), 'A long-run risks explanation of predictability puzzles in bond and currency markets', *The Review of Financial Studies* **26**(1), 1–33.

Basak, S. (2005), 'Asset pricing with heterogeneous beliefs', *Journal of Banking & Finance* **29**(11), 2849–2881.

Brandt, M. W., Cochrane, J. H. and Santa-Clara, P. (2006), 'International risk sharing is better than you think, or exchange rates are too smooth', *Journal of Monetary Economics* **53**(4), 671–698.

Burnside, A. C. and Graveline, J. J. (2019), 'On the asset market view of exchange rates', *The Review of Financial Studies* .

Burnside, C., Eichenbaum, M., Kleshchelski, I. and Rebelo, S. (2010), 'Do peso problems explain the returns to the carry trade?', *The Review of Financial Studies* **24**(3), 853–891.

Burnside, C., Han, B., Hirshleifer, D. and Wang, T. Y. (2011), 'Investor overconfidence and the forward premium puzzle', *The Review of Economic Studies* **78**(2), 523–558.

Colacito, R. and Croce, M. M. (2011), 'Risks for the long run and the real exchange rate', *Journal of Political economy* **119**(1), 153–181.

Dumas, B., Kurshev, A. and Uppal, R. (2009), 'Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility', *The Journal of Finance* **64**(2), 579–629.

Dumas, B., Lewis, K. K. and Osambela, E. (2017), 'Differences of opinion and international equity markets', *The Review of Financial Studies* **30**(3), 750–800.

Eichenbaum, M., Johannsen, B. K. and Rebelo, S. (2017), 'Monetary policy and the predictability of nominal exchange rates'.

Fama, E. F. (1984), 'Forward and spot exchange rates', *Journal of monetary economics* **14**(3), 319–338.

Farhi, E. and Gabaix, X. (2016), 'Rare disasters and exchange rates', *The Quarterly Journal of Economics* **131**(1), 1–52.

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

Heyerdahl-Larsen, C. (2014), 'Asset prices and real exchange rates with deep habits', *The Review of Financial Studies* **27**(11), 3280–3317.

Kollmann, R. M. W. (1991), Essays on international business cycles, PhD thesis, University of Chicago, Department of Economics.

Li, T. and Muzere, M. L. (2010), 'Heterogeneity and volatility puzzles in international finance', *Journal of Financial and Quantitative Analysis* **45**(6), 1485–1516.

Liptser, R. S. and Shiryaev, A. N. (2000), *Statistics of random processes: II. Application*, Vol. 5, Springer Science & Business Media.

Lucas, R. E. (1978), 'Asset prices in an exchange economy', *Econometrica: Journal of the Econometric Society* pp. 1429–1445.

Lustig, H. and Verdelhan, A. (2019), 'Does incomplete spanning in international financial markets help to explain exchange rates?', *American Economic Review* **109**(6), 2208–44.

Maurer, T. A. and Tran, N.-K. (2016), 'Entangled risks in incomplete fx markets'.

Maurer, T. A. and Tran, N.-K. (2018), 'Incomplete asset market view of the exchange rate determination'.

Meese, R. A. and Rogoff, K. (1983), 'Do they fit out of sample?', *Journal of international economics* **14**, 3–24.

Mueller, P., Tahbaz-Salehi, A. and Vedolin, A. (2017), 'Exchange rates and monetary policy uncertainty', *The Journal of Finance* **72**(3), 1213–1252.

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

Øksendal, B. (2003), *Stochastic differential equations*, Springer.

Pavlova, A. and Rigobon, R. (2007), 'Asset prices and exchange rates', *The Review of Financial Studies* **20**(4), 1139–1180.

Sandulescu, M., Trojani, F. and Vedolin, A. (2019), 'Model-free international stochastic discount factors', *The Journal of Finance* (Forthcoming).

Scheinkman, J. A. and Xiong, W. (2003), 'Overconfidence and speculative bubbles', *Journal of political Economy* **111**(6), 1183–1220.

Stathopoulos, A. (2016), 'Asset prices and risk sharing in open economies', *The Review of Financial Studies* **30**(2), 363–415.

Verdelhan, A. (2010), 'A habit-based explanation of the exchange rate risk premium', *The Journal of Finance* **65**(1), 123–146.

Xiong, W. and Yan, H. (2009), 'Heterogeneous expectations and bond markets', *The Review of Financial Studies* **23**(4), 1433–1466.

Yu, J. (2013), 'A sentiment-based explanation of the forward premium puzzle', *Journal of Monetary Economics* **60**(4), 474–491.

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 much 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

$$\max_{\{C^A_{At}, C^A_{Bt}, C^B_{At}, C^B_{Bt}\}} \mathbb{E}^{\mathcal{E}}_0 \left[\int_0^\infty e^{-\beta t} \left(\lambda_A \eta^A_t U(C_{At}) + \lambda_B \eta^B_t U(C_{Bt}) \right) dt \right], \tag{39}$$

where

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

subject to the resource constraint on the tradable good

$$\begin{aligned} C^A_{At} + C^A_{Bt} &= X_{At} \\ C^B_{At} + C^B_{Bt} &= X_{Bt}. \end{aligned}$$

A.2 Equilibrium allocations

Denoting ν_{it} as the Lagrange multiplier on the resource constraint, we obtain the first order conditions for country A

$$\begin{aligned} \lambda_A \eta^A_t \alpha C^{1-\gamma}_{At} \frac{1}{C^A_{At}} &= \nu_{At} \\ \lambda_A \eta^A_t (1-\alpha) C^{1-\gamma}_{At} \frac{1}{C^A_{Bt}} &= \nu_{Bt}, \end{aligned}$$

and for country B

$$\begin{aligned} \lambda_B \eta^B_t (1-\alpha) C^{1-\gamma}_{Bt} \frac{1}{C^B_{At}} &= \nu_{At} \\ \lambda_B \eta^B_t \alpha C^{1-\gamma}_{Bt} \frac{1}{C^B_{Bt}} &= \nu_{Bt}. \end{aligned}$$

We immediately see that the two countries equalize the marginal utilities on the same good due to frictionless trade, which implies

$$\frac{\lambda_A \alpha C_{At}^{1-\gamma} C_{At}^{B}}{\lambda_B (1-\alpha) C_{Bt}^{1-\gamma} C_{At}^{A}} \frac{\eta_t^A}{\eta_t^B} = 1$$
$$\frac{\lambda_A (1-\alpha) C_{At}^{1-\gamma} C_{Bt}^{B}}{\lambda_B \alpha C_{Bt}^{1-\gamma} C_{Bt}^{A}} \frac{\eta_t^A}{\eta_t^B} = 1.$$

Re-arranging the two equations above, we obtain the following

$$k_t \equiv \frac{\alpha C_{Bt}^A}{(1-\alpha) C_{At}^A} = \frac{(1-\alpha) C_{Bt}^B}{\alpha C_{At}^B}.$$

Define $\omega_t \equiv \frac{C_{At}^A}{C_{At}^B}$ as the (equilibrium) ratio of two countries' consumption of good A. In other words, the equilibrium consumption allocations for good A are

$$C_{At}^A = \frac{\omega_t}{1+\omega_t} X_{At}$$
$$C_{At}^B = \frac{1}{1+\omega_t} X_{At}$$

The first order conditions then imply

$$\begin{aligned} 1 &= \frac{\lambda_A}{\lambda_B} \frac{\alpha}{1-\alpha} \frac{\left(\frac{1-\alpha}{\alpha}\right)^{(1-\alpha)(1-\gamma)} (C_{At}^A)^{(1-\gamma)} k_t^{(1-\alpha)(1-\gamma)}}{\left(\frac{\alpha}{1-\alpha}\right)^{\alpha(1-\gamma)} (C_{At}^B)^{(1-\gamma)} k_t^{\alpha(1-\gamma)}} \frac{C_{At}^B}{C_{At}^A} \frac{\eta_t^A}{\eta_t^B} \\ &= \frac{\lambda_A}{\lambda_B} \left(\frac{1-\alpha}{\alpha}\right)^{-\gamma} \omega_t^{-\gamma} k_t^{(1-2\alpha)(1-\gamma)} \frac{\eta_t^A}{\eta_t^B}. \end{aligned} \tag{40}$$

The market clearing conditions for good B gives

$$\left(\frac{\alpha}{1-\alpha} C_{At}^B + \frac{1-\alpha}{\alpha} C_{At}^A\right) k_t = X_{Bt}.$$

Substituting C_{At}^A and C_{At}^B into the above equation yields

$$k_t = \frac{X_{Bt}}{X_{At}} \left[\frac{\alpha(1-\alpha)(1+\omega_t)}{\alpha^2 + (1-\alpha)^2 \omega_t}\right]. \tag{41}$$

I can numerically solve for ω_t using (40) and (41).

Proposition 8. *The consumption basket for country A is*

$$\begin{aligned} C_{At} &= \left(\frac{\omega_t}{1+\omega_t} X_{At}\right)^{\alpha} \left(\frac{1-\alpha}{\alpha} C_{At}^A k_t\right)^{1-\alpha} \\ &= \left(\frac{\omega_t}{1+\omega_t} X_{At}\right)^{\alpha} \left(\frac{(1-\alpha)^2\omega_t}{\alpha^2+(1-\alpha)^2\omega_t} X_{Bt}\right)^{1-\alpha}, \end{aligned}$$

and for country B is

$$\begin{aligned} C_{Bt} &= \left(\frac{\alpha}{1-\alpha} C_{Bt}^A k_t\right)^{\alpha} \left(\frac{1}{1+\omega_t} X_{At}\right)^{1-\alpha} \\ &= \left(\frac{\alpha^2}{\alpha^2+(1-\alpha)^2\omega_t} X_{Bt}\right)^{\alpha} \left(\frac{1}{1+\omega_t} X_{At}\right)^{1-\alpha}. \end{aligned}$$

A.3 Dynamics of good A consumption ratio ω_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 ω_t.

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

$$d\omega_t = \mu_{\omega t} dt + \sigma_{At}^{\omega} dZ_{At} + \sigma_{Bt}^{\omega} dZ_{Bt} + \boldsymbol{\sigma_{\mu t}^{\omega}}^{\top} \mathbf{dW}_{\boldsymbol{\mu}\mathbf{t}}^{\mathcal{E}},$$

where drift and volatilities are given in Appendix H.

A.4 Consumption dynamics

Proposition 10. *The process for the equilibrium consumption follows*

$$\frac{dC_{it}}{C_{it}} = \mu_{C_i t} dt + \sigma_A^{C_i} dZ_{At} + \sigma_B^{C_i} dZ_{Bt} + \boldsymbol{\sigma_{\mu t}^{C_i}}^{\top} \mathbf{dW}_{\boldsymbol{\mu}\mathbf{t}}^{\mathcal{E}}, \quad i = A, B,$$

where drift and volatilities are given in Appendix H.

A.5 Stochastic discount factors

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

$$\begin{aligned} M_{At}^{\mathcal{E}} &= e^{-\beta t} C_{At}^{-\gamma} \eta_t^A \\ M_{Bt}^{\mathcal{E}} &= e^{-\beta t} C_{Bt}^{-\gamma} \eta_t^B. \end{aligned}$$

Proposition 11. *The process for the stochastic discount factor follows*

$$\frac{dM_{it}^{\mathcal{E}}}{M_{it}^{\mathcal{E}}} = \mu_{M_i t} dt - \sigma_A^{M_i} dZ_{At} - \sigma_B^{M_i} dZ_{Bt} - \boldsymbol{\sigma_{\mu t}^{M_i}}^{\top} \mathbf{dW}_{\boldsymbol{\mu}\mathbf{t}}^{\mathcal{E}}, \quad i = A, B,$$

where drift and volatilities are given in Appendix H.

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

$$\begin{aligned} r_{it} &= \beta + \gamma\mu_{C_it} + \frac{1}{2}\boldsymbol{\xi}_{it}^\top\boldsymbol{\xi}_{it} - \frac{1}{2}\gamma\left({\sigma_A^{C_i}}^2 + {\sigma_B^{C_i}}^2 + 2\rho\sigma_A^{C_i}\sigma_B^{C_i} + {\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{C_i}}}^\top\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{C_i}}\right) \\ &\quad - \frac{1}{2}\left({\sigma_A^{M_i}}^2 + {\sigma_B^{M_i}}^2 + 2\rho\sigma_A^{M_i}\sigma_B^{M_i} + {\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{M_i}}}^\top\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{M_i}}\right). \end{aligned}$$

A.7 Exchange rate

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

$$\begin{aligned} S_t &= \frac{\eta_t^B}{\eta_t^A}\frac{M_{Bt}^{\mathcal{B}}}{M_{At}^{\mathcal{A}}} \\ &= \frac{\eta_t^B}{\eta_t^A}\left(\frac{C_{At}}{C_{Bt}}\right)^\gamma. \end{aligned}$$

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

$$\frac{dS_t}{S_t} = \mu_{St}dt + \sigma_A^S dZ_{At} + \sigma_B^S dZ_{Bt} + {\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{S}}}^\top \mathbf{dW}_{\boldsymbol{\mu}\mathbf{t}}^{\mathcal{E}},$$

where drift and volatilities are given in Appendix H.

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 output 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:

$$P_{At} \equiv \min_{\{C^A_{At}, C^A_{Bt}\}} q_{At} C^A_{At} + q_{Bt} C^A_{Bt} \quad \text{s.t. } C_{At} \equiv \left(C^A_{At}\right)^{\alpha} \left(C^A_{Bt}\right)^{1-\alpha} = 1.$$

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

$$\begin{aligned} P_{At} &= \left(\frac{q_{At}}{\alpha}\right)^{\alpha} \left(\frac{q_{Bt}}{1-\alpha}\right)^{1-\alpha} \\ P_{Bt} &= \left(\frac{q_{Bt}}{\alpha}\right)^{\alpha} \left(\frac{q_{At}}{1-\alpha}\right)^{1-\alpha}. \end{aligned}$$

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

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

subject to the budget constraint

$$\mathbb{E}^{\mathcal{E}}_0 \left[\int_0^{\infty} M^{\mathcal{E}}_{At} \left(q_{At} C^A_{At} + q_{Bt} C^A_{Bt}\right) dt\right] = \mathbb{E}^{\mathcal{E}}_0 \left[\int_0^{\infty} M^{\mathcal{E}}_{At} q_{At} X_{At} dt\right],$$

where $M^{\mathcal{E}}_{At}$ denotes country A's pricing kernel under measure $\mathcal{E}$. Notice that country A's optimization problem includes a change-of-measure variable η_{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

$$\begin{aligned} \eta_{At} e^{-\beta t} \alpha \frac{1}{C^A_{At}} &= \nu_A M^{\mathcal{E}}_{At} q_{At} \\ \eta_{At} e^{-\beta t} (1-\alpha) \frac{1}{C^A_{Bt}} &= \nu_A M^{\mathcal{E}}_{At} q_{Bt}, \end{aligned}$$

where ν_A is the Lagrange multiplier on A's budget constraint. Country B's problem is defined similarly, and the FOCs are

$$\begin{aligned} \eta_{Bt} e^{-\beta t} (1-\alpha) \frac{1}{C^B_{At}} &= \nu_B M^{\mathcal{E}}_{At} q_{At} \\ \eta_{Bt} e^{-\beta t} \alpha \frac{1}{C^B_{Bt}} &= \nu_B M^{\mathcal{E}}_{At} q_{Bt}. \end{aligned}$$

We immediately obtain the optimal consumption allocations

$$\begin{aligned} C_{At}^A &= \frac{\eta_{At} e^{-\beta t} \alpha}{\nu_A M_{At}^{\mathcal{E}} q_{At}} \\ C_{Bt}^A &= \frac{\eta_{At} e^{-\beta t} (1-\alpha)}{\nu_A M_{At}^{\mathcal{E}} q_{Bt}} \\ C_{At}^B &= \frac{\eta_{Bt} e^{-\beta t} (1-\alpha)}{\nu_B M_{At}^{\mathcal{E}} q_{At}} \\ C_{Bt}^B &= \frac{\eta_{Bt} e^{-\beta t} \alpha}{\nu_B M_{At}^{\mathcal{E}} q_{Bt}}. \end{aligned}$$

Given the resource constraint that $C_{At}^A + C_{At}^B = X_{At}$ and $C_{Bt}^A + C_{Bt}^B = X_{Bt}$, we obtain $M_{At}^{\mathcal{E}} q_{At}$ and $M_{At}^{\mathcal{E}} q_{Bt}$ as follows.

$$M_{At}^{\mathcal{E}} q_{At} = \frac{e^{-\beta t}}{X_{At}} \left[\frac{\eta_{At} \alpha}{\nu_A} + \frac{\eta_{Bt}(1-\alpha)}{\nu_B} \right] \tag{42}$$

$$M_{At}^{\mathcal{E}} q_{Bt} = \frac{e^{-\beta t}}{X_{Bt}} \left[\frac{\eta_{At}(1-\alpha)}{\nu_A} + \frac{\eta_{Bt} \alpha}{\nu_B} \right]. \tag{43}$$

The ratio k_t of prices for the output goods is therefore

$$k_t \equiv \frac{q_{At}}{q_{Bt}} = \frac{X_{Bt}}{X_{At}} \left[\frac{\eta_{At} \alpha / \nu_A + \eta_{Bt}(1-\alpha)/\nu_B}{\eta_{At}(1-\alpha)/\nu_A + \eta_{Bt} \alpha / \nu_B} \right]. \tag{44}$$

Since we normalize P_{At} to be one, we get the prices q_{At} and q_{Bt} as[27]

$$q_{At} = \alpha^\alpha (1-\alpha)^{1-\alpha} k_t^{1-\alpha} \tag{45}$$

$$q_{Bt} = \alpha^\alpha (1-\alpha)^{1-\alpha} k_t^{-\alpha} \tag{46}$$

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

$$\begin{aligned} C_{At} &= \left(C_{At}^A\right)^\alpha \left(C_{Bt}^A\right)^{1-\alpha} \\ &= \frac{e^{-\beta t}}{\nu_A M_{At}^{\mathcal{E}}} \eta_{At} \underbrace{\left(\frac{\alpha}{q_{At}}\right)^\alpha \left(\frac{1-\alpha}{q_{Bt}}\right)^{1-\alpha}}_{=1} \\ &= \frac{e^{-\beta t}}{\nu_A M_{At}^{\mathcal{E}}} \eta_{At}. \end{aligned}$$

[27]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

$$M^{\mathcal{E}}_{At} = e^{-\beta t}\frac{\eta_{At}}{\nu_A C_{At}}.$$

Note that the Lagrange multiplier is constant and does not affect pricing.[28] The expression for $M^{\mathcal{E}}_{Bt}$ can be obtained in a similar way.

In order to derive explicit expressions for $M^{\mathcal{E}}_{At}$ and C_{At}, I combine the equations (42) and (45). 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 γ, we still have $\left(\frac{\alpha}{q_{At}}\right)^{\alpha}\left(\frac{1-\alpha}{q_{Bt}}\right)^{1-\alpha} = 1$. Thus, we get

$$M^{\mathcal{E}}_{At} = \frac{e^{-\beta t}}{\nu_A}\eta_{At}C_{At}^{-\gamma}$$
$$M^{\mathcal{E}}_{Bt} = \frac{e^{-\beta t}}{\nu_B}\eta_{Bt}C_{Bt}^{-\gamma}.$$

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 movement, 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.

[28] The price of an asset payoff Y_T at time $t < T$ is $\mathbb{E}^{\mathcal{E}}_t\left[\frac{M^{\mathcal{E}}_{AT}}{M^{\mathcal{E}}_{At}}Y_T\right]$. 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 Additional Examples of Belief Specification

In this section, I provide two additional examples of belief specification and learning. In the first example, the unobserved variables are the long-run productivity growth in a production economy. In the second example, the unobservables are the persistence levels of business cycles in an endowment economy.

D.1 Long-run productivity growth

The productivity of country i is denoted as A_{it}. Country i takes labor L_{it} as input and produces output X_{it} with the production function $X_{it} = A_{it}L_{it}^{\psi}$, where $0 < \psi < 1$. For simplicity, I assume the labor is exogenous and follows

$$\frac{dL_{it}}{L_{it}} = \mu_{L_i}dt + \sigma_{L_i}dZ_{L_i t}.$$

The productivity follows

$$\frac{dA_{it}}{A_{it}} = \mu_{A_i t}dt + \sigma_{A_i}dZ_{A_i t},$$

where the growth rate $\mu_{A_i t}$ is time-varying and given by

$$d\mu_{it} = a\left(\theta_{it} - \mu_{it}\right)dt + \sigma_{\mu_i}dZ_{\mu_i t}$$

and the long-run growth rate θ_{it} is unobserved and follows

$$d\theta_{it} = b\left(\bar{\theta}_i - \theta_{it}\right)dt + \sigma_{\theta_i}dZ_{\theta_i t}.$$

All of the Brownian motions above are independent of each other. The dynamics of the output therefore is

$$\frac{dX_{it}}{X_{it}} = \frac{\psi}{L_{it}}dL_{it} + \frac{1}{A_{it}}dA_{it}.$$

Similar to Section 5, households 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

$$ds_{it} = \sigma_s dZ_{s_i t},$$

where $Z_{s_A t}$ and $Z_{s_B t}$ are independent Brownian motions. Assume each household i attaches correlations ϕ_A^i and ϕ_B^i 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

$$ds_{jt} = \sigma_s\phi_j^i dZ_{\theta_j t} + \sigma_s\sqrt{1-(\phi_j^i)^2}dZ_{s_j t}.$$

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

$$\phi^A_A = \phi^B_B = \phi \qquad\qquad \phi^A_B = \phi^B_A = -\phi$$
$$\sigma^A_{\theta_A} = \sigma^B_{\theta_A} = \frac{1}{\sqrt{1-\phi^2}}\sigma_{\theta_A} \qquad\qquad \sigma^A_{\theta_B} = \sigma^B_{\theta_B} = \frac{1}{\sqrt{1-\phi^2}}\sigma_{\theta_B}.$$

We then obtain similar results in Section 5.1.2.

D.2 Persistence level of business cycle

In this example, we are back to an endowment economy. Instead of having a time-varying long-run growth rate, the instantaneous expected growth rate now chases a constant target. The speed of convergence (or the mean reversion parameter), however, is time-varying and unobserved as in Andrei, Hasler and Jeanneret (2018). Specifically, the expected growth for country i's output follows

$$d\mu_{it} = a_{it}\left(\bar{\mu}_i - \mu_{it}\right)dt + \sigma_{\mu_i} dZ_{\mu_i t}$$

The time-varying meas-reversion parameter a_{it} equals $\bar{a} + \theta_{it}$, where θ_{it} follows

$$d\theta_{it} = -b\theta_{it}dt + \sigma_{\theta_i} dZ_{\theta_i t}.$$

Households observe μ_{it} but not θ_{it}. They receive public signals s_{At} and s_{Bt} regarding the shocks $dZ_{\theta_A t}$ and $dZ_{\theta_B t}$. Using similar signal structures and interpretations in Section 5, the benchmark estimate is

$$d\widehat{\theta}^{\mathcal{E}}_{it} = -b\widehat{\theta}^{\mathcal{E}}_{it}dt + \left(\bar{\mu}_i - \mu_{it}\right)\frac{\gamma_{\theta_i}}{\sigma_{\mu_i}} dW^{\mathcal{E}}_{\mu_i t}.$$

The household i's estimates $\widehat{\theta}^i_{it}$ and $\widehat{\theta}^i_{jt}$, where $i \neq j$, are

$$d\widehat{\theta}^i_{it} = -b\widehat{\theta}^i_{it}dt + \left(\bar{\mu}_i - \mu_{it}\right)\frac{\gamma_{\theta_i}}{\sigma_{\mu_i}} dW^i_{\mu_i t} + \frac{\phi\sigma_{\theta_i}}{\sqrt{1-\phi^2}} dW^i_{s_i t}$$
$$d\widehat{\theta}^i_{jt} = -b\widehat{\theta}^i_{jt}dt + \left(\bar{\mu}_j - \mu_{jt}\right)\frac{\gamma_{\theta_j}}{\sigma_{\mu_j}} dW^i_{\mu_j t} - \frac{\phi\sigma_{\theta_j}}{\sqrt{1-\phi^2}} dW^i_{s_j t},$$

where the estimation error γ_{θ_i} is

$$\gamma_{\theta_i} = \frac{-b\sigma^2_{\mu_i} + \sigma^2_{\mu_i}\sqrt{b^2 + \left(\bar{\mu}_i - \mu_{it}\right)^2 \frac{\sigma^2_{\theta_i}}{\sigma^2_{\mu_i}}}}{\left(\bar{\mu}_i - \mu_{it}\right)^2}.$$

All shocks are Brownian motions under country i's probability measure and are mutually independent. The rest follow similar results in Section 5.1.2.

E Additional Data Description

This section contains information on other data sets used in the empirical analysis in Section 6. The description on the forecast data can be found in Section 6.1. All data are from 2010Q1 to 2018Q4 at quarterly frequency.

E.1 Nominal and real bilateral exchange rates

The baseline data uses USD as the quote currency. The nominal exchange rate data for AUD/USD, CAD/USD, EUR/USD, JPY/USD, and GBP/USD are from MSCI. As mentioned in the main text, each country takes turns to serves as the quote currency (home country). To obtain real exchange rates, I multiply the nominal exchange rate with local CPI and then divide by the home country's CPI. The CPI data is from Global Financial Data (GFD) and is normalized to 100 in 2010Q1. The CPI for Euro is the average of the four countries in EMU4.

E.2 Real GDP

The real GDP data are from GFD at quarterly frequency with seasonal adjustments. Although each county sets their own base year, it will not matter when dealing with log growth rates. The real GDP volatility is constructed using trailing moving standard deviation with a window size of 4. When there are fewer elements than the window size in the beginning of the sample, I take the standard deviation over the elements that are available.

E.3 Real consumption

The real consumption data are from OECD with seasonal adjustments and are in local currencies. Each county sets their own base year, and it will not matter when dealing with log growth rates.

E.4 Population

The population data are from the World Bank World Development Indicators and are at annual level. In order to derive per capita consumption at quarterly level, I interpolate data between two adjacent years by assuming a linear quarterly population growth rate within each year.

E.5 Nominal and real interest rates

The nominal risk-free interest rate is defined to be the three-month treasury bill or government bond rate in each country obtained from GFD, which gathers information from each central bank. The interest rate for Euro bond is the average of the rates of four countries in EMU4. To obtain real interest rates, each country's nominal interest rates are deflated by the respective inflation rate, calculated from quarterly percentage changes in the country-specific CPI.

F 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

$$dY_t = \mu_Y (t, Y_t)\, dt + \boldsymbol{\sigma_Y} (t, Y_t)^\top \mathbf{dW_t^{\mathcal{E}}},$$

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

$$P_t^Y M_{At}^{\mathcal{E}} = \mathbb{E}_t^{\mathcal{E}} \left[M_{AT}^{\mathcal{E}} Y_T \right]. \tag{47}$$

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

$$dP_t^Y = \mu_{P^Y} \left(t, P_t^Y\right) dt + \boldsymbol{\sigma_{P^Y}} \left(t, P_t^Y\right)^\top \mathbf{dW_t^{\mathcal{E}}}. \tag{48}$$

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

Proposition 13. *Given the price process (48), the drift $\mu_{P^Y} \left(t, P_t^Y\right)$ and volatility $\boldsymbol{\sigma_{P^Y}} \left(t, P_t^Y\right)$ are*

$$\begin{aligned}
\mu_{P^Y} \left(t, P_t^Y\right) &= r_{At} P_t^Y + \boldsymbol{\sigma}_{\boldsymbol{M_{At}}}^\top \Omega_W \boldsymbol{\sigma_{P^Y}} \left(t, P_t^Y\right) \\
\boldsymbol{\sigma_{P^Y}} \left(t, P_t^Y\right) &= \boldsymbol{\sigma_{M_{At}}} M_{At}^{\mathcal{E}} + \left(\mathbb{E}_t^{\mathcal{E}} \left[M_{AT}^{\mathcal{E}} Y_T\right]\right)^{-1} \mathbb{E}_t^{\mathcal{E}} \left[M_{AT}^{\mathcal{E}} \left(\frac{\widetilde{Y}_T}{\widetilde{Y}_t} \boldsymbol{\sigma_Y}(t, Y_t) - Y_T \boldsymbol{\sigma_{M_{At}}} \right) \right],
\end{aligned}$$

where

$$\widetilde{Y}_t = \exp \left[\int_0^t \left(\mu_Y'(s, Y_s) - \frac{1}{2} \boldsymbol{\sigma_Y'}(s, Y_s)^\top \boldsymbol{\sigma_Y'}(s, Y_s) \right) ds + \int_0^t \boldsymbol{\sigma_Y'}(s, Y_s)^\top \mathbf{dW_s} \right].$$

Similar expressions can be derived for assets denominated in B's currency. The derivations are provided in Appendix H.

F.1 Stock price

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

$$E_t^i M_{At}^{\mathcal{E}} = \mathbb{E}_t^{\mathcal{E}}\left[\int_t^\infty M_{As}^{\mathcal{E}} q_{is} X_{is} ds\right], \tag{49}$$

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

$$\frac{dE_t^i + q_{it}X_{it}dt}{E_t^i} = \mu_{E^i t}dt + \boldsymbol{\sigma}_{\boldsymbol{E^i t}}^\top \mathbf{dW_t^{\mathcal{E}}}, \tag{50}$$

where $\mu_{E^i t}$ and $\boldsymbol{\sigma}_{\boldsymbol{E^i t}}$ 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_t^i, in units of the global numeraire, follows* (50)*, where drift and volatilities are given by*

$$\begin{aligned}\mu_{E^i t} &= r_{At} + \boldsymbol{\sigma}_{\boldsymbol{M_A t}}^\top \Omega_W \boldsymbol{\sigma}_{\boldsymbol{E^i t}} \\ \boldsymbol{\sigma}_{\boldsymbol{E^i t}} &= \boldsymbol{\sigma}_{\boldsymbol{M_A t}} + \frac{\mathbb{E}_t^{\mathcal{E}}\left[\int_t^\infty \mathrm{D}_t\left(M_{As}^{\mathcal{E}} q_{is} X_{is}\right) ds\right]}{E_t^i M_{At}^{\mathcal{E}}},\end{aligned}$$

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

G 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

$$dX_t = \mu(t, X_t)\, dt + \boldsymbol{\sigma}(t, X_t)^\top \mathbf{dW_t}, \quad X_0 = x.$$

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

$$d\widetilde{X}_t = \mu'(t, X_t)\, \widetilde{X}_t dt + \boldsymbol{\sigma}'(t, X_t)^\top \widetilde{X}_t \mathbf{dW_t},$$

or

$$\widetilde{X}_t = \exp\left[\int_0^t \left(\mu'(s, X_s) - \frac{1}{2}\boldsymbol{\sigma}'(s, X_s)^\top \boldsymbol{\sigma}'(s, X_s)\right) ds + \int_0^t \boldsymbol{\sigma}'(s, X_s)^\top \mathbf{dW_s}\right],$$

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

$$\mathrm{D}_s X_t = \frac{\widetilde{X}_t}{\widetilde{X}_s} \boldsymbol{\sigma}(t, X_t) \mathbb{1}_{\{s \le t\}}.$$

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*

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

then we have

$$F = \mathbb{E}(F) + \int_0^T \mathbb{E}(\mathrm{D}_s F|\mathcal{F}_s) dW_s.$$

H Proofs

Proof of Proposition 1

Proof. The welfare optimization in the planner's problem is

$$\max_{\{C^A_{At}, C^B_{At}, C^A_{Bt}, C^B_{Bt}\}} \mathbb{E}^{\mathcal{E}}_0 \left[\int_0^\infty e^{-\beta t}\left(\lambda_A \eta^A_t \log C_{At} + \lambda_B \eta^B_t \log C_{Bt}\right) dt\right] \tag{51}$$

subject to the resource constraints

$$C^A_{At} + C^B_{At} = X_{At} \qquad\qquad C^A_{Bt} + C^B_{Bt} = X_{Bt}. \tag{52}$$

Note that the household's original problem (2) is defined under their own probability measure. The change-of-measure variable η^i_t 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

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

Denote ν_{it} as the Lagrange multiplier on the resource constraint (52). The FOCs for country A in the planner's problem (51) imply

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

and for country B

$$\lambda_B \eta_t^B (1-\alpha) \frac{1}{C_{At}^B} = \nu_{At}$$
$$\lambda_B \eta_t^B \alpha \frac{1}{C_{Bt}^B} = \nu_{Bt}.$$

Dividing the two equations above for each ν_{it} and using (52), we achieve the equilibrium consumption allocations for country A as

$$C_{At}^A = \frac{\lambda_A \alpha \eta_t^A}{\lambda_A \eta_t^A \alpha + \lambda_B \eta_t^B (1-\alpha)} X_{At} \tag{53}$$
$$C_{Bt}^A = \frac{\lambda_A (1-\alpha) \eta_t^A}{\lambda_A \eta_t^A (1-\alpha) + \lambda_B \eta_t^B \alpha} X_{Bt}, \tag{54}$$

and for country B as

$$C_{At}^B = \frac{\lambda_B (1-\alpha) \eta_t^B}{\lambda_A \eta_t^A \alpha + \lambda_B \eta_t^B (1-\alpha)} X_{At} \tag{55}$$
$$C_{Bt}^B = \frac{\lambda_B \alpha \eta_t^B}{\lambda_A \eta_t^A (1-\alpha) + \lambda_B \eta_t^B \alpha} X_{Bt}. \tag{56}$$

Then we immediately obtain the optimal consumption baskets.

I now derive the equilibrium process for C_{At}, and the analogous argument will apply for C_{Bt}. Define

$$x(\boldsymbol{\eta_t}) \equiv \lambda_A \alpha \eta_t^A + \lambda_B (1-\alpha) \eta_t^B$$
$$y(\boldsymbol{\eta_t}) \equiv \lambda_A (1-\alpha) \eta_t^A + \lambda_B \alpha \eta_t^B.$$

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

$$\log C_{At} = \alpha \left[\log X_{At} - \log x(\boldsymbol{\eta_t})\right] + (1-\alpha) \left[\log X_{Bt} - \log y(\boldsymbol{\eta_t})\right] + \log \eta_t^A + const.$$

Applying Ito's lemma, we have

$$
\begin{aligned}
d\log C_{At} &= \frac{dC_{At}}{C_{At}} - \frac{1}{2}\frac{1}{C_{At}^2}\left(dC_{At}\right)^2 \\
&= \mu_{C_At}dt + \sigma_A^{C_A}dZ_{At} + \sigma_B^{C_A}dZ_{Bt} + \boldsymbol{\sigma_{\mu t}^{C_A}}^\top \mathbf{dW_{\mu t}^{\mathcal{E}}} \\
&\quad - \frac{1}{2}\left(\sigma_A^{C_A\,2} + \sigma_B^{C_A\,2} + 2\rho\sigma_A^{C_A}\sigma_B^{C_A} + \boldsymbol{\sigma_{\mu t}^{C_A}}^\top \boldsymbol{\sigma_{\mu t}^{C_A}}\right)dt \\
d\log X_{At} &= \frac{dX_{At}}{X_{At}} - \frac{1}{2}\frac{1}{X_{At}^2}\left(dX_{At}\right)^2 \\
&= \mu_{At}dt + \sigma_A dZ_{At} - \frac{1}{2}\sigma_A^2 dt \\
d\log X_{Bt} &= \frac{dX_{Bt}}{X_{Bt}} - \frac{1}{2}\frac{1}{X_{Bt}^2}\left(dX_{Bt}\right)^2 \\
&= \mu_{Bt}dt + \sigma_B dZ_{Bt} - \frac{1}{2}\sigma_B^2 dt
\end{aligned}
$$

$$
\begin{aligned}
d\log \eta_t^A &= \frac{d\eta_t^A}{\eta_t^A} - \frac{1}{2}\frac{1}{(\eta_t^A)^2}\left(d\eta_t^A\right)^2 \\
&= -\boldsymbol{\xi}_{\boldsymbol{At}}^\top \mathbf{dW_{\mu t}^{\mathcal{E}}} - \frac{1}{2}\boldsymbol{\xi}_{\boldsymbol{At}}^\top \boldsymbol{\xi}_{\boldsymbol{At}} dt \\
d\log \eta_t^B &= \frac{d\eta_t^B}{\eta_t^B} - \frac{1}{2}\frac{1}{(\eta_t^B)^2}\left(d\eta_t^B\right)^2 \\
&= -\boldsymbol{\xi}_{\boldsymbol{Bt}}^\top \mathbf{dW_{\mu t}^{\mathcal{E}}} - \frac{1}{2}\boldsymbol{\xi}_{\boldsymbol{Bt}}^\top \boldsymbol{\xi}_{\boldsymbol{Bt}} dt,
\end{aligned}
$$

and

$$
\begin{aligned}
d\log x(\boldsymbol{\eta_t}) &= \frac{\lambda_A\alpha}{x(\boldsymbol{\eta_t})}d\eta_t^A + \frac{\lambda_B(1-\alpha)}{x(\boldsymbol{\eta_t})}d\eta_t^B - \frac{1}{2}\frac{\lambda_A^2\alpha^2}{x(\boldsymbol{\eta_t})^2}\left(d\eta_t^A\right)^2 - \frac{1}{2}\frac{\lambda_B^2(1-\alpha)^2}{x(\boldsymbol{\eta_t})^2}\left(d\eta_t^B\right)^2 \\
&= -\frac{\lambda_A\alpha\eta_t^A}{\lambda_A\eta_t^A\alpha + \lambda_B\eta_t^B(1-\alpha)}\boldsymbol{\xi}_{\boldsymbol{At}}^\top \mathbf{dW_{\mu t}^{\mathcal{E}}} - \frac{\lambda_B(1-\alpha)\eta_t^B}{\lambda_A\eta_t^A\alpha + \lambda_B\eta_t^B(1-\alpha)}\boldsymbol{\xi}_{\boldsymbol{Bt}}^\top \mathbf{dW_{\mu t}^{\mathcal{E}}} \\
&\quad - \frac{1}{2}\frac{\lambda_A^2\alpha^2(\eta_t^A)^2}{\left(\lambda_A\eta_t^A\alpha + \lambda_B\eta_t^B(1-\alpha)\right)^2}\boldsymbol{\xi}_{\boldsymbol{At}}^\top \boldsymbol{\xi}_{\boldsymbol{At}} dt - \frac{1}{2}\frac{\lambda_B^2(1-\alpha)^2(\eta_t^B)^2}{\left(\lambda_A\eta_t^A\alpha + \lambda_B\eta_t^B(1-\alpha)\right)^2}\boldsymbol{\xi}_{\boldsymbol{Bt}}^\top \boldsymbol{\xi}_{\boldsymbol{Bt}} dt \\
d\log y(\boldsymbol{\eta_t}) &= \frac{\lambda_A(1-\alpha)}{y(\boldsymbol{\eta_t})}d\eta_t^A + \frac{\lambda_B\alpha}{y(\boldsymbol{\eta_t})}d\eta_t^B - \frac{1}{2}\frac{\lambda_A^2(1-\alpha)^2}{y(\boldsymbol{\eta_t})^2}\left(d\eta_t^A\right)^2 - \frac{1}{2}\frac{\lambda_B^2\alpha^2}{y(\boldsymbol{\eta_t})^2}\left(d\eta_t^B\right)^2 \\
&= -\frac{\lambda_A(1-\alpha)\eta_t^A}{\lambda_A\eta_t^A(1-\alpha) + \lambda_B\eta_t^B\alpha}\boldsymbol{\xi}_{\boldsymbol{At}}^\top \mathbf{dW_{\mu t}^{\mathcal{E}}} - \frac{\lambda_B\alpha\eta_t^B}{\lambda_A\eta_t^A(1-\alpha) + \lambda_B\eta_t^B\alpha}\boldsymbol{\xi}_{\boldsymbol{Bt}}^\top \mathbf{dW_{\mu t}^{\mathcal{E}}} \\
&\quad - \frac{1}{2}\frac{\lambda_A^2(1-\alpha)^2(\eta_t^A)^2}{\left(\lambda_A\eta_t^A(1-\alpha) + \lambda_B\eta_t^B\alpha\right)^2}\boldsymbol{\xi}_{\boldsymbol{At}}^\top \boldsymbol{\xi}_{\boldsymbol{At}} dt - \frac{1}{2}\frac{\lambda_B^2\alpha^2(\eta_t^B)^2}{\left(\lambda_A\eta_t^A(1-\alpha) + \lambda_B\eta_t^B\alpha\right)^2}\boldsymbol{\xi}_{\boldsymbol{Bt}}^\top \boldsymbol{\xi}_{\boldsymbol{Bt}} dt.
\end{aligned}
$$

Matching coefficients, we have for C_A that

$$\begin{aligned}
\mu_{C_A t} &= \alpha\mu_{At} + (1-\alpha)\mu_{Bt} - \frac{1}{2}\alpha\sigma_A^2 - \frac{1}{2}(1-\alpha)\sigma_B^2 - \frac{1}{2}\boldsymbol{\xi_{At}}^\top\boldsymbol{\xi_{At}} \\
&\quad + \frac{1}{2}\frac{\lambda_A^2\alpha^3(\eta_t^A)^2}{\left(\lambda_A\eta_t^A\alpha + \lambda_B\eta_t^B(1-\alpha)\right)^2}\boldsymbol{\xi_{At}}^\top\boldsymbol{\xi_{At}} + \frac{1}{2}\frac{\lambda_B^2\alpha\left(1-\alpha\right)^2(\eta_t^B)^2}{\left(\lambda_A\eta_t^A\alpha + \lambda_B\eta_t^B(1-\alpha)\right)^2}\boldsymbol{\xi_{Bt}}^\top\boldsymbol{\xi_{Bt}} \\
&\quad + \frac{1}{2}\frac{\lambda_A^2(1-\alpha)^3(\eta_t^A)^2}{\left(\lambda_A\eta_t^A(1-\alpha) + \lambda_B\eta_t^B\alpha\right)^2}\boldsymbol{\xi_{At}}^\top\boldsymbol{\xi_{At}} + \frac{1}{2}\frac{\lambda_B^2(1-\alpha)\alpha^2(\eta_t^B)^2}{\left(\lambda_A\eta_t^A(1-\alpha) + \lambda_B\eta_t^B\alpha\right)^2}\boldsymbol{\xi_{Bt}}^\top\boldsymbol{\xi_{Bt}} \\
&\quad + \frac{1}{2}\left({\sigma_A^{C_A}}^2 + {\sigma_B^{C_A}}^2 + 2\rho\sigma_A^{C_A}\sigma_B^{C_A} + \boldsymbol{\sigma_{\mu t}^{C_A}}^\top\boldsymbol{\sigma_{\mu t}^{C_A}}\right) \\
\sigma_A^{C_A} &= \alpha\sigma_A \\
\sigma_B^{C_A} &= (1-\alpha)\sigma_B \\
\boldsymbol{\sigma_{\mu t}^{C_A}} &= -\boldsymbol{\xi_{At}} + \frac{\lambda_A\alpha^2\eta_t^A}{\lambda_A\eta_t^A\alpha + \lambda_B\eta_t^B(1-\alpha)}\boldsymbol{\xi_{At}} + \frac{\lambda_A(1-\alpha)^2\eta_t^A}{\lambda_A\eta_t^A(1-\alpha) + \lambda_B\eta_t^B\alpha}\boldsymbol{\xi_{At}} \\
&\quad + \frac{\lambda_B\alpha\left(1-\alpha\right)\eta_t^B}{\lambda_A\eta_t^A\alpha + \lambda_B\eta_t^B(1-\alpha)}\boldsymbol{\xi_{Bt}} + \frac{\lambda_B(1-\alpha)\alpha\eta_t^B}{\lambda_A\eta_t^A(1-\alpha) + \lambda_B\eta_t^B\alpha}\boldsymbol{\xi_{Bt}},
\end{aligned}$$

and for C_B that

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

$\square$

Proof of Proposition 2

Proof. Taking log on both sides of (8) for $i = A$, we have $\log M^{\mathcal{E}}_{At} = -\beta t + \log \eta^A_t - \log C_{At}$. Applying Ito's lemma gives

$$\begin{aligned}
d\log M^{\mathcal{E}}_{At} &= \frac{dM^{\mathcal{E}}_{At}}{M^{\mathcal{E}}_{At}} - \frac{1}{2}\frac{1}{M^2_{At}}\left(dM_{At}\right)^2 \\
&= -\beta dt + d\log \eta^A_t - d\log C_{At} \\
&= \mu_{M_A t}dt - \sigma^{M_A}_A dZ_{At} - \sigma^{M_A}_B dZ_{Bt} - \boldsymbol{\sigma}^{\boldsymbol{M_A}\top}_{\boldsymbol{\mu t}} \mathbf{dW}^{\mathcal{E}}_{\boldsymbol{\mu}\mathbf{t}} \\
&\quad - \frac{1}{2}\left({\sigma^{M_A}_A}^2 + {\sigma^{M_A}_B}^2 + 2\rho\sigma^{M_A}_A\sigma^{M_A}_B + \boldsymbol{\sigma}^{\boldsymbol{M_A}\top}_{\boldsymbol{\mu t}}\boldsymbol{\sigma}^{\boldsymbol{M_A}}_{\boldsymbol{\mu t}}\right)dt.
\end{aligned}$$

Matching coefficients, this yields

$$\begin{aligned}
\mu_{M_A t} &= -\beta - \mu_{C_A t} - \frac{1}{2}\boldsymbol{\xi}^\top_{\boldsymbol{At}}\boldsymbol{\xi}_{\boldsymbol{At}} + \frac{1}{2}\left({\sigma^{C_A}_A}^2 + {\sigma^{C_A}_B}^2 + 2\rho\sigma^{C_A}_A\sigma^{C_A}_B + \boldsymbol{\sigma}^{\boldsymbol{C_A}\top}_{\boldsymbol{\mu t}}\boldsymbol{\sigma}^{\boldsymbol{C_A}}_{\boldsymbol{\mu t}}\right) \\
&\quad + \frac{1}{2}\left({\sigma^{M_A}_A}^2 + {\sigma^{M_A}_{Bt}}^2 + 2\rho\sigma^{M_A}_A\sigma^{M_A}_B + \boldsymbol{\sigma}^{\boldsymbol{M_A}\top}_{\boldsymbol{\mu t}}\boldsymbol{\sigma}^{\boldsymbol{M_A}}_{\boldsymbol{\mu t}}\right) \\
\sigma^{M_A}_A &= \sigma^{C_A}_A \\
\sigma^{M_A}_B &= \sigma^{C_A}_B \\
\boldsymbol{\sigma}^{\boldsymbol{M_A}}_{\boldsymbol{\mu t}} &= \boldsymbol{\xi}_{\boldsymbol{At}} + \boldsymbol{\sigma}^{\boldsymbol{C_A}}_{\boldsymbol{\mu t}}.
\end{aligned}$$

Similarly, the coefficients for $M^{\mathcal{E}}_{Bt}$ are

$$\begin{aligned}
\mu_{M_B t} &= -\beta - \mu_{C_B t} - \frac{1}{2}\boldsymbol{\xi}^\top_{\boldsymbol{Bt}}\boldsymbol{\xi}_{\boldsymbol{Bt}} + \frac{1}{2}\left({\sigma^{C_B}_A}^2 + {\sigma^{C_B}_B}^2 + 2\rho\sigma^{C_B}_A\sigma^{C_B}_B + \boldsymbol{\sigma}^{\boldsymbol{C_B}\top}_{\boldsymbol{\mu t}}\boldsymbol{\sigma}^{\boldsymbol{C_B}}_{\boldsymbol{\mu t}}\right) \\
&\quad + \frac{1}{2}\left({\sigma^{M_B}_A}^2 + {\sigma^{M_B}_B}^2 + 2\rho\sigma^{M_B}_A\sigma^{M_B}_B + \boldsymbol{\sigma}^{\boldsymbol{M_B}\top}_{\boldsymbol{\mu t}}\boldsymbol{\sigma}^{\boldsymbol{M_B}}_{\boldsymbol{\mu t}}\right) \\
\sigma^{M_B}_A &= \sigma^{C_B}_A \\
\sigma^{M_B}_B &= \sigma^{C_B}_B \\
\boldsymbol{\sigma}^{\boldsymbol{M_B}}_{\boldsymbol{\mu t}} &= \boldsymbol{\xi}_{\boldsymbol{Bt}} + \boldsymbol{\sigma}^{\boldsymbol{C_B}}_{\boldsymbol{\mu t}}.
\end{aligned}$$

□

Proof of Proposition 3

Proof. Taking log on both sides of exchange rate equation in Proposition 3, we have

$$\log S_t = \log \eta^B_t - \log \eta^A_t + \log C_{At} - \log C_{Bt}.$$

Applying Ito's lemma, we get

$$\begin{aligned}
d\log S_t &= \frac{dS_t}{S_t} - \frac{1}{2}\frac{1}{S_t^2}(dS_t)^2 \\
&= d\log\eta_t^B - d\log\eta_t^A + d\log C_{At} - d\log C_{Bt} \\
&= \mu_{St}dt + \sigma_A^S dZ_{At} + \sigma_B^S dZ_{Bt} + \boldsymbol{\sigma_{\mu t}^{S}}^\top \mathbf{dW}_{\boldsymbol{\mu t}}^{\mathcal{E}} \\
&\quad - \frac{1}{2}\left(\sigma_A^{S\,2} + \sigma_B^{S\,2} + 2\rho\sigma_A^S\sigma_B^S + \boldsymbol{\sigma_{\mu t}^{S}}^\top \boldsymbol{\sigma_{\mu t}^{S}}\right)dt.
\end{aligned}$$

Matching coefficients, this yields

$$\begin{aligned}
\mu_{St} &= \mu_{C_A t} - \mu_{C_B t} + \frac{1}{2}\boldsymbol{\xi}_{\boldsymbol{At}}^\top\boldsymbol{\xi}_{\boldsymbol{At}} - \frac{1}{2}\boldsymbol{\xi}_{\boldsymbol{Bt}}^\top\boldsymbol{\xi}_{\boldsymbol{Bt}} + \frac{1}{2}\left(\sigma_A^{S\,2} + \sigma_B^{S\,2} + 2\rho\sigma_A^S\sigma_B^S + \boldsymbol{\sigma_{\mu t}^{S}}^\top \boldsymbol{\sigma_{\mu t}^{S}}\right) \\
&\quad - \frac{1}{2}\left(\sigma_A^{C_A\,2} + \sigma_B^{C_A\,2} + 2\rho\sigma_A^{C_A}\sigma_B^{C_A} + \boldsymbol{\sigma_{\mu t}^{C_A}}^\top \boldsymbol{\sigma_{\mu t}^{C_A}}\right) + \frac{1}{2}\left(\sigma_A^{C_B\,2} + \sigma_B^{C_B\,2} + 2\rho\sigma_A^{C_B}\sigma_B^{C_B} + \boldsymbol{\sigma_{\mu t}^{C_B}}^\top \boldsymbol{\sigma_{\mu t}^{C_B}}\right) \\
\sigma_A^S &= \sigma_A^{C_A} - \sigma_A^{C_B} \\
\sigma_B^S &= \sigma_B^{C_A} - \sigma_B^{C_B} \\
\boldsymbol{\sigma_{\mu t}^{S}} &= \boldsymbol{\xi}_{\boldsymbol{At}} - \boldsymbol{\xi}_{\boldsymbol{Bt}} + \boldsymbol{\sigma_{\mu t}^{C_A}} - \boldsymbol{\sigma_{\mu t}^{C_B}}.
\end{aligned}$$

It is easy to show that $\mu_{St} = r_{At} - r_{Bt} + \boldsymbol{\sigma}_{\boldsymbol{M_A t}}^\top \Omega_W(\boldsymbol{\sigma}_{\boldsymbol{M_A t}} - \boldsymbol{\sigma}_{\boldsymbol{M_B t}})$, 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 $\mathbf{dW}_{\boldsymbol{\mu t}}^{\mathcal{E}}$ is two-dimensional. □

Proof of Proposition 4

Proof. Taking the difference between $\boldsymbol{\sigma_{\mu t}^{C_A}}$ and $\boldsymbol{\sigma_{\mu t}^{C_B}}$, we get

$$\begin{aligned}
\boldsymbol{\sigma_{\mu t}^{C_A}} - \boldsymbol{\sigma_{\mu t}^{C_B}} &= \left[\frac{\lambda_A\alpha(2\alpha-1)}{\lambda_A\eta_t^A\alpha + \lambda_B\eta_t^B(1-\alpha)} + \frac{\lambda_A(1-\alpha)(1-2\alpha)}{\lambda_A\eta_t^A(1-\alpha) + \lambda_B\eta_t^B\alpha}\right](\boldsymbol{\xi}_{\boldsymbol{At}} - \boldsymbol{\xi}_{\boldsymbol{Bt}})\eta_t^A \\
&\quad - \left[\frac{\lambda_B(1-\alpha)(2\alpha-1)}{\lambda_A\eta_t^A\alpha + \lambda_B\eta_t^B(1-\alpha)} + \frac{\lambda_B\alpha(1-2\alpha)}{\lambda_A\eta_t^A(1-\alpha) + \lambda_B\eta_t^B\alpha}\right](\boldsymbol{\xi}_{\boldsymbol{At}} - \boldsymbol{\xi}_{\boldsymbol{Bt}})\eta_t^B - (\boldsymbol{\xi}_{\boldsymbol{At}} - \boldsymbol{\xi}_{\boldsymbol{Bt}}) \\
&= \frac{(2\alpha-1)^2\lambda_A\lambda_B\eta_t^A\eta_t^B}{(\lambda_A\eta_t^A\alpha + \lambda_B\eta_t^B(1-\alpha))(\lambda_A\eta_t^A(1-\alpha) + \lambda_B\eta_t^B\alpha)}(\boldsymbol{\xi}_{\boldsymbol{At}} - \boldsymbol{\xi}_{\boldsymbol{Bt}}) - (\boldsymbol{\xi}_{\boldsymbol{At}} - \boldsymbol{\xi}_{\boldsymbol{Bt}}) \\
&\equiv \frac{H_t}{K_t}(\boldsymbol{\xi}_{\boldsymbol{At}} - \boldsymbol{\xi}_{\boldsymbol{Bt}}) - (\boldsymbol{\xi}_{\boldsymbol{At}} - \boldsymbol{\xi}_{\boldsymbol{Bt}}).
\end{aligned}$$

We can see that after some algebra we can see that $K_t > H_t$.[29] Therefore, we can express $\boldsymbol{\sigma}_{\mu t}^{C_A} - \boldsymbol{\sigma}_{\mu t}^{C_B}$ as $-f(\boldsymbol{\eta_t})(\boldsymbol{\xi_{At}} - \boldsymbol{\xi_{Bt}})$, where $0 < f(\boldsymbol{\eta_t}) < 1$.

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

Proof of Proposition 5

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

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

$$
\begin{aligned}
RX_{t+dt} &= \frac{S_{t+dt}}{S_t}(1 + r_{Bt}dt) - (1 + r_{At}dt) \\
&= \frac{S_{t+dt}(1 + r_{Bt}dt)}{S_t(1 + r_{At}dt)}, \quad \text{since } dt \text{ is small.}
\end{aligned}
$$

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

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

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

$$
\left(\boldsymbol{\sigma}_{\mu t}^{M_A} + \boldsymbol{\sigma}_{\mu t}^{M_B}\right)^\top \left(\boldsymbol{\sigma}_{\mu t}^{M_A} - \boldsymbol{\sigma}_{\mu t}^{M_B}\right) = \left(\boldsymbol{\sigma}_{\mu t}^{C_A} + \boldsymbol{\sigma}_{\mu t}^{C_B}\right)^\top \left(\boldsymbol{\xi_{At}} - \boldsymbol{\xi_{Bt}} + \boldsymbol{\sigma}_{\mu t}^{C_A} - \boldsymbol{\sigma}_{\mu t}^{C_B}\right).
$$

Using expressions for $\boldsymbol{\sigma}_{\mu t}^{C_A}$ and $\boldsymbol{\sigma}_{\mu t}^{C_B}$, it is straightforward to show that with Definition 1, the RHS above equals

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

Therefore, we have

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

[29] This is because $\alpha^2 + (1-\alpha)^2 - (2\alpha-1)^2 = 2\alpha - 2\alpha^2 > 0$ as $\alpha < 1$.

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

$$\begin{aligned}\frac{1}{dt}\mathbb{E}_t^{\mathcal{E}}\left[RX_{t+dt}\right] &= \frac{1}{dt}\mathbb{E}_t^{\mathcal{E}}\left[\left(\frac{dM_{At}^{\mathcal{E}}}{M_{At}^{\mathcal{E}}}\right)^2 - \frac{dM_{At}^{\mathcal{E}}}{M_{At}^{\mathcal{E}}}\frac{dM_{Bt}^{\mathcal{E}}}{M_{Bt}^{\mathcal{E}}}\right] \\ &= \frac{1}{dt}\mathrm{Cov}_t\left(\frac{dM_{At}^{\mathcal{E}}}{M_{At}^{\mathcal{E}}}, \frac{dM_{At}^{\mathcal{E}}}{M_{At}^{\mathcal{E}}} - \frac{dM_{Bt}^{\mathcal{E}}}{M_{Bt}^{\mathcal{E}}}\right) \\ &= \boldsymbol{\sigma_{\mu t}^{M_A}}^\top\left(\boldsymbol{\sigma_{\mu t}^{M_A}} - \boldsymbol{\sigma_{\mu t}^{M_B}}\right).\end{aligned}$$

Writing this out explicitly gives

$$\begin{aligned}\frac{1}{dt}\mathbb{E}_t^{\mathcal{E}}\left[RX_{t+dt}\right] = &\frac{\lambda_A\lambda_B(2\alpha-1)^2}{x(\boldsymbol{\eta_t})y(\boldsymbol{\eta_t})}\eta_t^A\eta_t^B\left(\boldsymbol{\xi_{At}} - \boldsymbol{\xi_{Bt}}\right)^\top\left(\boldsymbol{\xi_{At}} - \boldsymbol{\xi_{Bt}}\right) \\ &\times\left[\alpha(1-\alpha)\left(\eta_t^A + \eta_t^B\right)\left(\eta_t^A - \eta_t^B\right) + (2\alpha-1)^2\eta_t^A\eta_t^B\right]\end{aligned}$$

□

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

$$\begin{aligned}d\log\omega_t &= \frac{1}{\omega_t}d\omega_t - \frac{1}{2}\frac{1}{\omega_t^2}(d\omega_t)^2 \\ &= \frac{1}{\omega_t}\left(\mu_{\omega t}dt + \sigma_{At}^{\omega}dZ_{At} + \sigma_{Bt}^{\omega}dZ_{Bt} + \boldsymbol{\sigma_{\mu t}^{\omega}}^\top\mathbf{dW}_{\boldsymbol{\mu t}}^{\mathcal{E}}\right) \\ &\quad - \frac{1}{2}\frac{1}{\omega_t^2}\left({\sigma_{At}^{\omega}}^2 + {\sigma_{Bt}^{\omega}}^2 + 2\rho\sigma_{At}^{\omega}\sigma_{Bt}^{\omega} + \boldsymbol{\sigma_{\mu t}^{\omega}}^\top\boldsymbol{\sigma_{\mu t}^{\omega}}\right)dt\end{aligned}$$

$$\begin{aligned}
d\log\eta_t^A &= -\boldsymbol{\xi_{At}}^\top \mathbf{dW}^{\mathcal{E}}_{\boldsymbol{\mu}\mathbf{t}} - \frac{1}{2}\boldsymbol{\xi}_{\boldsymbol{At}}^\top\boldsymbol{\xi_{At}}dt \\
d\log\eta_t^B &= -\boldsymbol{\xi_{Bt}}^\top \mathbf{dW}^{\mathcal{E}}_{\boldsymbol{\mu}\mathbf{t}} - \frac{1}{2}\boldsymbol{\xi}_{\boldsymbol{Bt}}^\top\boldsymbol{\xi_{Bt}}dt \\
d\log X_{Bt} - d\log X_{At} &= (\mu_{Bt} - \mu_{At})\,dt + \sigma_B dZ_{Bt} - \sigma_A dZ_{At} - \frac{1}{2}\left(\sigma_B^2 - \sigma_A^2\right)dt \\
d\log(1+\omega_t) &= \frac{1}{1+\omega_t}\left(\mu_{\omega t}dt + \sigma^\omega_{At}dZ_{At} + \sigma^\omega_{Bt}dZ_{Bt} + \boldsymbol{\sigma^{\omega}_{\mu t}}^\top \mathbf{dW}^{\mathcal{E}}_{\boldsymbol{\mu}\mathbf{t}}\right) \\
&\quad - \frac{1}{2}\frac{1}{(1+\omega_t)^2}\left({\sigma^\omega_{At}}^2 + {\sigma^\omega_{Bt}}^2 + 2\rho\sigma^\omega_{At}\sigma^\omega_{Bt} + \boldsymbol{\sigma^{\omega}_{\mu t}}^\top\boldsymbol{\sigma^{\omega}_{\mu t}}\right)dt \\
d\log(\alpha^2 + (1-\alpha)^2\omega_t) &= \frac{(1-\alpha)^2}{\alpha^2+(1-\alpha)^2\omega_t}\left(\mu_{\omega t}dt + \sigma^\omega_{At}dZ_{At} + \sigma^\omega_{Bt}dZ_{Bt} + \boldsymbol{\sigma^{\omega}_{\mu t}}^\top \mathbf{dW}^{\mathcal{E}}_{\boldsymbol{\mu}\mathbf{t}}\right) \\
&\quad - \frac{1}{2}\frac{(1-\alpha)^4}{(\alpha^2+(1-\alpha)^2\omega_t)^2}\left({\sigma^\omega_{At}}^2 + {\sigma^\omega_{Bt}}^2 + 2\rho\sigma^\omega_{At}\sigma^\omega_{Bt} + \boldsymbol{\sigma^{\omega}_{\mu t}}^\top\boldsymbol{\sigma^{\omega}_{\mu t}}\right)dt.
\end{aligned}$$

Taking logs on both sides of (40) and then differentiating both sides, we get

$$\begin{aligned}
0 = &-\gamma d\log\omega_t + d\log\eta_t^A - d\log\eta_t^B + (1-2\alpha)(1-\gamma)\left(d\log X_{Bt} - d\log X_{At}\right) \\
&+ (1-2\alpha)(1-\gamma)\left[d\log(1+\omega_t) - d\log(\alpha^2+(1-\alpha)^2\omega_t)\right].
\end{aligned}$$

Matching coefficients, we obtain the drift part as

$$\begin{aligned}
\mu_{\omega t} = &\left[\frac{\gamma}{\omega_t} - (1-2\alpha)(1-\gamma)\frac{1}{1+\omega_t} + (1-2\alpha)(1-\gamma)\frac{(1-\alpha)^2}{\alpha^2+(1-\alpha)^2\omega_t}\right]^{-1} \\
&\times\Bigg[(1-2\alpha)(1-\gamma)(\mu_{Bt}-\mu_{At}) + \frac{1}{2}\frac{\gamma}{\omega_t^2}\left({\sigma^\omega_{At}}^2 + {\sigma^\omega_{Bt}}^2 + 2\rho\sigma^\omega_{At}\sigma^\omega_{Bt} + \boldsymbol{\sigma^{\omega}_{\mu t}}^\top\boldsymbol{\sigma^{\omega}_{\mu t}}\right) \\
&\quad -\frac{1}{2}(1-2\alpha)(1-\gamma)(\sigma_B^2-\sigma_A^2) - \frac{1}{2}\frac{(1-2\alpha)(1-\gamma)}{(1+\omega_t)^2}\left({\sigma^\omega_{At}}^2 + {\sigma^\omega_{Bt}}^2 + 2\rho\sigma^\omega_{At}\sigma^\omega_{Bt} + \boldsymbol{\sigma^{\omega}_{\mu t}}^\top\boldsymbol{\sigma^{\omega}_{\mu t}}\right) \\
&+\frac{1}{2}(1-2\alpha)(1-\gamma)\frac{(1-\alpha)^4}{(\alpha^2+(1-\alpha)^2\omega_t)^2}\left({\sigma^\omega_{At}}^2 + {\sigma^\omega_{Bt}}^2 + 2\rho\sigma^\omega_{At}\sigma^\omega_{Bt} + \boldsymbol{\sigma^{\omega}_{\mu t}}^\top\boldsymbol{\sigma^{\omega}_{\mu t}}\right)\Bigg]
\end{aligned}$$

and the volatilities as

$$\begin{aligned}
\sigma^\omega_{At} &= \left[-\frac{\gamma}{\omega_t} + (1-2\alpha)(1-\gamma)\frac{1}{1+\omega_t} - (1-2\alpha)(1-\gamma)\frac{(1-\alpha)^2}{\alpha^2+(1-\alpha)^2\omega_t}\right]^{-1}(1-2\alpha)(1-\gamma)\sigma_A \\
\sigma^\omega_{Bt} &= \left[-\frac{\gamma}{\omega_t} + (1-2\alpha)(1-\gamma)\frac{1}{1+\omega_t} - (1-2\alpha)(1-\gamma)\frac{(1-\alpha)^2}{\alpha^2+(1-\alpha)^2\omega_t}\right]^{-1}(1-2\alpha)(\gamma-1)\sigma_B \\
\boldsymbol{\sigma^{\omega}_{\mu t}} &= \left[-\frac{\gamma}{\omega_t} + (1-2\alpha)(1-\gamma)\frac{1}{1+\omega_t} - (1-2\alpha)(1-\gamma)\frac{(1-\alpha)^2}{\alpha^2+(1-\alpha)^2\omega_t}\right]^{-1}\left(\boldsymbol{\xi_{At}} - \boldsymbol{\xi_{Bt}}\right).
\end{aligned}$$

$\square$

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

$$
\begin{aligned}
\mu_{C_A t} = & \frac{\mu_{\omega t}}{\omega_t} - \frac{1}{2\omega_t^2}\left({\sigma_{At}^{\omega}}^2 + {\sigma_{Bt}^{\omega}}^2 + 2\rho\sigma_{At}^{\omega}\sigma_{Bt}^{\omega} + {\boldsymbol{\sigma_{\mu t}^{\omega}}}^\top \boldsymbol{\sigma_{\mu t}^{\omega}}\right) + \alpha\mu_{At} - \frac{1}{2}\alpha\sigma_A^2 + (1-\alpha)\mu_{Bt} \\
& - \frac{1}{2}(1-\alpha)\sigma_B^2 - \frac{1}{1+\omega_t}\alpha\mu_{\omega t} + \frac{1}{2}\alpha\frac{1}{(1+\omega_t)^2}\left({\sigma_{At}^{\omega}}^2 + {\sigma_{Bt}^{\omega}}^2 + 2\rho\sigma_{At}^{\omega}\sigma_{Bt}^{\omega} + {\boldsymbol{\sigma_{\mu t}^{\omega}}}^\top \boldsymbol{\sigma_{\mu t}^{\omega}}\right) \\
& - \frac{(1-\alpha)^3}{\alpha^2 + (1-\alpha)^2\omega_t}\mu_{\omega t} + \frac{1}{2}\frac{(1-\alpha)^5}{(\alpha^2 + (1-\alpha)^2\omega_t)^2}\left({\sigma_{At}^{\omega}}^2 + {\sigma_{Bt}^{\omega}}^2 + 2\rho\sigma_{At}^{\omega}\sigma_{Bt}^{\omega} + {\boldsymbol{\sigma_{\mu t}^{\omega}}}^\top \boldsymbol{\sigma_{\mu t}^{\omega}}\right) \\
& + \frac{1}{2}\left({\sigma_A^{C_A}}^2 + {\sigma_B^{C_A}}^2 + 2\rho\sigma_A^{C_A}\sigma_B^{C_A} + {\boldsymbol{\sigma_{\mu t}^{C_A}}}^\top \boldsymbol{\sigma_{\mu t}^{C_A}}\right),
\end{aligned}
$$

and for C_{Bt} as

$$
\begin{aligned}
\mu_{C_B t} = & (1-\alpha)\mu_{At} - \frac{1}{2}(1-\alpha)\sigma_A^2 + \alpha\mu_{Bt} - \frac{1}{2}\alpha\sigma_B^2 \\
& - \frac{1}{1+\omega_t}(1-\alpha)\mu_{\omega t} + \frac{1}{2}(1-\alpha)\frac{1}{(1+\omega_t)^2}\left({\sigma_{At}^{\omega}}^2 + {\sigma_{Bt}^{\omega}}^2 + 2\rho\sigma_{At}^{\omega}\sigma_{Bt}^{\omega} + {\boldsymbol{\sigma_{\mu t}^{\omega}}}^\top \boldsymbol{\sigma_{\mu t}^{\omega}}\right) \\
& - \frac{\alpha(1-\alpha)^2}{\alpha^2 + (1-\alpha)^2\omega_t}\mu_{\omega t} + \frac{1}{2}\frac{\alpha(1-\alpha)^4}{(\alpha^2 + (1-\alpha)^2\omega_t)^2}\left({\sigma_{At}^{\omega}}^2 + {\sigma_{Bt}^{\omega}}^2 + 2\rho\sigma_{At}^{\omega}\sigma_{Bt}^{\omega} + {\boldsymbol{\sigma_{\mu t}^{\omega}}}^\top \boldsymbol{\sigma_{\mu t}^{\omega}}\right) \\
& + \frac{1}{2}\left({\sigma_A^{C_B}}^2 + {\sigma_B^{C_B}}^2 + 2\rho\sigma_A^{C_B}\sigma_B^{C_B} + {\boldsymbol{\sigma_{\mu t}^{C_B}}}^\top \boldsymbol{\sigma_{\mu t}^{C_B}}\right).
\end{aligned}
$$

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

$$
\begin{aligned}
\sigma_{At}^{C_A} &= \alpha\sigma_A + \frac{1}{\omega_t}\sigma_{At}^{\omega} - \frac{1}{1+\omega_t}\alpha\sigma_{At}^{\omega} - \frac{(1-\alpha)^3}{\alpha^2 + (1-\alpha)^2\omega_t}\sigma_{At}^{\omega} \\
\sigma_{Bt}^{C_A} &= (1-\alpha)\sigma_B + \frac{1}{\omega_t}\sigma_{Bt}^{\omega} - \frac{1}{1+\omega_t}\alpha\sigma_{Bt}^{\omega} - \frac{(1-\alpha)^3}{\alpha^2 + (1-\alpha)^2\omega_t}\sigma_{Bt}^{\omega} \\
\boldsymbol{\sigma_{\mu t}^{C_A}} &= \frac{1}{\omega_t}\boldsymbol{\sigma_{\mu t}^{\omega}} - \frac{1}{1+\omega_t}\alpha\boldsymbol{\sigma_{\mu t}^{\omega}} - \frac{(1-\alpha)^3}{\alpha^2 + (1-\alpha)^2\omega_t}\boldsymbol{\sigma_{\mu t}^{\omega}} \\
\sigma_{At}^{C_B} &= (1-\alpha)\sigma_A - \frac{1}{1+\omega_t}(1-\alpha)\sigma_{At}^{\omega} - \frac{\alpha(1-\alpha)^2}{\alpha^2 + (1-\alpha)^2\omega_t}\sigma_{At}^{\omega} \\
\sigma_{Bt}^{C_B} &= \alpha\sigma_B - \frac{1}{1+\omega_t}(1-\alpha)\sigma_{Bt}^{\omega} - \frac{\alpha(1-\alpha)^2}{\alpha^2 + (1-\alpha)^2\omega_t}\sigma_{Bt}^{\omega} \\
\boldsymbol{\sigma_{\mu t}^{C_B}} &= -\frac{1}{1+\omega_t}(1-\alpha)\boldsymbol{\sigma_{\mu t}^{\omega}} - \frac{\alpha(1-\alpha)^2}{\alpha^2 + (1-\alpha)^2\omega_t}\boldsymbol{\sigma_{\mu t}^{\omega}}.
\end{aligned}
$$

□

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 $M^{\mathcal{E}}_{At}$ as

$$\begin{aligned}
\mu_{M_A t} &= -\beta - \gamma\mu_{C_A t} - \frac{1}{2}\boldsymbol{\xi}_{\boldsymbol{At}}^\top \boldsymbol{\xi}_{\boldsymbol{At}} + \frac{1}{2}\gamma\left({\sigma_A^{C_A}}^2 + {\sigma_B^{C_A}}^2 + 2\rho\sigma_A^{C_A}\sigma_B^{C_A} + {\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{C_A}}}^\top \boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{C_A}}\right) \\
&\quad + \frac{1}{2}\left({\sigma_A^{M_A}}^2 + {\sigma_{Bt}^{M_A}}^2 + 2\rho\sigma_A^{M_A}\sigma_B^{M_A} + {\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{M_A}}}^\top \boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{M_A}}\right) \\
\sigma_A^{M_A} &= \gamma\sigma_A^{C_A} \\
\sigma_B^{M_A} &= \gamma\sigma_B^{C_A} \\
\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{M_A}} &= \boldsymbol{\xi}_{\boldsymbol{At}} + \gamma\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{C_A}}.
\end{aligned}$$

The drift and volatilities for $M^{\mathcal{E}}_{Bt}$ are

$$\begin{aligned}
\mu_{M_B t} &= -\beta - \gamma\mu_{C_B t} - \frac{1}{2}\boldsymbol{\xi}_{\boldsymbol{Bt}}^\top \boldsymbol{\xi}_{\boldsymbol{Bt}} + \frac{1}{2}\gamma\left({\sigma_A^{C_B}}^2 + {\sigma_B^{C_B}}^2 + 2\rho\sigma_A^{C_B}\sigma_B^{C_B} + {\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{C_B}}}^\top \boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{C_B}}\right) \\
&\quad + \frac{1}{2}\left({\sigma_A^{M_B}}^2 + {\sigma_B^{M_B}}^2 + 2\rho\sigma_A^{M_B}\sigma_B^{M_B} + {\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{M_B}}}^\top \boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{M_B}}\right) \\
\sigma_A^{M_B} &= \gamma\sigma_A^{C_B} \\
\sigma_B^{M_B} &= \gamma\sigma_B^{C_B} \\
\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{M_B}} &= \boldsymbol{\xi}_{\boldsymbol{Bt}} + \gamma\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{C_B}}.
\end{aligned}$$

□

Proof of Proposition 12

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

$$\begin{aligned}
\mu_{St} &= \gamma\mu_{C_A t} - \gamma\mu_{C_B t} - \frac{1}{2}\gamma\left({\sigma_A^{C_A}}^2 + {\sigma_B^{C_A}}^2 + 2\rho\sigma_A^{C_A}\sigma_B^{C_A} + {\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{C_A}}}^\top \boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{C_A}}\right) \\
&\quad + \frac{1}{2}\gamma\left({\sigma_A^{C_B}}^2 + {\sigma_B^{C_B}}^2 + 2\rho\sigma_A^{C_B}\sigma_B^{C_B} + {\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{C_B}}}^\top \boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{C_B}}\right) + \frac{1}{2}\left({\sigma_A^S}^2 + {\sigma_B^S}^2 + 2\rho\sigma_A^S\sigma_B^S + {\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{S}}}^\top \boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{S}}\right) \\
\sigma_A^S &= \gamma\sigma_A^{C_A} - \gamma\sigma_A^{C_B} \\
\sigma_B^S &= \gamma\sigma_B^{C_A} - \gamma\sigma_B^{C_B} \\
\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{S}} &= \boldsymbol{\xi}_{\boldsymbol{At}} - \boldsymbol{\xi}_{\boldsymbol{Bt}} + \gamma\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{C_A}} - \gamma\boldsymbol{\sigma}_{\boldsymbol{\mu t}}^{\boldsymbol{C_B}}.
\end{aligned}$$

□

Proof of Proposition 13

Proof. I first note that when T is fixed $P_t^Y M_{At}^{\mathcal{E}}$ is an $\mathcal{E}$-martingale. Applying Ito's lemma to the both sides of (47), we get the following for the LHS

$$P_t^Y M_{At}^{\mathcal{E}} \left(\boldsymbol{\sigma}_{\boldsymbol{P^Y}}\left(t, P_t^Y\right) - \boldsymbol{\sigma}_{\boldsymbol{M_A t}} M_{At}^{\mathcal{E}}\right)^\top \mathbf{dW_t^{\mathcal{E}}}.$$

I then apply Theorem 2 to the RHS of (47) and obtain the RHS as

$$\mathbb{E}_t^{\mathcal{E}}\left[\mathrm{D}_t\left(M_{AT}^{\mathcal{E}} Y_T\right)\right]^\top \mathbf{dW_t}^{\mathcal{E}},$$

where D_t denotes the Malliavin derivative operator defined in Appendix G.

To compute $\mathrm{D}_t\left(M_{AT}^{\mathcal{E}} Y_T\right)$, note that

$$\begin{aligned}
\mathrm{D}_t\left(M_{AT}^{\mathcal{E}} Y_T\right) &= M_{AT}^{\mathcal{E}} Y_T\left(\mathrm{D}_t \log M_{AT}^{\mathcal{E}} + \mathrm{D}_t \log Y_T\right)\\
&= M_{AT}^{\mathcal{E}} Y_T\left(\frac{1}{M_{AT}^{\mathcal{E}}}\mathrm{D}_t M_{AT}^{\mathcal{E}} + \frac{1}{Y_T}\mathrm{D}_t Y_T\right).
\end{aligned}$$

The first variation processes for $M_{At}^{\mathcal{E}}$ and Y_t are given by

$$\begin{aligned}
d\widetilde{M}_{At}^{\mathcal{E}} &= \mu_{M_A t}\widetilde{M}_{At}^{\mathcal{E}} dt - \boldsymbol{\sigma}_{M_A t}^\top \widetilde{M}_{At}^{\mathcal{E}} \mathbf{dW_t^{\mathcal{E}}}\\
d\widetilde{Y}_t &= \mu_Y'\left(t, Y_t\right)\widetilde{Y}_t dt + \boldsymbol{\sigma}_{\boldsymbol{Y}}'\left(t, Y_t\right)^\top \widetilde{Y}_t \mathbf{dW_t^{\mathcal{E}}}.
\end{aligned}$$

Therefore, the Malliavin derivatives are

$$\begin{aligned}
\mathrm{D}_t M_{AT}^{\mathcal{E}} &= -M_{AT}^{\mathcal{E}} \boldsymbol{\sigma}_{\boldsymbol{M_A t}}\\
\mathrm{D}_t Y_T &= \frac{\widetilde{Y}_T}{\widetilde{Y}_t}\boldsymbol{\sigma}_{\boldsymbol{Y}}\left(t, Y_t\right).
\end{aligned}$$

The volatility of price P_t^Y is then

$$\begin{aligned}
\boldsymbol{\sigma}_{\boldsymbol{P^Y}}\left(t, P_t^Y\right) &= \boldsymbol{\sigma}_{\boldsymbol{M_A t}} M_{At}^{\mathcal{E}} + \frac{\mathbb{E}_t^{\mathcal{E}}\left[\mathrm{D}_t(M_{AT}^{\mathcal{E}} Y_T)\right]}{P_t^Y M_{At}^{\mathcal{E}}}\\
&= \boldsymbol{\sigma}_{\boldsymbol{M_A t}} M_{At}^{\mathcal{E}} + \frac{\mathbb{E}_t^{\mathcal{E}}\left[M_{AT}^{\mathcal{E}} Y_T \mathrm{D}_t(\log M_{AT}^{\mathcal{E}} + \log Y_T)\right]}{P_t^Y M_{At}^{\mathcal{E}}}\\
&= \boldsymbol{\sigma}_{\boldsymbol{M_A t}} M_{At}^{\mathcal{E}} + \frac{1}{P_t^Y M_{At}^{\mathcal{E}}}\mathbb{E}_t^{\mathcal{E}}\left[M_{AT}^{\mathcal{E}} Y_T\left(\frac{\widetilde{Y}_T}{Y_T\widetilde{Y}_t}\boldsymbol{\sigma}_{\boldsymbol{Y}}\left(t, Y_t\right) - \boldsymbol{\sigma}_{\boldsymbol{M_A t}}\right)\right].
\end{aligned}$$

The drift of the price process has the following usual form as a result of zero drift on LHS of (47)

$$\mu_{P^Y}\left(t, P_t^Y\right) = r_{At} P_t^Y + \boldsymbol{\sigma}_{\boldsymbol{M_A t}}^\top \Omega_W \boldsymbol{\sigma}_{\boldsymbol{P^Y}}\left(t, P_t^Y\right).$$

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 (49), we get the diffusion part as

$$E_t^A M_{At}^{\mathcal{E}} \left(\boldsymbol{\sigma}_{\boldsymbol{E^A t}} - \boldsymbol{\sigma}_{\boldsymbol{M_A t}}\right)^\top \mathbf{dW_t^{\mathcal{E}}}. \tag{57}$$

By Theorem 2, the diffusion part of RHS of (49) is

$$\mathbb{E}_t^{\mathcal{E}} \left[\int_t^\infty \mathrm{D}_t \left(M_{As}^{\mathcal{E}} q_{As} X_{As}\right) ds\right], \tag{58}$$

where we know from Appendix B that

$$M_{At}^{\mathcal{E}} q_{At} X_{At} = e^{-\beta t} \left(\frac{\eta_{At} \alpha C_{At}^{1-\gamma}}{\nu_A} + \frac{\eta_{Bt}(1-\alpha) C_{Bt}^{1-\gamma}}{\nu_B}\right),$$

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

$$M_{At}^{\mathcal{E}} q_{At} X_{At} = e^{-\beta t} \left(\frac{\eta_{At} \alpha}{\nu_A} + \frac{\eta_{Bt}(1-\alpha)}{\nu_B}\right).$$

Let η_{it} follow

$$\frac{d\eta_{it}}{\eta_{it}} = \boldsymbol{\sigma}_{\boldsymbol{\eta_i t}}^\top \mathbf{dW_t}^{\mathcal{E}}$$

where $\boldsymbol{\sigma}_{\boldsymbol{\eta_A t}} = (0, 0, -\boldsymbol{\xi}_{\boldsymbol{At}}^\top)^\top, \boldsymbol{\sigma}_{\boldsymbol{\eta_B t}} = (0, 0, -\boldsymbol{\xi}_{\boldsymbol{Bt}}^\top)^\top$. The first variation processes for η_{it} and C_{it} are

$$\begin{aligned} d\widetilde{\eta}_{it} &= \boldsymbol{\sigma}_{\boldsymbol{\eta_i t}}^\top \widetilde{\eta}_{it} \mathbf{dW_t}^{\mathcal{E}} \\ d\widetilde{C}_{it} &= \mu_{C_i t} \widetilde{C}_{it} dt + \boldsymbol{\sigma}_{\boldsymbol{C_i t}}^\top \widetilde{C}_{it} \mathbf{dW_t}^{\mathcal{E}}, \end{aligned}$$

or

$$\begin{aligned} \widetilde{\eta}_{it} &= \exp\left[\int_0^t \left(-\frac{1}{2} \boldsymbol{\sigma}_{\boldsymbol{\eta_i s}}^\top \boldsymbol{\sigma}_{\boldsymbol{\eta_i s}}\right) ds + \int_0^t \boldsymbol{\sigma}_{\boldsymbol{\eta_A s}}^\top \mathbf{dW_s^{\mathcal{E}}}\right] \\ \widetilde{C}_{it} &= \exp\left[\int_0^t \left(\mu_{C_i s} - \frac{1}{2} \boldsymbol{\sigma}_{\boldsymbol{C_i s}}^\top \boldsymbol{\sigma}_{\boldsymbol{C_i s}}\right) ds + \int_0^t \boldsymbol{\sigma}_{\boldsymbol{C_i s}}^\top \mathbf{dW_s^{\mathcal{E}}}\right]. \end{aligned}$$

When $\gamma \neq 1$, we have for $s \geq t$

$$\mathrm{D}_t \left(\eta_{As} C_{As}^{1-\gamma}\right) = C_{As}^{1-\gamma} \mathrm{D}_t \eta_{As} + (1-\gamma) \eta_{As} C_{As}^{-\gamma} \mathrm{D}_t C_s,$$

where

$$\mathrm{D}_t \eta_{As} = \eta_{As} \boldsymbol{\sigma_{\eta_A t}}$$
$$\mathrm{D}_t C_{As} = C_{As} \boldsymbol{\sigma_{C_A t}}.$$

The expression for $\mathrm{D}_t \left(\eta_{Bs} C_{Bs}^{1-\gamma}\right)$ can be obtained similarly. Substituting these expressions in (58) and equating (57) and (58), we obtain desired results. In particular, we have

$$\begin{aligned}\mathrm{D}_t \left(M_{As}^{\mathcal{E}} q_{As} X_{As}\right) = e^{-\beta s} \big[& \lambda_A \alpha \left(\eta_{As} C_{As}^{1-\gamma} \boldsymbol{\sigma_{\eta_A t}} + (1-\gamma) \eta_{As} C_{As}^{1-\gamma} \boldsymbol{\sigma_{C_A t}}\right) \\ & + \lambda_B (1-\alpha) \left(\eta_{Bs} C_{Bs}^{1-\gamma} \boldsymbol{\sigma_{\eta_B t}} + (1-\gamma) \eta_{Bs} C_{Bs}^{1-\gamma} \boldsymbol{\sigma_{C_B t}}\right)\big].\end{aligned}$$

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