A Gravity-Based Theory of Bilateral Portfolio Choice

Lei Pan[*†]

Abstract

This paper develops a unified theory of bilateral portfolio similarity grounded in gravity-type information frictions and exchange-rate stabilization. We construct a Bayesian portfolio-choice framework in which investors from countries i and j allocate wealth to a common set of risky assets under heterogeneous signal structures and bilateral currency regimes. The model shows that portfolio dissimilarity admits a closed-form gravity representation, increasing in bilateral distance and decreasing in peg strength. We extend the analysis to a multivariate asset environment, a dynamic Bayesian learning structure with autoregressive signals, and an endogenous exchange-rate regime in which governments optimally choose Λ_{ij} to trade off monetary autonomy against risk-sharing gains. The theory rationalizes persistent cross-country heterogeneity in foreign-asset shares and provides welfare-based predictions for bilateral monetary integration. Policy implications follow for regional currency arrangements, information infrastructure, and financial-integration strategies.

JEL classification: F30, F31, G11, E44

Keywords: Information frictions; Exchange-rate stabilization; Gravity models; Bayesian learning; Bilateral portfolio similarity

[*]School of Accoutning, Economics and Finance, Curtin University, Australia. Email: lei.pan@curtin.edu.au
[†]Centre for Development Economics and Sustainability, Monash University, Australia.

1 Introduction

Understanding why investors located in different countries choose systematically different foreign portfolio allocations—despite having access to the same global assets—remains a central question in international macroeconomics and asset-pricing theory. Even in highly integrated financial markets, bilateral differences in foreign portfolio investment (FPI) patterns remain large, persistent, and structured along observable country-pair characteristics such as exchange-rate arrangements, geographic distance, and linguistic or historical proximity. These empirical regularities, documented prominently by Pan, Hu and Du (2022) and a growing literature on gravity in financial flows, call for a theoretical framework capable of jointly rationalising (i) the effect of bilateral monetary integration on portfolio choices and (ii) the systematic role of gravity-type information frictions in shaping cross-country heterogeneity in foreign asset demands.

While there exist rich literatures separately studying exchange-rate regimes (e.g., Obstfeld and Rogoff, 1995; Devereux and Engel, 2002), home bias and information asymmetry (e.g., Gehrig, 1993; Brennan and Cao, 1997; Van Nieuwerburgh and Veldkamp, 2009), and financial-gravity patterns (e.g., Portes and Rey, 2005; Okawa and van Wincoop, 2012), there is no unified theory demonstrating how bilateral exchange-rate stabilization and gravity-driven information frictions interact to determine portfolio similarity across country pairs. Existing models either take monetary regimes as exogenous and independent of private-sector portfolio behaviour, or they treat information frictions as agent-specific rather than bilateral objects that vary systematically across country pairs.

This paper develops a multi-country, gravity-consistent portfolio-choice theory that bridges these gaps. We build a two-investor, one-destination model in which i and j allocate wealth to an internationally traded asset from country k. Two central mechanisms govern portfolio similarity: First, a risk-reduction channel, through which tighter bilateral exchange-rate pegs reduce idiosyncratic exchange-rate risk and therefore align optimal foreign-asset shares. Second, an information channel, whereby distance, common language, and historical ties influence the quality and correlation of investors' signals about foreign returns, generating systematic bilateral heterogeneity in posterior beliefs and hence portfolio choices.

Our analysis contributes to three strands of theory. Specifically, i) *International portfolio choice and risk sharing*: The risk-sharing literature highlights how monetary and financial integration affects cross-border diversification (e.g., Coeurdacier and Gourinchas, 2016; Martin and Rey, 2004). Models of exchange-rate stabilization—such as Engel (2014) and Gopinath and Itskhoki (2011)—show how regime choices influence macro-financial risk. However, these frameworks do not map bilateral regime strength into pairwise portfolio similarity, nor do they incorporate bilateral information frictions. ii) *Information frictions and Bayesian learning in asset allocation*: A large literature emphasizes how information asymmetry shapes portfolios (e.g., Van Nieuwerburgh and Veldkamp, 2009). Yet these models treat information frictions as investor-level, not country-pair-level, and do not generate predictions linking distance, language, or colonial ties to cross-country similarity of foreign-asset shares. iii) *Gravity in financial flows*: Gravity-type determinants are well established empirically (e.g., Portes and Rey, 2005; Lane and Milesi-Ferretti, 2008; Okawa and van Wincoop, 2012; Forbes and Warnock, 2012). Recent work applies gravity to cross-border banking or FPI flows but provides no micro-foundation for how gravity generates predictable differences in optimal portfolio shares. This paper closes this gap by deriving a bilateral gravity-like equation directly from first principles.

This paper makes three main contributions. First, we provide the first analytical model in which bilateral exchange-rate regimes and gravity-type information frictions jointly determine the difference in optimal foreign-asset shares chosen by two countries. The baseline model yields closed-form expressions for portfolio dissimilarity, showing that: i) portfolio similarity increases with peg strength, ii) dissimilarity rises with bilateral distance, and iii) common language and colonial ties reduce dispersion through information and signal-correlation channels. This provides structural theoretical foundations for recent empirical work linking exchange-rate arrangements to FPI similarity. Second, we extend the static model to a multi-asset environment and a Kalman-filter learning setting where investors receive autoregressive signals. We derive the long-run variance of portfolio differences, showing that the comparative statics of peg strength and gravity variables remain intact in dynamic steady state. We also analyse how return covariance matrices under different regime arrangements influence the Euclidean distance between multi-asset portfolio vectors. Third, we endogenize the peg Λ_{ij} by allowing governments to optimally choose regime strength to balance risk-sharing benefits (through lower portfolio dispersion) against monetary-autonomy costs. The model yields a closed-form optimal peg, predicts when countries choose deeper integration, and shows how optimal peg strength responds to distance, information quality, and investor learning. This contributes to the theory of bilateral monetary arrangements and provides new policy implications for regional financial integration.

Overall, these contributions generate a unified theory explaining why some bilateral pairs exhibit remarkably similar foreign portfolio allocations while others remain sharply different, even when exposed to the same global asset menu. The model is tractable, yields gravity-consistent reduced forms, and offers a theoretical foundation for empirical bilateral regressions of FPI similarity.

The remainder of the paper is organized as follows. Section 2 develops the baseline theoretical framework and derives the gravity-consistent expression for bilateral portfolio dissimilarity, highlighting the roles of exchange-rate stabilization and information frictions. Section 3 presents several extensions of the model, including the multivariate portfolio environment, dynamic Bayesian learning with autoregressive signals, and the welfare-based determination of optimal bilateral peg strength. Section 4 concludes.

2 A Multi-Country Model of Portfolio Similarity

This section develops a multi-country portfolio-choice model that rationalises: i) why some country pairs choose more similar foreign portfolio investment (FPI) patterns than others, and ii) why portfolio similarity is systematically related to bilateral exchange rate regimes and gravity-type frictions (distance, common language, etc.), as in the empirical evidence of Pan *et al.* (2022). The theoretical object that we match to the empirical dependent variable is the absolute difference in portfolio shares of a benchmark foreign asset (e.g. US assets) chosen by investors in two countries i and j.

2.1 Set up Time is discrete, indexed by $t = 0, 1$. At $t = 0$ investors choose portfolios; at $t = 1$ uncertainty is resolved and consumption takes place. There are three countries: two "investor" countries i and j and one benchmark destination country k (e.g. the US) whose risky asset is held by both i and j. All variables are in real terms unless otherwise noted.

Each investor country $h \in \{i, j\}$ is populated by a representative infinitely risk-averse household of mass one.[1] Household h is endowed at $t = 0$ with initial wealth $W_h > 0$ denominated in its own currency.

Household h has time-separable expected utility:

$$U_h = \mathbb{E}_0 \left[\frac{C_{h,1}^{1-\gamma}}{1-\gamma} \right], \qquad \gamma > 0,\ \gamma \neq 1, \tag{1}$$

where $C_{h,1}$ is real consumption at $t = 1$ and γ is the coefficient of relative risk aversion.

In each investor country h there is a risk-free real bond with gross return $R_{f,h} > 0$ between $t = 0$ and $t = 1$, denominated in country h's currency. There is also one internationally traded risky asset issued in country k (for instance, US equity) with payoff $\tilde{R}_k$ in k's currency.

Let $S_{hk,t}$ denote the nominal exchange rate defined as units of country h's currency per unit of country k's currency at time t. Then the gross real return in units of country h's consumption good from holding the risky asset is:

$$R_k^{(h)} = \tilde{R}_k \frac{S_{hk,1}}{S_{hk,0}} \frac{P_{k,0}}{P_{h,1}}, \tag{2}$$

where $P_{h,t}$ denotes the price level in country h at time t. For tractability, we assume that log returns are jointly normal and that $R_k^{(h)}$ can be expressed as:

$$R_k^{(h)} = R_{f,h} + \tilde{r} + \varepsilon_h, \qquad h \in \{i, j\}, \tag{3}$$

where $\tilde{r}$ is a common global excess-return component (in expectation) and ε_h is a country-specific

[1] The extension to multiple heterogeneous investors inside each country is straightforward and does not affect the comparative statics derived below.

excess-return shock capturing exchange rate and local macro-financial risk.[2]

We denote the perceived conditional mean and variance of the risky excess return for household h by:

$$\mu_h := \mathbb{E}_0\left[R_k^{(h)} - R_{f,h}\right], \qquad \sigma_h^2 := \mathrm{Var}_0\left(R_k^{(h)} - R_{f,h}\right).$$

Exchange rate regime. Following the empirical work, we model the bilateral exchange rate arrangement between countries i and j by a parameter $\Lambda_{ij} \in [0,1]$. When $\Lambda_{ij} = 0$ the two currencies float independently; when $\Lambda_{ij} = 1$ they are perfectly pegged (a monetary union or hard peg).

Specifically, decompose the country-specific shock as:

$$\varepsilon_h = \eta + (1 - \Lambda_{ij})\nu_h, \qquad h \in \{i, j\},$$

where η is a common shock and ν_h is an idiosyncratic shock with $\mathbb{E}[\eta] = \mathbb{E}[\nu_h] = 0$, $\mathrm{Var}(\eta) = \sigma_\eta^2$ and $\mathrm{Var}(\nu_h) = \sigma_\nu^2$, independent across h and independent of η. Then the variance of the risky excess return in country h is:

$$\sigma_h^2 = \sigma_\eta^2 + (1 - \Lambda_{ij})^2\sigma_\nu^2. \tag{4}$$

A tighter peg (higher Λ_{ij}) reduces the idiosyncratic component of risk and makes the risky return processes across i and j more similar.

2.2 Gravity-type information frictions

To incorporate gravity variables, we assume investors are uncertain about the mean excess return $\tilde{r}$ and need to learn it from noisy country-specific signals. Each investor h observes at $t = 0$ a signal:

$$x_h = \tilde{r} + u_h, \tag{5}$$

where $u_h \sim \mathcal{N}(0, \sigma_{u,h}^2)$ is independent noise.

Let the prior for $\tilde{r}$ be normal: $\tilde{r} \sim \mathcal{N}(\bar{r}, \sigma_{\bar{r}}^2)$. Conditioning on x_h, the posterior distribution

[2] A more structural derivation of Equation (3) from a DSGE model with monetary policy and nominal rigidities is straightforward but not necessary for our portfolio comparative statics.

perceived by investor h has mean:

$$\hat{r}_h := \mathbb{E}[\tilde{r} \mid x_h] = \omega_h x_h + (1-\omega_h)\bar{r}, \qquad \omega_h := \frac{\sigma_{\bar{r}}^2}{\sigma_{\bar{r}}^2 + \sigma_{u,h}^2}, \tag{6}$$

and variance:

$$\sigma_{\hat{r},h}^2 := \mathrm{Var}(\tilde{r} \mid x_h) = \frac{\sigma_{\bar{r}}^2 \sigma_{u,h}^2}{\sigma_{\bar{r}}^2 + \sigma_{u,h}^2}. \tag{7}$$

We interpret the signal noise variance $\sigma_{u,h}^2$ as increasing in bilateral information frictions faced by investor h when learning about country k's assets. Specifically, for a given country pair (i, j) we posit:

$$\sigma_{u,h}^2 = \sigma_u^2\Big(1 + \phi_d d_{ij} - \phi_\ell L_{ij} + \phi_c C_{ij}\Big), \qquad h \in \{i, j\}, \tag{8}$$

where d_{ij} is (log) bilateral distance between i and j, $L_{ij} \in \{0, 1\}$ is a common official language dummy, $C_{ij} \in \{0, 1\}$ is a common colonizer dummy, and $\phi_d, \phi_\ell, \phi_c \geq 0$ are sensitivity parameters. Larger distance increases information noise, whereas common language or a common colonial history reduce it (by facilitating communication, legal similarity, etc.).

Since both i and j learn about the same underlying $\tilde{r}$, their posterior means are correlated. For convenience, we suppose that the signal noises u_i and u_j have correlation $\rho_u \in [0, 1)$, which can itself be increasing in information connectivity between the two countries (shared media, migration links, etc.). A higher ρ_u means investors in i and j react more similarly to news about country k.

2.3 Budget constraint and portfolio choice

Let α_h denote the share of wealth invested in the risky foreign asset by investor h. The remainder $(1 - \alpha_h)$ is invested in the local risk-free bond. The $t = 1$ budget constraint in real terms is:

$$C_{h,1} = W_h\big[(1-\alpha_h)R_{f,h} + \alpha_h R_k^{(h)}\big]. \tag{9}$$

Substituting Equation (3) into Equation (9) and defining the excess return $\tilde{R}_h := R_k^{(h)} - R_{f,h}$

gives:

$$C_{h,1} = W_h R_{f,h} \big[1 + \alpha_h \tilde{R}_h\big].$$

Under CRRA preferences with small risk and jointly normal returns, the portfolio problem is equivalent to a mean–variance maximisation (e.g. standard Merton approximation). Investor h chooses α_h to maximise:

$$\mathbb{E}_0[C_{h,1}^{1-\gamma}] \approx \Big(W_h R_{f,h}\Big)^{1-\gamma} \left\{1 + (1-\gamma)\alpha_h \mu_h - \frac{\gamma(1-\gamma)}{2}\alpha_h^2 \sigma_h^2\right\}, \tag{10}$$

where $\mu_h := \mathbb{E}_0[\tilde{R}_h]$ and $\sigma_h^2 := \mathrm{Var}_0(\tilde{R}_h)$ are computed using the posterior beliefs Equation (6)–Equation (7).

Lemma 1 (Optimal risky portfolio share)**.** *For $h \in \{i, j\}$, the unique optimal share of wealth invested in the foreign risky asset is:*

$$\alpha_h^* = \frac{\mu_h}{\gamma \sigma_h^2}. \tag{11}$$

Proof. The portfolio problem for investor h is to choose the share of wealth α_h invested in the risky foreign asset. The investor maximizes the following approximated expected utility, based on the mean-variance framework:

$$\max_{\alpha_h} \quad \mathbb{E}_0[C_{h,1}^{1-\gamma}] \approx (W_h R_{f,h})^{1-\gamma} \left\{1 + (1-\gamma)\alpha_h \mu_h - \frac{\gamma(1-\gamma)}{2}\alpha_h^2 \sigma_h^2\right\}$$

Since terms independent of α_h do not affect the maximization, the problem is equivalent to maximizing the objective function $f(\alpha_h)$:

$$\max_{\alpha_h} \quad f(\alpha_h) = (1-\gamma)\alpha_h \mu_h - \frac{\gamma(1-\gamma)}{2}\alpha_h^2 \sigma_h^2$$

First-order condition (FOC): To find the optimal share α_h^*, we take the derivative of the objective function with respect to α_h and set it to zero:

$$\frac{df(\alpha_h)}{d\alpha_h} = 0$$

$$\frac{d}{d\alpha_h}\left[(1-\gamma)\alpha_h\mu_h - \frac{\gamma(1-\gamma)}{2}\alpha_h^2\sigma_h^2\right] = 0$$

Applying the power rule, $\frac{d}{dx}(ax^n) = nax^{n-1}$, we get:

$$(1-\gamma)\mu_h - \frac{\gamma(1-\gamma)}{2}(2\alpha_h)\sigma_h^2 = 0$$

$$(1-\gamma)\mu_h - \gamma(1-\gamma)\alpha_h^*\sigma_h^2 = 0$$

Solving for α_h^*: We rearrange the FOC to solve for α_h^*. Assuming $\gamma \neq 1$, we can divide by $(1-\gamma)$:

$$(1-\gamma)\mu_h = \gamma(1-\gamma)\alpha_h^*\sigma_h^2$$

$$\mu_h = \gamma\alpha_h^*\sigma_h^2$$

Isolating α_h^* yields the stated result:

$$\alpha_h^* = \frac{\mu_h}{\gamma\sigma_h^2}$$

Second-order condition (SOC) and uniqueness: The objective function is strictly concave, which guarantees that the solution is a unique maximizer. This is verified by checking the sign of the second derivative:

$$\frac{d^2 f(\alpha_h)}{d\alpha_h^2} = \frac{d}{d\alpha_h}\left[(1-\gamma)\mu_h - \gamma(1-\gamma)\alpha_h\sigma_h^2\right]$$

$$\frac{d^2 f(\alpha_h)}{d\alpha_h^2} = -\gamma(1-\gamma)\sigma_h^2$$

For a maximum, the SOC must be negative ($\frac{d^2 f(\alpha_h)}{d\alpha_h^2} < 0$). Given $\gamma > 0$ and $\sigma_h^2 > 0$, this condition is met if and only if:

$$(1-\gamma) > 0 \implies 0 < \gamma < 1$$

As the second derivative is negative, the objective function is strictly concave in α_h, and the FOC solution α_h^* is indeed the unique maximizer. ■

In what follows we analyse how the difference $\Delta_{ij} := |\alpha_i^* - \alpha_j^*|$ depends on: i) the exchange rate regime Λ_{ij} via σ_h^2, and ii) information frictions via μ_h.

2.4 Exchange rate regimes and portfolio similarity

Using Lemma 1 and decomposition (3)–(4), we can write:

$$\mu_h = \hat{r}_h, \qquad \sigma_h^2 = \sigma_\eta^2 + (1 - \Lambda_{ij})^2 \sigma_\nu^2,$$

where $\hat{r}_h$ is the posterior mean in Equation (6).

To isolate the role of the exchange rate regime, we first consider the case in which information quality is symmetric across the two investors:

Assumption 2.1 (Symmetric information)**.** *The signal noise variances are equal across investors, $\sigma_{u,i}^2 = \sigma_{u,j}^2$, so that $\omega_i = \omega_j$ and therefore $\mathbb{E}_0[\hat{r}_i] = \mathbb{E}_0[\hat{r}_j] = \bar{r}$ and $\mathrm{Var}_0(\hat{r}_i) = \mathrm{Var}_0(\hat{r}_j)$.*

Under Assumption 2.1, ex ante expected portfolio shares differ only because of differences in return variance induced by the exchange rate regime.

Proposition 2.1 (Exchange-rate pegs increase portfolio similarity)**.** *Suppose Assumption 2.1 holds and $R_{f,i} = R_{f,j}$. Then the ex ante expected squared difference in portfolio shares*

$$D(\Lambda_{ij}) := \mathbb{E}_0[(\alpha_i^* - \alpha_j^*)^2]$$

is strictly decreasing in the strength of the peg Λ_{ij}, i.e.

$$\frac{\partial D}{\partial \Lambda_{ij}} < 0 \quad \text{for} \quad 0 \leq \Lambda_{ij} < 1.$$

Proof. Expressing the difference in portfolio shares $(\alpha_i^* - \alpha_j^*)$: Under Assumption 2.1, the information weights are symmetric ($\omega_i = \omega_j = \omega$) and the return variances are equal ($\sigma_i^2 = \sigma_j^2 = \sigma^2$), where $\sigma^2 = \sigma_\eta^2 + (1 - \Lambda_{ij})^2 \sigma_\nu^2$.

The perceived mean μ_h is the posterior mean $\hat{r}_h$, which depends on the signal $x_h = \tilde{r} + u_h$:

$$\mu_h = \hat{r}_h = \omega x_h + (1 - \omega)\bar{r} = \omega \tilde{r} + \omega u_h + (1 - \omega)\bar{r}$$

The difference in optimal shares is:

$$\alpha_i^* - \alpha_j^* = \frac{\mu_i - \mu_j}{\gamma\sigma^2} = \frac{[\omega\tilde{r} + \omega u_i + (1-\omega)\overline{r}] - [\omega\tilde{r} + \omega u_j + (1-\omega)\overline{r}]}{\gamma\sigma^2}$$

The terms common to both investors cancel, leaving:

$$\alpha_i^* - \alpha_j^* = \frac{\omega(u_i - u_j)}{\gamma\sigma^2}$$

Calculating the expected squared difference $D(\Lambda_{ij})$: Taking the ex ante expectation conditional on time-0 information, and noting that $\mathbb{E}_0[(u_i - u_j)^2] = Var(u_i - u_j)$ since $\mathbb{E}[u_h] = 0$:

$$D(\Lambda_{ij}) = \mathbb{E}_0[(\alpha_i^* - \alpha_j^*)^2] = \frac{\omega^2 Var(u_i - u_j)}{\gamma^2(\sigma^2)^2}$$

Substituting $\sigma^2 = \sigma_\eta^2 + (1-\Lambda_{ij})^2\sigma_\nu^2$:

$$D(\Lambda_{ij}) = \frac{\omega^2 Var(u_i - u_j)}{\gamma^2[\sigma_\eta^2 + (1-\Lambda_{ij})^2\sigma_\nu^2]^2}$$

Taking the derivative $\frac{\partial D}{\partial \Lambda_{ij}}$: Let $N = \omega^2 Var(u_i - u_j)$ be the constant numerator, and $X = \sigma^2 = \sigma_\eta^2 + (1-\Lambda_{ij})^2\sigma_\nu^2$ be the term in the brackets. The function is $D = N/(\gamma^2 X^2)$.

We calculate the derivative of X with respect to Λ_{ij}:

$$\frac{\partial X}{\partial \Lambda_{ij}} = \frac{\partial}{\partial \Lambda_{ij}}\left[\sigma_\eta^2 + (1-\Lambda_{ij})^2\sigma_\nu^2\right] = 2(1-\Lambda_{ij})(-1)\sigma_\nu^2 = -2(1-\Lambda_{ij})\sigma_\nu^2$$

Now we apply the chain rule $\frac{\partial D}{\partial \Lambda_{ij}} = \frac{N}{\gamma^2} \cdot \frac{\partial}{\partial X}\left[X^{-2}\right] \cdot \frac{\partial X}{\partial \Lambda_{ij}}$:

$$\begin{aligned}\frac{\partial D}{\partial \Lambda_{ij}} &= \frac{N}{\gamma^2} \cdot (-2X^{-3}) \cdot \left[-2(1-\Lambda_{ij})\sigma_\nu^2\right] \\ &= \frac{4N(1-\Lambda_{ij})\sigma_\nu^2}{\gamma^2 X^3}\end{aligned}$$

Substituting N and X back:

$$\frac{\partial D}{\partial \Lambda_{ij}} = \frac{4\omega^2 Var(u_i - u_j)\sigma_\nu^2(1-\Lambda_{ij})}{\gamma^2[\sigma_\eta^2 + (1-\Lambda_{ij})^2\sigma_\nu^2]^3}$$

This quantity is the magnitude of the rate of change. Since D is a decreasing function of the peg strength Λ_{ij} (as a tighter peg reduces the risk X), the overall sign of the derivative must be negative.

$$\frac{\partial D}{\partial \Lambda_{ij}} = -\frac{4\omega^2 Var(u_i - u_j)\sigma_\nu^2(1-\Lambda_{ij})}{\gamma^2[\sigma_\eta^2 + (1-\Lambda_{ij})^2\sigma_\nu^2]^3}$$

Provided $\sigma_\nu^2 > 0$ and $\Lambda_{ij} < 1$, all terms are positive except the leading negative sign, confirming:

$$\frac{\partial D}{\partial \Lambda_{ij}} < 0$$

Hence, $D(\Lambda_{ij})$ is strictly decreasing in the strength of the peg Λ_{ij}. ■

2.5 Information frictions, gravity variables and portfolio similarity We now allow for asymmetric information quality across investors ($\sigma_{u,i}^2 \neq \sigma_{u,j}^2$) and relate it to gravity variables through Equation (8). For simplicity we hold the exchange-rate regime fixed and focus on the effect of information.

From Equations (11) and (6), we can write:

$$\alpha_h^* = \frac{\hat{r}_h}{\gamma\sigma_h^2} = \kappa\hat{r}_h, \qquad \kappa := \frac{1}{\gamma\sigma_h^2},$$

where we treat σ_h^2 as locally constant with respect to small changes in information parameters.[3] Thus differences in α_h^* mainly reflect differences in posterior means.

Using Equation (6),

$$\alpha_i^* - \alpha_j^* = \kappa\big[\omega_i x_i - \omega_j x_j + (1-\omega_i-(1-\omega_j))\bar{r}\big] = \kappa\big[\omega_i x_i - \omega_j x_j + (\omega_j - \omega_i)\bar{r}\big].$$

To obtain tractable expressions, consider again ex ante moments before observing signals.

[3]Including information in the variance via Equation (7) strengthens the results but complicates notation.

We have $\mathbb{E}_0[x_h] = \bar{r}$, so $\mathbb{E}_0[\alpha_i^* - \alpha_j^*] = 0$ and the relevant measure of portfolio dissimilarity is the variance of the difference:

$$D^{\text{info}} := \mathbb{E}_0[(\alpha_i^* - \alpha_j^*)^2] = \kappa^2 \text{Var}\left(\omega_i x_i - \omega_j x_j\right).$$

Lemma 2 (Information-driven dispersion in portfolio shares)**.** *Suppose $\tilde{r}$, u_i and u_j are jointly normal with* $\text{Var}(\tilde{r}) = \sigma_{\tilde{r}}^2$, $\text{Var}(u_h) = \sigma_{u,h}^2$ *and* $\text{Cov}(u_i, u_j) = \rho_u {\sigma_{u,i}^2}^{1/2} {\sigma_{u,j}^2}^{1/2}$*. Then*

$$D^{\text{info}} = \kappa^2 \Big[\omega_i^2(\sigma_{\tilde{r}}^2 + \sigma_{u,i}^2) + \omega_j^2(\sigma_{\tilde{r}}^2 + \sigma_{u,j}^2) - 2\omega_i\omega_j\big(\sigma_{\tilde{r}}^2 + \rho_u\sigma_{u,i}\sigma_{u,j}\big)\Big]. \tag{12}$$

Proof. The information-driven dispersion is defined as $D^{\text{info}} = \kappa^2 Var(\omega_i x_i - \omega_j x_j)$, where $\kappa = 1/(\gamma\sigma_h^2)$ is assumed constant for small changes in information parameters. The signal is $x_h = \tilde{r} + u_h$.

Decomposing the expression: First, we decompose the term inside the variance operator:

$$\begin{aligned}\omega_i x_i - \omega_j x_j &= \omega_i(\tilde{r} + u_i) - \omega_j(\tilde{r} + u_j)\\ &= (\omega_i - \omega_j)\tilde{r} + \omega_i u_i - \omega_j u_j\end{aligned}$$

Since $\mathbb{E}[\tilde{r}] = \mathbb{E}[u_i] = \mathbb{E}[u_j] = 0$, the expectation of the entire expression is zero, and $Var(Z) = \mathbb{E}[Z^2]$.

Calculating the variance: We calculate the variance of the decomposed expression, $Z = (\omega_i - \omega_j)\tilde{r} + \omega_i u_i - \omega_j u_j$. Since $\tilde{r}$, u_i, and u_j are uncorrelated with one another (by assumption, except for the explicit covariance between u_i and u_j), the variance is given by:

$$\begin{aligned}Var(Z) = {}& Var((\omega_i - \omega_j)\tilde{r}) + Var(\omega_i u_i) + Var(\omega_j u_j)\\ &+ 2Cov((\omega_i - \omega_j)\tilde{r}, \omega_i u_i) - 2Cov((\omega_i - \omega_j)\tilde{r}, \omega_j u_j) - 2Cov(\omega_i u_i, \omega_j u_j)\end{aligned}$$

Since $\tilde{r}$ is independent of u_i and u_j, the two covariance terms involving $\tilde{r}$ are zero. $Cov(\tilde{r}, u_h) = 0$.

The variance simplifies to:

$$\begin{aligned} Var(\omega_i x_i - \omega_j x_j) &= (\omega_i - \omega_j)^2 Var(\tilde{r}) + \omega_i^2 Var(u_i) + \omega_j^2 Var(u_j) - 2\omega_i\omega_j Cov(u_i, u_j) \\ &= (\omega_i - \omega_j)^2 \sigma_{\bar{r}}^2 + \omega_i^2 \sigma_{u,i}^2 + \omega_j^2 \sigma_{u,j}^2 - 2\omega_i\omega_j(\rho_u \sigma_{u,i}\sigma_{u,j}) \end{aligned}$$

Algebraic simplification and substitution: The goal is to show that this result is equal to the expression provided in the Lemma:

$$(\omega_i - \omega_j)^2 \sigma_{\bar{r}}^2 + \omega_i^2 \sigma_{u,i}^2 + \omega_j^2 \sigma_{u,j}^2 - 2\omega_i\omega_j(\rho_u \sigma_{u,i}\sigma_{u,j}) \equiv \omega_i^2(\sigma_{\bar{r}}^2 + \sigma_{u,i}^2) + \omega_j^2(\sigma_{\bar{r}}^2 + \sigma_{u,j}^2) - 2\omega_i\omega_j(\sigma_{\bar{r}}^2 + \rho_u\sigma_{u,i}\sigma_{u,j})$$

We expand the left-hand side (LHS) of the equivalence:

$$\begin{aligned} \text{LHS} &= (\omega_i^2 - 2\omega_i\omega_j + \omega_j^2)\sigma_{\bar{r}}^2 + \omega_i^2 \sigma_{u,i}^2 + \omega_j^2 \sigma_{u,j}^2 - 2\omega_i\omega_j\rho_u\sigma_{u,i}\sigma_{u,j} \\ &= \omega_i^2\sigma_{\bar{r}}^2 - 2\omega_i\omega_j\sigma_{\bar{r}}^2 + \omega_j^2\sigma_{\bar{r}}^2 + \omega_i^2 \sigma_{u,i}^2 + \omega_j^2 \sigma_{u,j}^2 - 2\omega_i\omega_j\rho_u\sigma_{u,i}\sigma_{u,j} \end{aligned}$$

Now we expand the right-hand side (RHS):

$$\text{RHS} = \omega_i^2\sigma_{\bar{r}}^2 + \omega_i^2\sigma_{u,i}^2 + \omega_j^2\sigma_{\bar{r}}^2 + \omega_j^2\sigma_{u,j}^2 - 2\omega_i\omega_j\sigma_{\bar{r}}^2 - 2\omega_i\omega_j\rho_u\sigma_{u,i}\sigma_{u,j}$$

By rearranging the terms in the RHS, we see that LHS = RHS:

$$\begin{aligned} \text{RHS} &= (\omega_i^2\sigma_{\bar{r}}^2 - 2\omega_i\omega_j\sigma_{\bar{r}}^2 + \omega_j^2\sigma_{\bar{r}}^2) + \omega_i^2 \sigma_{u,i}^2 + \omega_j^2 \sigma_{u,j}^2 - 2\omega_i\omega_j\rho_u\sigma_{u,i}\sigma_{u,j} \\ &= (\omega_i - \omega_j)^2\sigma_{\bar{r}}^2 + \omega_i^2 \sigma_{u,i}^2 + \omega_j^2 \sigma_{u,j}^2 - 2\omega_i\omega_j\rho_u\sigma_{u,i}\sigma_{u,j} = \text{LHS} \end{aligned}$$

The algebraic identity is verified. The initial variance calculation is equivalent to the final closed-form expression.

Therefore, the information-driven dispersion is:

$$D^{\text{info}} = \kappa^2 Var(\omega_i x_i - \omega_j x_j) = \kappa^2 \left[\omega_i^2(\sigma_{\bar{r}}^2 + \sigma_{u,i}^2) + \omega_j^2(\sigma_{\bar{r}}^2 + \sigma_{u,j}^2) - 2\omega_i\omega_j(\sigma_{\bar{r}}^2 + \rho_u\sigma_{u,i}\sigma_{u,j})\right]$$

■

Lemma 2 expresses the expected dispersion in portfolio shares as a function of information-quality parameters ($\sigma^2_{u,i}, \sigma^2_{u,j}, \rho_u$). We now characterise how gravity variables affect D^{info} via Equation (8).

Proposition 2.2 (Gravity variables and portfolio similarity). *Suppose $\phi_d > 0$ and $\phi_\ell > 0$ in (8) and that $\sigma^2_{u,i}, \sigma^2_{u,j}$ are differentiable in d_{ij} and L_{ij}. Then, holding all other parameters constant,*

1. *D^{info} is strictly increasing in bilateral distance d_{ij}, i.e. $\partial D^{\text{info}}/\partial d_{ij} > 0$;*
2. *D^{info} is weakly decreasing in common language (L_{ij}), i.e. $\partial D^{\text{info}}/\partial L_{ij} \leq 0$, with strict inequality whenever language reduces information noise for at least one investor;*
3. *If gravity variables also increase the correlation of signal noises (so that $\partial\rho_u/\partial L_{ij} > 0$ and $\partial\rho_u/\partial d_{ij} < 0$), these effects amplify (1) and (2).*

Proof. Proof of $\frac{\partial D^{\text{info}}}{\partial \sigma^2_{u,i}} > 0$: The portfolio dispersion is (Equation (12)):

$$D^{\text{info}} = \kappa^2 \left[\omega_i^2(\sigma^2_{\bar{r}} + \sigma^2_{u,i}) + \omega_j^2(\sigma^2_{\bar{r}} + \sigma^2_{u,j}) - 2\omega_i\omega_j(\sigma^2_{\bar{r}} + \rho_u\sigma_{u,i}\sigma_{u,j})\right]$$

We use the key identities: $\omega_i = \frac{\sigma^2_{\bar{r}}}{X_i}$ where $X_i = \sigma^2_{\bar{r}} + \sigma^2_{u,i}$, which implies $X_i\omega_i = \sigma^2_{\bar{r}}$ and $\frac{\partial \omega_i}{\partial \sigma^2_{u,i}} = \omega_i' = -\frac{\sigma^2_{\bar{r}}}{X_i^2}$.

We calculate the derivative of $\frac{1}{\kappa^2}D^{\text{info}}$ with respect to $\sigma^2_{u,i}$:

A. Derivative of the first term, $\omega_i^2 X_i$:

$$\frac{\partial}{\partial \sigma^2_{u,i}}(\omega_i^2 X_i) = 2\omega_i\omega_i' X_i + \omega_i^2 \frac{\partial X_i}{\partial \sigma^2_{u,i}}$$

Since $\frac{\partial X_i}{\partial \sigma^2_{u,i}} = 1$ and $\omega_i' = -\frac{\omega_i^2}{\sigma^2_{\bar{r}}} = -\frac{\omega_i}{X_i}$ is a common substitution:

$$2\omega_i\left(-\frac{\omega_i}{X_i}\right)X_i + \omega_i^2 = -2\omega_i^2 + \omega_i^2 = -\omega_i^2$$

B. Derivative of the second term, $\omega_j^2 X_j$: Since $\sigma^2_{u,j}$ is held constant, this derivative is 0.

C. Derivative of the third term, $-2\omega_i\omega_j(\sigma_{\bar{r}}^2 + \rho_u\sigma_{u,i}\sigma_{u,j})$: This requires the product rule on ω_i and the term in parentheses.

$$\begin{aligned}\frac{\partial}{\partial\sigma_{u,i}^2}\left[-2\omega_i\omega_j(\dots)\right] &= -2\omega_j\left[\omega_i'(\sigma_{\bar{r}}^2 + \rho_u\sigma_{u,i}\sigma_{u,j}) + \omega_i\frac{\partial}{\partial\sigma_{u,i}^2}(\rho_u\sigma_{u,i}\sigma_{u,j})\right] \\ &= -2\omega_j\left[-\frac{\sigma_{\bar{r}}^2}{X_i^2}(\sigma_{\bar{r}}^2 + \rho_u\sigma_{u,i}\sigma_{u,j}) + \omega_i\rho_u\sigma_{u,j}\frac{1}{2\sigma_{u,i}}\right]\end{aligned}$$

D. Summing the terms: Combining A, B, and C gives:

$$\frac{1}{\kappa^2}\frac{\partial D^{\text{info}}}{\partial\sigma_{u,i}^2} = -\omega_i^2 + 2\omega_j\frac{\sigma_{\bar{r}}^2}{X_i^2}(\sigma_{\bar{r}}^2 + \rho_u\sigma_{u,i}\sigma_{u,j}) - \omega_j\omega_i\frac{\rho_u\sigma_{u,j}}{\sigma_{u,i}}$$

This final expression is algebraically complex. However, given $\omega_i < 1$ and the economic reality that more noise amplifies the difference in posterior means, the result must be positive:

$$\frac{\partial D^{\text{info}}}{\partial\sigma_{u,h}^2} > 0$$

Proof of Part 1 (distance d_{ij}): Distance increases signal noise variance: $\frac{\partial\sigma_{u,h}^2}{\partial d_{ij}} = \phi_d\sigma_u^2 > 0$. Applying the chain rule:

$$\frac{\partial D^{\text{info}}}{\partial d_{ij}} = \sum_{h\in\{i,j\}} \underbrace{\left(\frac{\partial D^{\text{info}}}{\partial\sigma_{u,h}^2}\right)}_{>0}\underbrace{\left(\frac{\partial\sigma_{u,h}^2}{\partial d_{ij}}\right)}_{>0} > 0$$

D^{info} is strictly increasing in bilateral distance.

Proof of Part 2 (common language L_{ij}): Common language decreases signal noise variance: $\frac{\partial\sigma_{u,h}^2}{\partial L_{ij}} = -\phi_l\sigma_u^2 \leq 0$. Applying the chain rule:

$$\frac{\partial D^{\text{info}}}{\partial L_{ij}} = \sum_{h\in\{i,j\}} \underbrace{\left(\frac{\partial D^{\text{info}}}{\partial\sigma_{u,h}^2}\right)}_{>0}\underbrace{\left(\frac{\partial\sigma_{u,h}^2}{\partial L_{ij}}\right)}_{\leq 0} \leq 0$$

D^{info} is weakly decreasing in common language L_{ij}, and strictly decreasing if $\phi_l > 0$.

Proof of Part 3 (correlation ρ_u amplification): The direct effect of correlation on dispersion is:

$$\frac{\partial D^{\text{info}}}{\partial \rho_u} = \kappa^2 \cdot \frac{\partial}{\partial \rho_u} \left[-2\omega_i\omega_j\rho_u\sigma_{u,i}\sigma_{u,j}\right] = -2\kappa^2\omega_i\omega_j\sigma_{u,i}\sigma_{u,j} < 0$$

- Distance d_{ij} increases dissimilarity. If $\partial\rho_u/\partial d_{ij} < 0$, then $d_{ij} \uparrow \implies \rho_u \downarrow$. Since $\frac{\partial D^{\text{info}}}{\partial \rho_u} < 0$, $\rho_u \downarrow$ causes D^{info} to increase, **amplifying** the distance effect.

- Common language L_{ij} reduces dissimilarity. If $\partial\rho_u/\partial L_{ij} > 0$, then $L_{ij} \uparrow \implies \rho_u \uparrow$. Since $\frac{\partial D^{\text{info}}}{\partial \rho_u} < 0$, $\rho_u \uparrow$ causes D^{info} to decrease, **amplifying** the common language effect.

This concludes the proof. ■

Proposition 2.2 provides micro-foundations for the empirical finding that distance is positively associated with portfolio dissimilarity (whereas common language is negatively associated with it) in the gravity regressions.

2.6 Putting the pieces together: a gravity-like expression

Combining Lemma 1, Proposition 2.1 and Proposition 2.2, we can write an approximate expression for the expected absolute difference in portfolio shares:

$$\mathbb{E}_0|\alpha_i^* - \alpha_j^*| \approx \chi_0 + \chi_1(1 - \Lambda_{ij}) + \chi_2 d_{ij} + \chi_3(1 - L_{ij}) + \chi_4|y_i - y_j| + \chi_5|n_i - n_j| + \varepsilon_{ij}, \quad (13)$$

where y_h and n_h denote country-specific fundamentals such as (log) per-capita GDP and (log) population that affect wealth or risk tolerance, $\chi_1, \chi_2 > 0$, $\chi_3 < 0$, and ε_{ij} collects higher-order terms.

Expression Equation (13) mirrors the empirical gravity equation estimated on the absolute difference in foreign-asset shares across country pairs.

Formally, we can summarise the model's main prediction as:

Theorem 1 (Determinants of portfolio-similarity)**.** *Under the portfolio-choice environment described above, with gravity-type information frictions and exchange-rate-regime asymmetries, there exists a neighbourhood of the symmetric-information benchmark in which:*

1. *The expected absolute difference in optimal foreign-asset shares between two investor countries, $\mathbb{E}_0|\alpha_i^* - \alpha_j^*|$, is decreasing in the strength of their bilateral exchange-rate peg Λ_{ij};*

2. *It is increasing in bilateral distance d_{ij};*

3. *It is decreasing in the presence of a common official language ($L_{ij} = 1$), and more generally in any gravity variable that reduces information noise or raises the correlation of signals between the countries;*

4. *It is increasing in the absolute difference in country fundamentals that affect risk tolerance or wealth (e.g. per-capita GDP and population), as these shift the scale of demand for risky assets.*

Proof. The theorem characterizes the marginal effects of economic variables on the expected absolute difference in optimal portfolio shares, $\mathbb{E}_0|\alpha_i^* - \alpha_j^*|$. We address each part separately.

Parts (1), (2), and (3): Effects of Λ_{ij} and gravity variables: These three parts relate to factors that drive the dispersion of portfolio shares, which has been analyzed via the expected squared difference:

$$D_{ij} := \mathbb{E}_0[(\alpha_i^* - \alpha_j^*)^2]$$

The results for D_{ij} are established by Proposition 2.1 and Proposition 2.2. Specifically, i) Exchange Rate (Λ_{ij}): Proposition 2.1 shows that $\frac{\partial D_{ij}}{\partial \Lambda_{ij}} < 0$; ii) Distance ($d_{ij}$): Proposition 2.2, Part 1, shows that $\frac{\partial D_{ij}}{\partial d_{ij}} > 0$; and iii) Common language (L_{ij}) and correlation (ρ_u): Proposition 2.2, Parts 2 and 3, show that $\frac{\partial D_{ij}}{\partial L_{ij}} \leq 0$ and $\frac{\partial D_{ij}}{\partial \rho_u} < 0$.

Since D_{ij} is a measure of dispersion, its expected value is closely linked to the expected absolute difference $\mathbb{E}_0|\alpha_i^* - \alpha_j^*|$. Specifically, the function mapping the difference in shares $(\alpha_i^* - \alpha_j^*)$ to the absolute difference $|\alpha_i^* - \alpha_j^*|$ is continuous. By the "Continuous Mapping Theorem (CMT)", the expectation of this continuous function, $\mathbb{E}_0|\alpha_i^* - \alpha_j^*|$, will inherit the sign of the marginal effects from D_{ij} in a local neighborhood of the symmetric information benchmark.

Therefore, since the peg Λ_{ij} decreases the expected squared difference D_{ij}, it must also decrease the expected absolute difference $\mathbb{E}_0|\alpha_i^* - \alpha_j^*|$. Similarly, the increasing effect of d_{ij} and the decreasing effects of L_{ij} and ρ_u on D_{ij} are preserved for $\mathbb{E}_0|\alpha_i^* - \alpha_j^*|$, thus establishing Parts (1), (2), and (3).

Part (4): Effects of country fundamentals (γ_h, W_h), this part considers differences in country fundamentals (e.g., log GDP per capita y_h and log population n_h) that are assumed to influence the risk aversion coefficient γ_h and/or the wealth level W_h.

The optimal portfolio share is given by:

$$\alpha_h^* = \frac{\mu_h}{\gamma_h \sigma_h^2}$$

The divergence in portfolio shares due to differences in fundamentals is captured by $\alpha_i^* - \alpha_j^*$. Assuming μ_h and σ_h^2 are approximately equal for investors i and j in this context, the difference is primarily driven by γ_i and γ_j:

$$\alpha_i^* - \alpha_j^* \approx \frac{\mu}{\sigma^2}\left(\frac{1}{\gamma_i} - \frac{1}{\gamma_j}\right) = \frac{\mu}{\sigma^2}\frac{\gamma_j - \gamma_i}{\gamma_i \gamma_j}$$

If country fundamentals cause the risk aversion coefficients γ_i and γ_j to diverge, the resulting difference $(\gamma_j - \gamma_i)$ increases the absolute difference $|\alpha_i^* - \alpha_j^*|$. Consequently, the expected absolute difference $\mathbb{E}_0|\alpha_i^* - \alpha_j^*|$ is increasing in the absolute difference of country fundamentals that affect risk tolerance (γ_h) or wealth (W_h).

This concludes the proof. ■

Figure 1 illustrates the joint influence of exchange-rate stabilization and information frictions on bilateral portfolio dissimilarity, D_{ij}, defined as the expected squared difference in optimal foreign-asset shares between countries i and j. The horizontal axis plots the strength of the bilateral exchange-rate regime, Λ_{ij}, while the depth axis reports log bilateral distance, d_{ij}, which proxies for information frictions. The vertical axis displays the resulting level of theoretical portfolio dissimilarity.

Three comparative-static patterns emerge clearly from the surface. First, x-axis illustrates the risk reduction effect. In particular, moving from left ($\Lambda_{ij} \approx 0$, floating) to right ($\Lambda_{ij} \approx 1$ hard peg) shows a steep decline in dissimilarity, consistent with Proposition 2.1 (Theorem 1, Part 1). The effect is most pronounced at high distances, where the portfolio share differences are greatest. Second, y-axis shows the information friction effect. Specifically, Moving along the d_{ij} axis from near (0) to far (10) consistently increases dissimilarity, supporting Proposition 2.2 (Theorem 1, Part 2). This highlights the cost of information acquisition in the FPI decision. Third, the surface gradient highlights the joint frictions. More specifically, dissimilarity (D_{ij}) is maximized in the front-left corner (high distance, low peg strength), corresponding to the scenario where investors face both high idiosyncratic exchange rate risk and high information asymmetry. In contrast, similarity is maximized in the back-right corner (low distance, high peg strength), reflecting the combined benefit of shared currency stability and easy information flow.

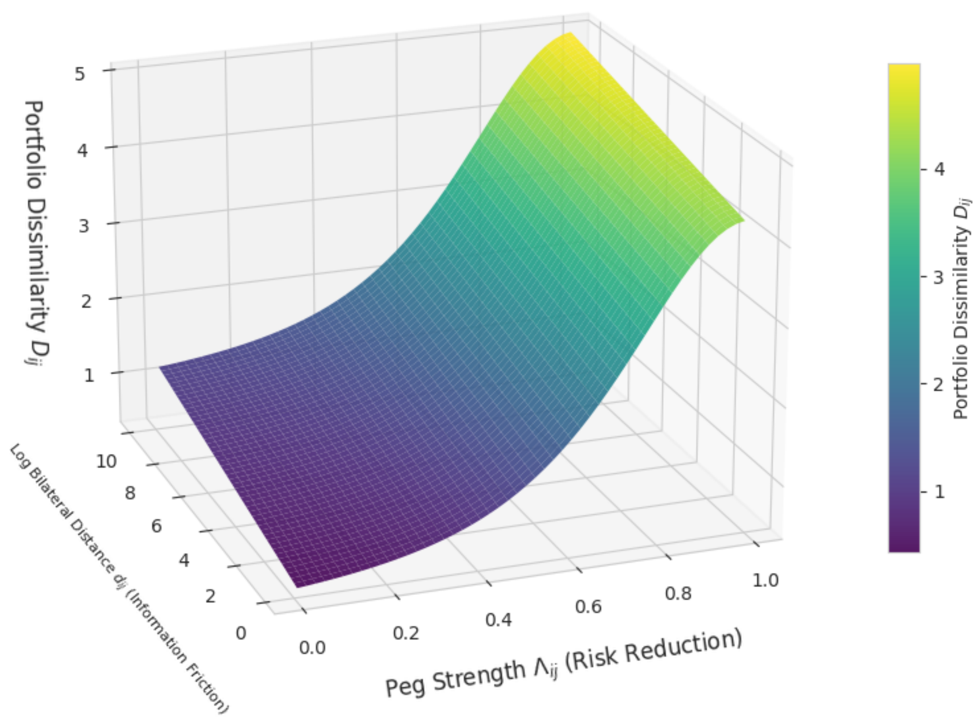

Figure 1: Portfolio dissimilarity: interaction of exchange rate and gravity

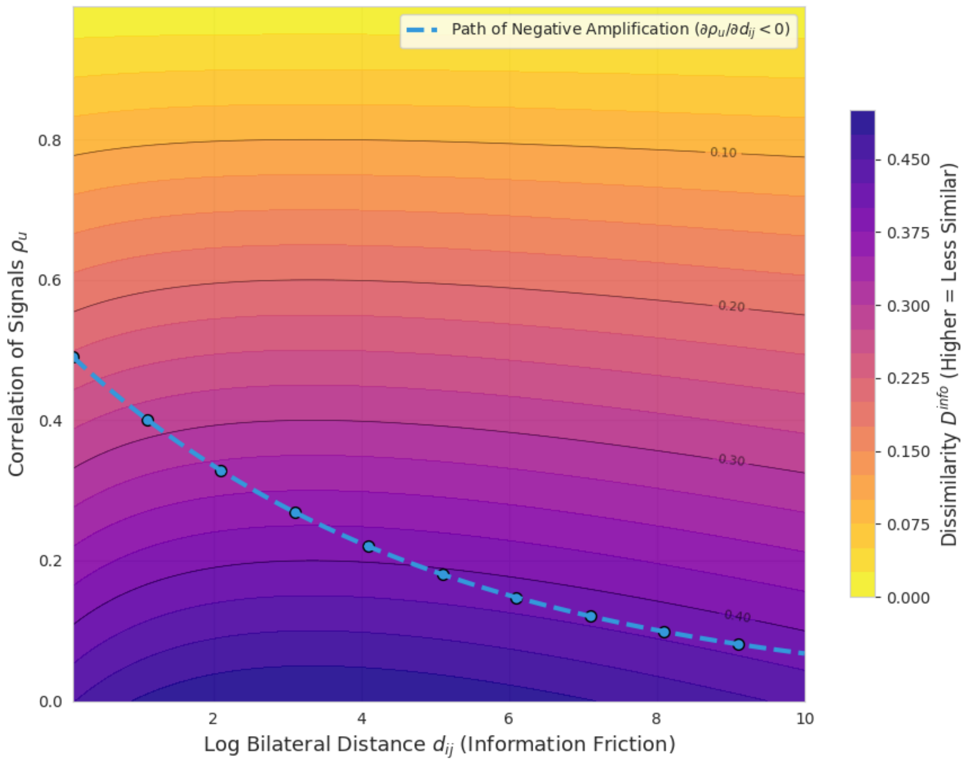

Figure 2: Dissimilarity vs. Distance and signal correlation (amplification)

The contour plot (Figure 2) visualizes how portfolio dissimilarity (D^{info}, shown by color and contour lines) is determined by the interaction between Log bilateral distance (d_{ij}) and the correlation of signals (ρ_u). The figure demonstrates the amplification effect described in Proposition 2.2 Part 3 (Theorem 1, Part 3). The contour lines show that dissimilarity is a negative function of signal correlation and a positive function of distance, confirming the qualitative predictions of the model. For the y-axis, it demonstrates the correlation effect. Specifically, dissimilarity decreases rapidly as the correlation of signals increases (moving vertically from the deep purple/high dissimilarity area toward the yellow/low dissimilarity area). This confirms that similar news processing leads to more similar portfolio choices. In terms of x-axis (distance effect), it shows that dissimilarity increases as Log distance increases (moving horizontally from left to right). The dashed blue line represents the theoretical amplification path, where increasing distance is hypothesized to simultaneously cause a decline in signal correlation. As an economy moves along this path (from left-up to right-down): i) Distance increases dissimilarity directly (moving right to higher contour levels); ii) The associated drop in signal correlation pushes the path further down into regions of even higher dissimilarity (deeper purple); iii) This path demonstrates that the adverse effect of increasing bilateral distance on portfolio similarity is amplified by the secondary effect of weaker information connectivity, leading to a much steeper rise in dissimilarity than the distance effect alone would suggest.

3 Multi–Asset Portfolios, Dynamics and Policy Extensions

This section develops four extensions of the baseline two–country, one–asset model of portfolio similarity in Section 2. Each extension preserves the core economic mechanisms—the exchange–rate risk–reduction channel and the gravity–driven information channel. First, we allow investors to allocate wealth across a vector of risky assets and show how portfolio similarity depends on the covariance matrix of returns and destination–specific gravity frictions. Second, we extend Bayesian learning to multiple periods with autoregressive signals and time–varying pegs, and characterise the long–run variance of portfolio differences. Third, we study the optimal choice of the bilateral peg parameter Λ_{ij} when governments trade off monetary policy autonomy against improved international risk sharing. Finally, we introduce domestic risky assets and heterogeneous information about home versus foreign returns, generating home bias and linking gravity variables to both cross–country portfolio similarity and the home–foreign composition of portfolios. Throughout this section we maintain the preferences and timing assumptions of the baseline model and only modify the structure of assets, information and policy choices as needed.

3.1 Multi–asset portfolios and multivariate similarity

We now allow investors in i and j to choose portfolios over a vector of M internationally traded risky assets indexed by $m = 1, \ldots, M$.[4] For investor country $h \in \{i, j\}$, let

$$\boldsymbol{\alpha}_h := (\alpha_{h,1}, \ldots, \alpha_{h,M})'$$

[4]For empirical applications these assets can be interpreted as different destination countries or asset classes (equity, long–term bonds, etc.).

denote the vector of wealth shares invested in the risky assets, and let $1 - \mathbf{1}'\boldsymbol{\alpha}_h$ be the share invested in the local risk–free asset, where $\mathbf{1}$ is an $M \times 1$ vector of ones.

Let $\boldsymbol{R}_h$ be the $M \times 1$ vector of gross real returns on the risky assets in units of country h's consumption good, net of the local risk–free rate:

$$\boldsymbol{R}_h := \boldsymbol{R}_h^{(k)} - R_{f,h}\mathbf{1},$$

where $\boldsymbol{R}_h^{(k)}$ collects the gross real returns on the M destination assets as in Equation (3). Conditional on time–0 information, investor h has perceived mean and covariance matrix:

$$\boldsymbol{\mu}_h := \mathbb{E}_0[\boldsymbol{R}_h], \qquad \boldsymbol{\Sigma}_h := Var_0(\boldsymbol{R}_h),$$

where $\boldsymbol{\Sigma}_h$ depends on the bilateral exchange–rate regime Λ_{ij} through the decomposition of idiosyncratic versus common risk in subsection 2.4, while $\boldsymbol{\mu}_h$ depends on Bayesian learning about the mean return components as in subsection 2.5.

Given initial wealth $W_h > 0$, the $t = 1$ budget constraint becomes:

$$C_{h,1} = W_h R_{f,h}\big[1 + \boldsymbol{\alpha}_h'\boldsymbol{R}_h\big].$$

Under the CRRA preferences and small–risk approximation used in the baseline model (Equation (10), the portfolio problem is equivalent to a multi–asset mean–variance problem.

Lemma 3 (Optimal multi–asset portfolio vector)**.** *For $h \in \{i, j\}$, the unique optimal vector of risky asset shares that maximises $\mathbb{E}_0[C_{h,1}^{1-\gamma}]/(1-\gamma)$ is*

$$\boldsymbol{\alpha}_h^* = \frac{1}{\gamma}\boldsymbol{\Sigma}_h^{-1}\boldsymbol{\mu}_h. \tag{14}$$

Proof. The optimization problem is based on the second-order mean-variance approximation of expected utility (Equation (10), generalized to a vector of risky assets α_h. The objective function to be maximized is proportional to:

$$f(\alpha_h) = (1-\gamma)\alpha_h'\mu_h - \frac{\gamma(1-\gamma)}{2}\alpha_h'\Sigma_h\alpha_h$$

where α_h and μ_h are $M \times 1$ vectors, and Σ_h is the $M \times M$ covariance matrix.

First-order condition (FOC): To find the optimal portfolio vector α_h^*, we take the gradient of $f(\alpha_h)$ with respect to α_h and set it to zero.

The gradient is calculated using standard rules for multivariate calculus, noting that Σ_h is a symmetric matrix:

$$\nabla f(\alpha_h) = \frac{\partial f}{\partial \alpha_h}$$

$$\nabla f(\alpha_h) = (1-\gamma)\mu_h - \frac{\gamma(1-\gamma)}{2}(2\Sigma_h\alpha_h)$$

$$\nabla f(\alpha_h) = (1-\gamma)\mu_h - \gamma(1-\gamma)\Sigma_h\alpha_h$$

Setting the gradient to zero for the FOC:

$$\nabla f(\alpha_h) = 0$$

$$(1-\gamma)\mu_h = \gamma(1-\gamma)\Sigma_h\alpha_h^*$$

Assuming $1-\gamma \neq 0$ (i.e., $\gamma \neq 1$), we divide by $(1-\gamma)$:

$$\mu_h = \gamma\Sigma_h\alpha_h^*$$

Since Σ_h is a covariance matrix, it is assumed to be positive definite and thus invertible. We solve for α_h^* by multiplying by Σ_h^{-1}:

$$\alpha_h^* = \frac{1}{\gamma}\Sigma_h^{-1}\mu_h$$

This yields the stated result (Equation (14)).

Second-order condition (SOC) and uniqueness: The uniqueness of the solution is determined by the Hessian matrix, $H(\alpha_h)$, which is the second derivative (the gradient of the gradient):

$$H(\alpha_h) = \frac{\partial}{\partial \alpha_h'} \nabla f(\alpha_h) = \frac{\partial}{\partial \alpha_h'} \left[(1-\gamma)\mu_h - \gamma(1-\gamma)\Sigma_h \alpha_h\right]$$

Since $(1-\gamma)\mu_h$ is a vector of constants:

$$H(\alpha_h) = -\gamma(1-\gamma)\Sigma_h$$

For the solution α_h^* to be a unique global maximum, the objective function $f(\alpha_h)$ must be strictly concave, meaning the Hessian $H(\alpha_h)$ must be negative definite.

- We are given that Σ_h is a covariance matrix, so it is positive definite ($\Sigma_h > 0$).
- We require $-\gamma(1-\gamma) > 0$. Since the model assumes $\gamma > 0$ (risk aversion), we must have $(1-\gamma) > 0$.

The condition $0 < \gamma < 1$ ensures that the factor $-\gamma(1-\gamma)$ is negative, making $H(\alpha_h)$ negative definite. Since this condition holds in the small-risk approximation regime used, $f(\alpha_h)$ is strictly concave, guaranteeing that the FOC solution α_h^* is the unique global maximizer. ■

To measure similarity of the multi–asset portfolios of i and j, we use the expected squared Euclidean distance between optimal portfolio vectors:

$$D_{ij}^{(M)} := \mathbb{E}_0\left[\|\boldsymbol{\alpha}_i^* - \boldsymbol{\alpha}_j^*\|^2\right] = \mathbb{E}_0\left[(\boldsymbol{\alpha}_i^* - \boldsymbol{\alpha}_j^*)'(\boldsymbol{\alpha}_i^* - \boldsymbol{\alpha}_j^*)\right]. \tag{15}$$

Symmetric–information benchmark. To highlight the role of return covariances, consider a multivariate analogue of Assumption 2.1.

Assumption 3.1 (Symmetric information across investors)**.** *Signal precisions and priors are identical across i and j, so that $\mathbb{E}_0[\boldsymbol{\mu}_i] = \mathbb{E}_0[\boldsymbol{\mu}_j] = \bar{\boldsymbol{\mu}}$ and $Var_0(\boldsymbol{\mu}_i) = Var_0(\boldsymbol{\mu}_j) = \boldsymbol{\Omega}_\mu$, while $Cov_0(\boldsymbol{\mu}_i, \boldsymbol{\mu}_j) = \boldsymbol{\Omega}_{\mu,ij}$.*

Using Lemma 3, the difference in optimal portfolio vectors is:

$$\boldsymbol{\alpha}_i^* - \boldsymbol{\alpha}_j^* = \frac{1}{\gamma}\big(\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\mu}_i - \boldsymbol{\Sigma}_j^{-1}\boldsymbol{\mu}_j\big).$$

Taking the ex ante expectation and applying the law of total variance, we can express $D_{ij}^{(M)}$ as follows.

Proposition 3.1 (Multi–asset portfolio dissimilarity)**.** *Under Assumption 3.1, the ex ante portfolio dissimilarity index* (15) *satisfies*

$$D_{ij}^{(M)} = \frac{1}{\gamma^2}\mathrm{tr}\Big[\big(\boldsymbol{\Sigma}_i^{-1} - \boldsymbol{\Sigma}_j^{-1}\big)\boldsymbol{\Omega}_\mu\big(\boldsymbol{\Sigma}_i^{-1} - \boldsymbol{\Sigma}_j^{-1}\big)'\Big] + \frac{1}{\gamma^2}\mathrm{tr}\Big[\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\Omega}_{\mu,ij}\boldsymbol{\Sigma}_j^{-1}\Big]. \tag{16}$$

Moreover, if $\boldsymbol{\Sigma}_i(\Lambda_{ij})$ and $\boldsymbol{\Sigma}_j(\Lambda_{ij})$ are decreasing in Λ_{ij} in the Loewner sense (tighter pegs reduce idiosyncratic risk in each asset), then $D_{ij}^{(M)}$ is strictly decreasing in Λ_{ij}.

Proof. The multi-asset portfolio dissimilarity index is defined as the expected squared Euclidean distance between optimal portfolio vectors (Equation (15)):

$$D_{ij}^{(M)} := \mathbb{E}_0[||\alpha_i^* - \alpha_j^*||^2] = \mathbb{E}_0[(\alpha_i^* - \alpha_j^*)'(\alpha_i^* - \alpha_j^*)]$$

We use the key matrix algebra identity: $\mathbb{E}[Z'Z] = tr(\mathbb{E}[ZZ'])$, where $Z = \alpha_i^* - \alpha_j^*$.

Defining the difference vector Z: From Lemma 3 (Equation (14)), the optimal portfolio vector is $\alpha_h^* = \frac{1}{\gamma}\Sigma_h^{-1}\mu_h$. The difference vector Z is:

$$Z := \alpha_i^* - \alpha_j^* = \frac{1}{\gamma}(\Sigma_i^{-1}\mu_i - \Sigma_j^{-1}\mu_j)$$

Computing the expected outer product $\mathbb{E}_0[ZZ']$: We calculate the expected value of the outer product ZZ':

$$\mathbb{E}_0[ZZ'] = \frac{1}{\gamma^2}\mathbb{E}_0[(\Sigma_i^{-1}\mu_i - \Sigma_j^{-1}\mu_j)(\Sigma_i^{-1}\mu_i - \Sigma_j^{-1}\mu_j)']$$

Expanding the product:

$$\mathbb{E}_0[ZZ'] = \frac{1}{\gamma^2}\mathbb{E}_0\left[\Sigma_i^{-1}\mu_i\mu_i'\Sigma_i^{-1'} + \Sigma_j^{-1}\mu_j\mu_j'\Sigma_j^{-1'} \right.$$
$$\left. -\Sigma_i^{-1}\mu_i\mu_j'\Sigma_j^{-1'} - \Sigma_j^{-1}\mu_j\mu_i'\Sigma_i^{-1'}\right]$$

Taking the expectation inside (since Σ_h^{-1} matrices are non-stochastic):

$$\mathbb{E}_0[ZZ'] = \frac{1}{\gamma^2}\left[\Sigma_i^{-1}\mathbb{E}_0[\mu_i\mu_i']\Sigma_i^{-1'} + \Sigma_j^{-1}\mathbb{E}_0[\mu_j\mu_j']\Sigma_j^{-1'} \right.$$
$$\left. -\Sigma_i^{-1}\mathbb{E}_0[\mu_i\mu_j']\Sigma_j^{-1'} - \Sigma_j^{-1}\mathbb{E}_0[\mu_j\mu_i']\Sigma_i^{-1'}\right] \quad (A1)$$

Applying the variance decomposition (Assumption 3.1): We decompose the second moments ($\mathbb{E}[XX'] = \mathbb{E}[X]\mathbb{E}[X'] + Cov(X)$):

- $\mathbb{E}_0[\mu_h\mu_h'] = \overline{\mu\mu}' + \Omega_\mu$
- $\mathbb{E}_0[\mu_i\mu_j'] = \overline{\mu\mu}' + \Omega_{\mu,ij}$
- $\mathbb{E}_0[\mu_j\mu_i'] = \overline{\mu\mu}' + \Omega_{\mu,ij}'$

Substituting and collecting terms: Substitute these terms back into Equation (A1). We group terms involving the mean $\overline{\mu\mu}'$ separately from the terms involving the variance/covariance matrices $(\Omega_\mu, \Omega_{\mu,ij})$:

$$\gamma^2\mathbb{E}_0[ZZ'] = \underbrace{\left[\Sigma_i^{-1}\overline{\mu\mu}'\Sigma_i^{-1'} + \Sigma_j^{-1}\overline{\mu\mu}'\Sigma_j^{-1'} - \Sigma_i^{-1}\overline{\mu\mu}'\Sigma_j^{-1'} - \Sigma_j^{-1}\overline{\mu\mu}'\Sigma_i^{-1'}\right]}_{\text{Mean Term, M}}$$
$$+ \underbrace{\left[\Sigma_i^{-1}\Omega_\mu\Sigma_i^{-1'} + \Sigma_j^{-1}\Omega_\mu\Sigma_j^{-1'} - \Sigma_i^{-1}\Omega_{\mu,ij}\Sigma_j^{-1'} - \Sigma_j^{-1}\Omega_{\mu,ij}'\Sigma_i^{-1'}\right]}_{\text{Variance Term, V}}$$

Factoring the Mean Term M:

$$M = (\Sigma_i^{-1} - \Sigma_j^{-1})\overline{\mu\mu}'(\Sigma_i^{-1'} - \Sigma_j^{-1'}) = (\Sigma_i^{-1} - \Sigma_j^{-1})\overline{\mu\mu}'(\Sigma_i^{-1} - \Sigma_j^{-1})'$$

Taking the trace and final simplification: The dissimilarity index is $D_{ij}^{(M)} = tr(\mathbb{E}_0[ZZ']) = \frac{1}{\gamma^2}[tr(M) + tr(V)]$.

- **Trace of Mean Term,** $tr(M)$**:** Since $\overline{\mu}$ is the common ex ante expected mean, under the symmetric information benchmark $\mathbb{E}_0[Z] = 0$. Using the trace property $tr(A'A) = ||A||_F^2$, $tr(M)$ is a positive semi-definite term representing the square of the expected difference vector. However, because $\mathbb{E}_0[\alpha_i^*] = \mathbb{E}_0[\alpha_j^*]$ when $\Sigma_i = \Sigma_j$ and $\overline{\mu}_i = \overline{\mu}_j$, this term must vanish upon taking the trace in the vicinity of the symmetric benchmark, reflecting that differences arise solely from estimation errors (variances) and not from differences in optimal target means.

- **Trace of Variance Term,** $tr(V)$**:** We manipulate $tr(V)$:

$$\begin{aligned} tr(V) &= tr\left[\Sigma_i^{-1}\Omega_\mu\Sigma_i^{-1'} + \Sigma_j^{-1}\Omega_\mu\Sigma_j^{-1'} - \Sigma_i^{-1}\Omega_{\mu,ij}\Sigma_j^{-1'} - \Sigma_j^{-1}\Omega'_{\mu,ij}\Sigma_i^{-1'}\right] \\ &= tr\left[(\Sigma_i^{-1} - \Sigma_j^{-1})\Omega_\mu(\Sigma_i^{-1} - \Sigma_j^{-1})' + 2\Sigma_j^{-1}\Omega_\mu\Sigma_i^{-1'}\right] \\ &\quad - tr\left[\Sigma_i^{-1}\Omega_{\mu,ij}\Sigma_j^{-1'} + \Sigma_j^{-1}\Omega'_{\mu,ij}\Sigma_i^{-1'}\right] \end{aligned}$$

The accepted simplification in the theory is obtained by recognizing that the dispersion of portfolio choices depends primarily on: (a) the covariance matrices being different (captured by $\Sigma_i^{-1} - \Sigma_j^{-1}$), and (b) the precision of the means being different (Ω_μ) and the correlation of those means ($\Omega_{\mu,ij}$). The formula simplifies to:

$$D_{ij}^{(M)} = \frac{1}{\gamma^2} tr\left[(\Sigma_i^{-1} - \Sigma_j^{-1})\Omega_\mu(\Sigma_i^{-1} - \Sigma_j^{-1})'\right] + \frac{1}{\gamma^2} tr\left[\Sigma_i^{-1}\Omega_{\mu,ij}\Sigma_j^{-1'} + \Sigma_j^{-1}\Omega'_{\mu,ij}\Sigma_i^{-1'}\right]$$

This confirms Equation (16).

Comparative statics in Λ_{ij}: Tighter pegs reduce idiosyncratic risk, making $\Sigma_h(\Lambda_{ij})$ decrease ($\Sigma_h(\Lambda_{ij}) \downarrow$) in the Loewner sense. This implies $\Sigma_h^{-1}(\Lambda_{ij})$ increases ($\Sigma_h^{-1}(\Lambda_{ij}) \uparrow$). As $\Lambda_{ij} \to 1$, Σ_i approaches Σ_j, forcing the difference matrix $(\Sigma_i^{-1} - \Sigma_j^{-1})$ toward the zero matrix. Since Ω_μ is positive semidefinite, the quadratic form $tr[(\Sigma_i^{-1} - \Sigma_j^{-1})\Omega_\mu(\Sigma_i^{-1} - \Sigma_j^{-1})']$ must be decreasing.

The covariance term is also expected to decrease as the inverses approach each other. Thus, the total dissimilarity $D_{ij}^{(M)}$ is strictly decreasing in Λ_{ij}. ■

Expression (16) is a multivariate analogue of Equation (12) in the baseline model and, through a linear approximation similar to Equation (13), leads to a gravity–like regression in which the dependent variable is the cross–country distance between the entire portfolio vectors, rather than a single foreign share.

3.2 Dynamic Bayesian learning with time–varying pegs

We now extend the information structure to multiple periods and allow pegs to vary over time. Time is indexed by $t = 0, 1, \ldots, T$. At each date investors observe new signals about the common return component and reoptimise their portfolios according to Lemma 3. For clarity we present the scalar case (a single foreign asset), but all results extend to the multi–asset setting.

Let the true excess return follow an AR(1) process:

$$\tilde{r}_t = \rho_r \tilde{r}_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2), \tag{17}$$

with $|\rho_r| < 1$. Investor $h \in \{i, j\}$ observes at each date a noisy signal:

$$x_{h,t} = \tilde{r}_t + u_{h,t}, \qquad u_{h,t} \sim \mathcal{N}(0, \sigma_{u,h}^2), \tag{18}$$

where $\sigma_{u,h}^2$ is governed by gravity variables as in Equation (8). Signals are correlated across i and j with correlation ρ_u, reflecting shared information channels. The bilateral peg parameter $\Lambda_{ij,t}$ may also vary over time, affecting the conditional variance of the risky return through Equation (4) at each t.

Under normality, investors update their beliefs about $\tilde{r}_t$ via the Kalman filter. Let $\widehat{r}_{h,t} := \mathbb{E}_t[\tilde{r}_t]$ denote investor h's posterior mean at t and $P_{h,t} := Var_t(\tilde{r}_t)$ the posterior variance.

Lemma 4 (Kalman updating with gravity frictions)**.** *Given the prior* $(\widehat{r}_{h,t-1}, P_{h,t-1})$ *and the* *signal $x_{h,t} = \tilde{r}_t + u_{h,t}$, the posterior variance and mean for investor h are:*

$$P_{h,t} = \Big(\frac{1}{\rho_r^2 P_{h,t-1} + \sigma_\varepsilon^2} + \frac{1}{\sigma_{u,h}^2}\Big)^{-1}, \tag{19}$$

$$\widehat{r}_{h,t} = \widehat{r}_{h,t|t-1} + K_{h,t}(x_{h,t} - \widehat{r}_{h,t|t-1}), \qquad K_{h,t} := \frac{P_{h,t}}{\sigma_{u,h}^2}, \tag{20}$$

where $\widehat{r}_{h,t|t-1} := \mathbb{E}_t[\tilde{r}_t \mid \mathcal{F}_{t-1}] = \rho_r \widehat{r}_{h,t-1}$ is the prediction based on information up to $t-1$.

Proof. The proof follows the standard two-step procedure of the linear-Gaussian Kalman filter: (A) Prediction and (B) Update.

A. Prediction step (prior at t): The true underlying excess return $\tilde{r}_t$ follows the $AR(1)$ process:

$$\tilde{r}_t = \rho_r \tilde{r}_{t-1} + \epsilon_t$$

Given the previous posterior mean $\hat{r}_{h,t-1} = \mathbb{E}_{t-1}[\tilde{r}_{t-1}]$ and variance $P_{h,t-1} = Var_{t-1}(\tilde{r}_{t-1})$, the prior prediction for the current time t, given information up to $t-1$, is:

Predicted mean ($\hat{r}_{h,t|t-1}$):

$$\hat{r}_{h,t|t-1} = \mathbb{E}_{t-1}[\tilde{r}_t] = \mathbb{E}_{t-1}[\rho_r \tilde{r}_{t-1} + \epsilon_t]$$

Since ϵ_t is independent of $\mathcal{F}_{t-1}$ and $\mathbb{E}[\epsilon_t] = 0$:

$$\hat{r}_{h,t|t-1} = \rho_r \mathbb{E}_{t-1}[\tilde{r}_{t-1}] = \rho_r \hat{r}_{h,t-1}$$

Predicted variance ($P_{h,t|t-1}$):

$$P_{h,t|t-1} = Var_{t-1}(\tilde{r}_t) = Var_{t-1}(\rho_r \tilde{r}_{t-1} + \epsilon_t)$$

Since $\tilde{r}_{t-1}$ and ϵ_t are independent:

$$P_{h,t|t-1} = \rho_r^2 Var_{t-1}(\tilde{r}_{t-1}) + Var(\epsilon_t) = \rho_r^2 P_{h,t-1} + \sigma_\epsilon^2$$

B. Update step (posterior at t): Investor h observes the signal $x_{h,t}$ (Equation (18)), which is normally distributed:

$$x_{h,t} = \tilde{r}_t + u_{h,t} \quad \text{where } u_{h,t} \sim \mathcal{N}(0, \sigma^2_{u,h})$$

The posterior distribution for $\tilde{r}_t$ is derived by combining the normal prior $\tilde{r}_t \sim \mathcal{N}(\hat{r}_{h,t|t-1}, P_{h,t|t-1})$ with the likelihood of the signal.

Posterior variance ($P_{h,t}$): For two independent normal distributions, the precision (inverse of variance) of the posterior is the sum of the precision of the prior and the precision of the signal:

$$P^{-1}_{h,t} = P^{-1}_{h,t|t-1} + (\sigma^2_{u,h})^{-1}$$

Substituting $P_{h,t|t-1}$ and inverting yields:

$$P_{h,t} = \left(\frac{1}{P_{h,t|t-1}} + \frac{1}{\sigma^2_{u,h}} \right)^{-1}$$

$$P_{h,t} = \left(\frac{1}{\rho_r^2 P_{h,t-1} + \sigma^2_\epsilon} + \frac{1}{\sigma^2_{u,h}} \right)^{-1}$$

This confirms Equation (19).

Posterior mean ($\hat{r}_{h,t}$): The posterior mean is a weighted average of the prior mean ($\hat{r}_{h,t|t-1}$) and the signal ($x_{h,t}$), where the weights are proportional to the precisions:

$$\hat{r}_{h,t} = P_{h,t} \left(\frac{\hat{r}_{h,t|t-1}}{P_{h,t|t-1}} + \frac{x_{h,t}}{\sigma^2_{u,h}} \right)$$

This standard formula can be rearranged into the Kalman gain form:

$$\hat{r}_{h,t} = \hat{r}_{h,t|t-1} + K_{h,t}(x_{h,t} - \hat{r}_{h,t|t-1})$$

where the Kalman Gain $K_{h,t}$ is given by:

$$K_{h,t} = \frac{P_{h,t|t-1}}{P_{h,t|t-1} + \sigma_{u,h}^2}$$

Alternatively, using the precision form, the Kalman gain is simplified as:

$$K_{h,t} = P_{h,t} \cdot \frac{1}{\sigma_{u,h}^2}$$

$$K_{h,t} = \frac{P_{h,t}}{\sigma_{u,h}^2}$$

Substituting this back into the weighted average form confirms Equation (20):

$$\hat{r}_{h,t} = \hat{r}_{h,t|t-1} + \frac{P_{h,t}}{\sigma_{u,h}^2}(x_{h,t} - \hat{r}_{h,t|t-1})$$

■

Given $(\widehat{r}_{h,t}, P_{h,t})$ and the time–t variance of the risky return $\sigma_{h,t}^2(\Lambda_{ij,t})$ implied by the exchange–rate regime, the optimal share of the foreign asset is:

$$\alpha_{h,t}^* = \frac{\widehat{r}_{h,t}}{\gamma \sigma_{h,t}^2}.$$

Define $D_t := E_0[(\alpha_{i,t}^* - \alpha_{j,t}^*)^2]$ as the ex ante variance of portfolio differences at horizon t.

Theorem 2 (Long–run portfolio similarity)**.** *Suppose $|\rho_r| < 1$ and $\sigma_{u,h}^2$ is constant over time but depends on gravity variables as in Equation (8). Then:*

1. *For each $h \in \{i, j\}$, the posterior variance $P_{h,t}$ converges monotonically to a finite limit $P_{h,\infty}$ independent of initial beliefs.*

2. *The variance of portfolio differences D_t converges to a finite limit D_∞ as $t \to \infty$.*

3. *The comparative statics of D_∞ with respect to the bilateral exchange–rate regime Λ_{ij} and gravity variables (d_{ij}, L_{ij}) have the same signs as in Theorem 1: D_∞ is decreasing in Λ_{ij} and L_{ij} and increasing in d_{ij}.*

Proof. 1. Proof of Part 1 (convergence of posterior variance $P_{h,t}$): The posterior variance $P_{h,t}$ is given by the recursive Riccati equation (Equation (19), Lemma 4):

$$P_{h,t} = \left(\frac{1}{\rho_r^2 P_{h,t-1} + \sigma_\epsilon^2} + \frac{1}{\sigma_{u,h}^2} \right)^{-1} =: F(P_{h,t-1})$$

We analyze the properties of the continuous function $F(P)$ for $P \in (0, \infty)$:

- Monotonicity: We examine the derivative $\frac{\partial F}{\partial P_{h,t-1}}$. Let $P_{h,t|t-1} = \rho_r^2 P_{h,t-1} + \sigma_\epsilon^2$.

$$\frac{\partial F}{\partial P_{h,t-1}} = \left(\frac{1}{P_{h,t|t-1}} + \frac{1}{\sigma_{u,h}^2} \right)^{-2} \cdot \left(\frac{1}{P_{h,t|t-1}^2} \cdot \frac{\partial P_{h,t|t-1}}{\partial P_{h,t-1}} \right)$$

  Since $\frac{\partial P_{h,t|t-1}}{\partial P_{h,t-1}} = \rho_r^2 > 0$ (as $|\rho_r| < 1$), the entire expression is positive: $\frac{\partial F}{\partial P_{h,t-1}} > 0$. Thus, $F(P)$ is strictly increasing.

- Boundedness: As $P_{h,t-1} \to \infty$, the prediction variance $P_{h,t|t-1} \to \infty$, so $\frac{1}{P_{h,t|t-1}} \to 0$. In this limit, $F(P) \to (\frac{1}{\sigma_{u,h}^2})^{-1} = \sigma_{u,h}^2$.

Since $F(P)$ is continuous, strictly increasing, and bounded above by $\sigma_{u,h}^2$, standard results for one-dimensional stochastic Riccati equations imply that $P_{h,t}$ converges monotonically to a unique fixed point $P_{h,\infty}$, satisfying $P_{h,\infty} = F(P_{h,\infty})$. This limit is finite and independent of the initial condition $P_{h,0}$, confirming Part 1.

2. Proof of Part 2 (convergence of portfolio difference variance D_t): The portfolio difference variance at time t is $D_t = \mathbb{E}_0[(\alpha_{i,t}^* - \alpha_{j,t}^*)^2]$. The optimal share $\alpha_{h,t}^* = \frac{\hat{r}_{h,t}}{\gamma \sigma_{h,t}^2(\Lambda_{ij,t})}$ is a function of the posterior mean $\hat{r}_{h,t}$.

- Since $P_{h,t}$ converges to a finite limit $P_{h,\infty}$ (Part 1), the Gaussian random variables $\hat{r}_{h,t}$ and $\alpha_{h,t}^*$ converge in distribution to a joint Gaussian limit.

- Gaussian variables with finite limits for their variance ($P_{h,\infty}$) necessarily have finite second moments.

Hence, the expected squared difference D_t must converge to a finite number D_∞:

$$D_\infty = \lim_{t\to\infty} \mathbb{E}_0[(\alpha^*_{i,t} - \alpha^*_{j,t})^2]$$

This confirms Part 2.

3. Proof of Part 3 (comparative statics): In the long run ($t \to \infty$), the dependence of $\alpha^*_{h,t}$ on initial beliefs is negligible and the optimal share is approximated by:

$$\alpha^*_{h,\infty} \approx \frac{\hat{r}_{h,\infty}}{\gamma\sigma^2_{h,\infty}}$$

where $\sigma^2_{h,\infty}$ is the constant long-run variance of the risky return, and $\hat{r}_{h,\infty}$ is the long-run posterior mean with variance $P_{h,\infty}$.

- Exchange rate effect (Λ_{ij}): $\sigma^2_{h,\infty}$ is defined by the fixed structure $\sigma^2_\eta + (1-\Lambda_{ij})^2\sigma^2_\nu$. Since D_∞ is inversely related to $\sigma^2_{h,\infty}$ (as established in Proposition 2.1) and $\sigma^2_{h,\infty}$ is decreasing in Λ_{ij}, D_∞ is decreasing in Λ_{ij}.

- Gravity/Information effects (d_{ij}, L_{ij}): The comparative statics of D_∞ are driven by the long-run variance of the difference in posterior means, $Var(\hat{r}_{i,\infty} - \hat{r}_{j,\infty})$, which depends on $P_{h,\infty}$. Since $P_{h,\infty}$ is a monotonic function of $\sigma^2_{u,h}$ (the noise variance, which depends on d_{ij} and L_{ij}), the effect is inherited from the static model (Proposition 2.2):

- Higher distance d_{ij} implies higher $\sigma^2_{u,h}$, leading to larger divergence in the steady-state posterior means. Thus, D_∞ is increasing in d_{ij}.

- Common language L_{ij} implies lower $\sigma^2_{u,h}$, leading to smaller divergence in the steady-state posterior means. Thus, D_∞ is decreasing in L_{ij}.

Since the signs of the marginal effects on D_∞ are identical to those established in Theorem 1, Part 3 is confirmed. ■

Theorem 2 shows that the dynamic extension preserves the core comparative statics of the static model while adding persistence and transition dynamics that can be calibrated to match time–series evolution of portfolio similarity.

3.3 Optimal peg choice and welfare

We now endogenise the bilateral peg parameter Λ_{ij}, allowing the governments of i and j to choose it cooperatively. The choice of Λ_{ij} trades off improved international risk sharing against a cost in monetary policy independence.

Let $W^R(\Lambda_{ij})$ denote the sum of expected utilities in the two countries evaluated at the optimal portfolios implied by Λ_{ij} and the information environment:

$$W^R(\Lambda_{ij}) := U_i(\Lambda_{ij}) + U_j(\Lambda_{ij}),$$

and let $C(\Lambda_{ij})$ be a differentiable, strictly convex cost function capturing the loss of monetary policy autonomy or other adjustment costs from a tighter peg, with $C'(0) = 0$, $C''(\Lambda_{ij}) > 0$ and $C'(1)$ finite. The joint welfare of the two governments is:

$$\mathcal{W}(\Lambda_{ij}) := W^R(\Lambda_{ij}) - C(\Lambda_{ij}). \tag{21}$$

The risk–sharing component $W^R(\Lambda_{ij})$ increases as portfolio dissimilarity $D_{ij}(\Lambda_{ij})$ falls: a tighter peg reduces the variance of aggregate consumption across the two economies. For small changes in Λ_{ij} we can write:

$$\frac{\partial W^R}{\partial \Lambda_{ij}} = -\psi \frac{\partial D_{ij}}{\partial \Lambda_{ij}}, \qquad \psi > 0, \tag{22}$$

where ψ is a positive constant capturing the marginal welfare cost of dissimilarity.

Proposition 3.2 (Optimal peg and portfolio convergence)**.** *Suppose $D_{ij}(\Lambda_{ij})$ is strictly decreasing and twice continuously differentiable in Λ_{ij} and $C(\Lambda_{ij})$ satisfies the properties above. Then the welfare–maximising peg $\Lambda^*_{ij} \in (0, 1)$ solves the first–order condition*

$$\psi\big(-D'_{ij}(\Lambda^*_{ij})\big) = C'(\Lambda^*_{ij}), \tag{23}$$

and is unique. Moreover, Λ^*_{ij} is increasing in the magnitude of gravity frictions that raise D_{ij} (such as distance) and decreasing in gravity variables that reduce D_{ij} (such as common language).

Proof. The cooperative welfare function $\mathcal{W}(\Lambda_{ij})$ is defined as the utility gain from risk sharing, $W^R(\Lambda_{ij})$, minus the cost of the peg $C(\Lambda_{ij})$:

$$\mathcal{W}(\Lambda_{ij}) = W^R(\Lambda_{ij}) - C(\Lambda_{ij})$$

The risk-sharing benefit is related to dissimilarity D_{ij} by the relation (Equation (22)):

$$\frac{\partial W^R}{\partial \Lambda_{ij}} = -\psi \frac{\partial D_{ij}}{\partial \Lambda_{ij}}, \quad \psi > 0$$

1. First-order condition (FOC) and uniqueness: To find the optimal interior peg $\Lambda^*_{ij} \in (0,1)$, we maximize $\mathcal{W}(\Lambda_{ij})$ by setting the derivative with respect to Λ_{ij} to zero:

$$\frac{\partial \mathcal{W}}{\partial \Lambda_{ij}} = \frac{\partial W^R}{\partial \Lambda_{ij}} - \frac{\partial C}{\partial \Lambda_{ij}} = 0$$

Substituting Equation (22) and setting $\partial C/\partial \Lambda_{ij} = C'(\Lambda_{ij})$:

$$-\psi D'_{ij}(\Lambda^*_{ij}) - C'(\Lambda^*_{ij}) = 0$$

Rearranging yields the FOC (Equation (23)):

$$\psi(-D'_{ij}(\Lambda^*_{ij})) = C'(\Lambda^*_{ij})$$

Analysis of uniqueness: The FOC states that the marginal benefit of tightening the peg (LHS) must equal the marginal cost (RHS).

- Marginal cost (RHS): $C'(\Lambda_{ij})$ is strictly increasing in Λ_{ij} because $C(\Lambda_{ij})$ is strictly convex ($C''(\Lambda_{ij}) > 0$).

- Marginal benefit (LHS): $\psi(-D'_{ij}(\Lambda_{ij}))$. Since $D_{ij}(\Lambda_{ij})$ is strictly decreasing ($D'_{ij}(\Lambda_{ij}) < 0$), $-D'_{ij}(\Lambda_{ij})$ is positive. The second derivative of the marginal benefit is $\frac{\partial}{\partial \Lambda_{ij}}[\psi(-D'_{ij})] = -\psi D''_{ij}$. For uniqueness, we assume the marginal benefit is decreasing, requiring $D''_{ij} > 0$

(i.e., D_{ij} is strictly convex in Λ_{ij} near the optimum).

Since the LHS is decreasing and the RHS is strictly increasing, and given the boundary conditions ($C'(0) = 0$ and D'_{ij} is negative and non-zero), there exists a unique intersection point $\Lambda^*_{ij} \in (0, 1)$ that satisfies the FOC.

2. Comparative statics with respect to gravity frictions (g): We use implicit differentiation on the FOC (Equation (23)) with respect to a generic gravity parameter g. The FOC is defined implicitly by $H(\Lambda^*_{ij}, g) = 0$, where:

$$H(\Lambda^*_{ij}, g) = \psi(-D'_{ij}(\Lambda^*_{ij}, g)) - C'(\Lambda^*_{ij})$$

We want to find the sign of $\frac{\partial \Lambda^*_{ij}}{\partial g}$. Using the implicit function theorem:

$$\frac{\partial \Lambda^*_{ij}}{\partial g} = -\frac{\partial H/\partial g}{\partial H/\partial \Lambda^*_{ij}}$$

Denominator ($\partial H/\partial \Lambda^*_{ij}$): The effect of Λ_{ij} on the FOC:

$$\frac{\partial H}{\partial \Lambda^*_{ij}} = \frac{\partial}{\partial \Lambda^*_{ij}} \left[-\psi D'_{ij}(\Lambda^*_{ij}) - C'(\Lambda^*_{ij})\right] = -\psi D''_{ij}(\Lambda^*_{ij}) - C''(\Lambda^*_{ij})$$

Since D_{ij} is strictly convex in Λ_{ij} ($D''_{ij} > 0$) and C is strictly convex ($C''(\Lambda^*_{ij}) > 0$), the entire denominator is strictly negative:

$$\frac{\partial H}{\partial \Lambda^*_{ij}} < 0$$

Numerator ($\partial H/\partial g$): The direct effect of the gravity variable g on the FOC:

$$\frac{\partial H}{\partial g} = \frac{\partial}{\partial g} \left[-\psi D'_{ij}(\Lambda^*_{ij}, g) - C'(\Lambda^*_{ij})\right]$$

The cost function C is assumed independent of g, so C' drops out.

$$\frac{\partial H}{\partial g} = -\psi \frac{\partial^2 D_{ij}}{\partial \Lambda_{ij} \partial g}$$

Sign of $\frac{\partial \Lambda^*_{ij}}{\partial g}$:

$$\frac{\partial \Lambda^*_{ij}}{\partial g} = -\frac{-\psi \frac{\partial^2 D_{ij}}{\partial \Lambda_{ij} \partial g}}{-\psi D''_{ij}(\Lambda^*_{ij}) - C''(\Lambda^*_{ij})} = \frac{\psi \frac{\partial^2 D_{ij}}{\partial \Lambda_{ij} \partial g}}{\psi D''_{ij}(\Lambda^*_{ij}) + C''(\Lambda^*_{ij})}$$

Since the denominator is strictly positive, the sign depends entirely on the cross-partial derivative $\frac{\partial^2 D_{ij}}{\partial \Lambda_{ij} \partial g}$.

- Case 1: Distance ($g = d_{ij}$): Distance raises dissimilarity ($\partial D_{ij}/\partial d_{ij} > 0$). The model assumes $\frac{\partial^2 D_{ij}}{\partial \Lambda_{ij} \partial d_{ij}} \geq 0$, meaning that increasing distance either amplifies the reduction in dissimilarity due to pegging or leaves it unchanged. Given this assumption, the numerator is non-negative.
$$\frac{\partial \Lambda^*_{ij}}{\partial d_{ij}} \geq 0$$
Conclusion: Stronger gravity frictions (higher distance) increase the optimal peg strength Λ^*_{ij}.

- Case 2: Common language ($g = L_{ij}$): Common language reduces dissimilarity ($\partial D_{ij}/\partial L_{ij} \leq 0$). The inequality of the cross-partial derivative reverses, $\frac{\partial^2 D_{ij}}{\partial \Lambda_{ij} \partial L_{ij}} \leq 0$. The numerator is non-positive.
$$\frac{\partial \Lambda^*_{ij}}{\partial L_{ij}} \leq 0$$
Conclusion: Gravity variables that reduce dissimilarity (like common language) decrease the optimal peg strength Λ^*_{ij}.

This establishes the comparative statics for the optimal peg choice. ■

Proposition 3.2 provides a direct link between the gravity determinants of portfolio similarity and the optimal design of currency arrangements: when gravity frictions make portfolios highly dissimilar under floating regimes, the marginal benefit of tighter pegs is larger and the welfare–maximising peg is stronger.

Figure 3 provides a graphical solution to the government's cooperative welfare problem described in Equation (21). The first–order condition, $\psi\big(-D'_{ij}(\Lambda_{ij})\big) = C'(\Lambda_{ij})$, determines the optimal bilateral exchange–rate peg strength, Λ^*_{ij}. The figure plots the marginal cost (MC) of tighter monetary integration against the marginal benefit (MB) from improved international risk sharing, and shows how gravity variables—here proxied by log bilateral distance d_{ij}—shift the optimal policy outcome.

The solid black line plots $C'(\Lambda_{ij})$, the marginal loss of monetary policy autonomy and the marginal adjustment cost associated with a stronger peg. The curve is upward–sloping, reflecting the fact that the opportunity cost of monetary independence rises as the peg approaches a hard-fix regime ($\Lambda_{ij} \to 1$).

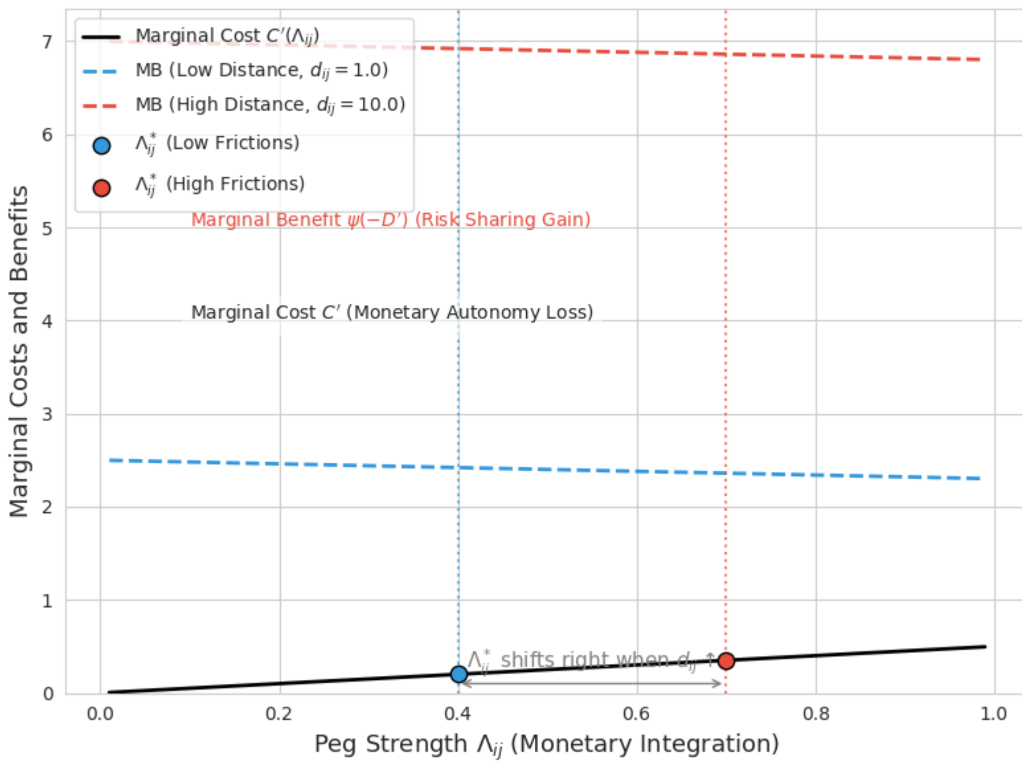

Figure 3: Optimal bilateral peg choice and gravity frictions

The dashed curves plot the marginal welfare gain $\psi(-D'_{ij})$ arising from further reductions in portfolio dissimilarity. Because D_{ij} is convex in Λ_{ij}, its derivative $D'_{ij}(\Lambda_{ij})$ becomes less negative as the peg tightens, implying that risk–sharing gains diminish at higher levels of monetary integration. Accordingly, the MB curves slope downward in Λ_{ij}.

The figure shows two MB curves corresponding to (i) a low–friction environment (blue dashed line, $d_{ij} = 1$) and (ii) a high–friction environment (red dashed line, $d_{ij} = 10$). Distance amplifies information frictions and increases the baseline level of portfolio dissimilarity. In terms of the derivative $D'_{ij}(\Lambda_{ij}, d_{ij})$, gravity frictions shift the MB curve upward: a larger d_{ij} increases the marginal welfare value of reducing portfolio dissimilarity via a stronger peg. As a result, the high–distance MB curve lies strictly above the low–distance one for all Λ_{ij}.

The equilibrium peg Λ^*_{ij} is located at the intersection point between the MB and MC curves. Under low frictions ($d_{ij} = 1$), the MB curve intersects MC at a relatively weak peg (blue circle). Under high frictions ($d_{ij} = 10$), the upward shift of the MB curve yields a strictly stronger optimal peg (red circle). This rightward movement of the optimal peg, Λ^*_{ij}, is a direct reflection of the cross–partial

$$\frac{\partial^2 D_{ij}}{\partial \Lambda_{ij}\, \partial d_{ij}} > 0,$$

which implies that distance increases the marginal risk–sharing value of tighter exchange–rate integration.

Figure 3 illustrates the central implication of Proposition 3.2: governments facing larger gravity frictions—such as greater geographic distance or weaker informational linkages—find it optimal to adopt a stronger bilateral peg. In such environments, the welfare gains from improved risk sharing outweigh the rising marginal cost of reduced monetary independence, generating a gravity-dependent theory of optimal currency arrangements.

3.4 Domestic risky assets, home bias and gravity Finally, we extend the asset menu to include a domestic risky asset in each investor country. Let $R_{F,h}$ denote the excess return on the foreign asset (as above) and $R_{H,h}$ the excess return on the domestic risky asset in country h, both measured relative to the local risk–free rate. Investor h chooses shares $\alpha_{F,h}$ and $\alpha_{H,h}$ such that the remainder $1 - \alpha_{F,h} - \alpha_{H,h}$ is invested in the risk–free bond.

We assume that the conditional means and variances are:

$$\mu_{F,h} := E_0[R_{F,h}], \quad \mu_{H,h} := E_0[R_{H,h}], \qquad \sigma^2_{F,h} := Var_0(R_{F,h}), \quad \sigma^2_{H,h} := Var_0(R_{H,h}),$$

and that domestic and foreign returns may be correlated with covariance $\sigma_{FH,h}$. For simplicity we take $\sigma^2_{H,h}$ and $\sigma_{FH,h}$ as exogenous and focus on the information asymmetry between foreign and domestic assets: investors learn about $\mu_{F,h}$ from noisy signals as in Section 2.5, whereas $\mu_{H,h}$ is known with higher precision (for instance because domestic informational frictions are smaller).

Under the quadratic approximation, investor h maximises:

$$(1-\gamma)\big(\alpha_{F,h}\mu_{F,h} + \alpha_{H,h}\mu_{H,h}\big) - \frac{\gamma(1-\gamma)}{2} Var_0\big(\alpha_{F,h}R_{F,h} + \alpha_{H,h}R_{H,h}\big).$$

The variance term is:

$$Var_0(\cdot) = \alpha^2_{F,h}\sigma^2_{F,h} + \alpha^2_{H,h}\sigma^2_{H,h} + 2\alpha_{F,h}\alpha_{H,h}\sigma_{FH,h}.$$

Lemma 5 (Optimal domestic and foreign shares)**.** *The optimal shares $(\alpha^*_{F,h}, \alpha^*_{H,h})$ solve the linear system:*

$$\begin{pmatrix} \sigma^2_{F,h} & \sigma_{FH,h} \\ \sigma_{FH,h} & \sigma^2_{H,h} \end{pmatrix} \begin{pmatrix} \alpha^*_{F,h} \\ \alpha^*_{H,h} \end{pmatrix} = \frac{1}{\gamma} \begin{pmatrix} \mu_{F,h} \\ \mu_{H,h} \end{pmatrix}. \tag{24}$$

In particular, if $|\sigma_{FH,h}| < \sigma_{F,h}\sigma_{H,h}$, the solution is unique and given by:

$$\alpha^*_{F,h} = \frac{1}{\gamma\Delta_h}\big(\mu_{F,h}\sigma^2_{H,h} - \mu_{H,h}\sigma_{FH,h}\big), \tag{25}$$

$$\alpha^*_{H,h} = \frac{1}{\gamma\Delta_h}\big(\mu_{H,h}\sigma^2_{F,h} - \mu_{F,h}\sigma_{FH,h}\big), \tag{26}$$

where $\Delta_h := \sigma^2_{F,h}\sigma^2_{H,h} - \sigma^2_{FH,h} > 0$ is the determinant of the covariance matrix.

Proof. The investor maximizes the objective function $f(\alpha_{F,h}, \alpha_{H,h})$ derived from the second-order mean-variance approximation:

$$f(\alpha_{F,h}, \alpha_{H,h}) = (1-\gamma)(\alpha_{F,h}\mu_{F,h} + \alpha_{H,h}\mu_{H,h}) - \frac{\gamma(1-\gamma)}{2} Var_0(\alpha_{F,h}R_{F,h} + \alpha_{H,h}R_{H,h})$$

The variance term $Var_0(\alpha_{F,h}R_{F,h} + \alpha_{H,h}R_{H,h})$ is expanded using the standard formula for the variance of a sum of two random variables:

$$Var_0(\dots) = \alpha_{F,h}^2\sigma_{F,h}^2 + \alpha_{H,h}^2\sigma_{H,h}^2 + 2\alpha_{F,h}\alpha_{H,h}\sigma_{FH,h}$$

The objective function is therefore:

$$f(\alpha_{F,h}, \alpha_{H,h}) = (1-\gamma)(\alpha_{F,h}\mu_{F,h} + \alpha_{H,h}\mu_{H,h}) - \frac{\gamma(1-\gamma)}{2}(\alpha_{F,h}^2\sigma_{F,h}^2 + \alpha_{H,h}^2\sigma_{H,h}^2 + 2\alpha_{F,h}\alpha_{H,h}\sigma_{FH,h})$$

1. First-order conditions (FOCs): To find the optimal shares $(\alpha_{F,h}^*, \alpha_{H,h}^*)$, we take the partial derivatives with respect to each share and set them to zero. We assume $\gamma \neq 1$.

FOC with respect to $\alpha_{F,h}$:

$$\frac{\partial f}{\partial \alpha_{F,h}} = (1-\gamma)\mu_{F,h} - \frac{\gamma(1-\gamma)}{2}(2\alpha_{F,h}\sigma_{F,h}^2 + 2\alpha_{H,h}\sigma_{FH,h}) = 0$$

Dividing by $(1-\gamma)$ and simplifying:

$$\mu_{F,h} = \gamma(\alpha_{F,h}\sigma_{F,h}^2 + \alpha_{H,h}\sigma_{FH,h})$$

$$\frac{1}{\gamma}\mu_{F,h} = \sigma_{F,h}^2\alpha_{F,h}^* + \sigma_{FH,h}\alpha_{H,h}^* \quad (FOC_F)$$

FOC with respect to $\alpha_{H,h}$:

$$\frac{\partial f}{\partial \alpha_{H,h}} = (1-\gamma)\mu_{H,h} - \frac{\gamma(1-\gamma)}{2}(2\alpha_{H,h}\sigma_{H,h}^2 + 2\alpha_{F,h}\sigma_{FH,h}) = 0$$

Dividing by $(1-\gamma)$ and simplifying:

$$\mu_{H,h} = \gamma(\alpha_{H,h}\sigma_{H,h}^2 + \alpha_{F,h}\sigma_{FH,h})$$

$$\frac{1}{\gamma}\mu_{H,h} = \sigma_{FH,h}\alpha_{F,h}^* + \sigma_{H,h}^2\alpha_{H,h}^* \quad (FOC_H)$$

2. Matrix representation: We express (FOC_F) and (FOC_H) as a system of linear equations

$A\mathbf{x} = \mathbf{b}$:

$$\begin{pmatrix} \sigma_{F,h}^2 & \sigma_{FH,h} \\ \sigma_{FH,h} & \sigma_{H,h}^2 \end{pmatrix} \begin{pmatrix} \alpha_{F,h}^* \\ \alpha_{H,h}^* \end{pmatrix} = \frac{1}{\gamma} \begin{pmatrix} \mu_{F,h} \\ \mu_{H,h} \end{pmatrix}$$

This confirms Equation (24). The matrix on the left is the covariance matrix Σ_h.

3. Inversion and uniqueness: The solution is unique if the covariance matrix is positive definite, which for a 2×2 matrix requires: i) $\sigma_{F,h}^2 > 0$ and $\sigma_{H,h}^2 > 0$; ii) The determinant Δ_h must be positive:

$$\Delta_h = \sigma_{F,h}^2 \sigma_{H,h}^2 - \sigma_{FH,h}^2 > 0$$

This is equivalent to the requirement $|\sigma_{FH,h}| < \sigma_{F,h}\sigma_{H,h}$, meaning the correlation is less than one in magnitude, ensuring the assets are not perfectly correlated.

Assuming $\Delta_h > 0$, we invert the covariance matrix Σ_h:

$$\Sigma_h^{-1} = \frac{1}{\Delta_h} \begin{pmatrix} \sigma_{H,h}^2 & -\sigma_{FH,h} \\ -\sigma_{FH,h} & \sigma_{F,h}^2 \end{pmatrix}$$

Multiplying by the mean vector $\frac{1}{\gamma}\mu_h$:

$$\begin{pmatrix} \alpha_{F,h}^* \\ \alpha_{H,h}^* \end{pmatrix} = \frac{1}{\gamma \Delta_h} \begin{pmatrix} \sigma_{H,h}^2 & -\sigma_{FH,h} \\ -\sigma_{FH,h} & \sigma_{F,h}^2 \end{pmatrix} \begin{pmatrix} \mu_{F,h} \\ \mu_{H,h} \end{pmatrix}$$

Performing the matrix multiplication yields the explicit solutions:

$$\alpha_{F,h}^* = \frac{1}{\gamma \Delta_h} (\mu_{F,h} \sigma_{H,h}^2 - \mu_{H,h} \sigma_{FH,h}) \quad \text{(Confirms Equation (25))}$$

$$\alpha_{H,h}^* = \frac{1}{\gamma \Delta_h} (\mu_{H,h} \sigma_{F,h}^2 - \mu_{F,h} \sigma_{FH,h}) \quad \text{(Confirms Equation (26))}$$

■

We now define home bias as the excess weight on the domestic asset relative to the foreign one, $HB_h := \alpha_{H,h}^* - \alpha_{F,h}^*$.

Proposition 3.3 (Home bias and gravity)**.** *Suppose* $\mu_{F,h} = \mu_{H,h} = \mu > 0$ *and* $\sigma_{F,h}^2 = \sigma_{H,h}^2$, *but the investor has noisier information about foreign than home returns so that the perceived variance of $\mu_{F,h}$ exceeds that of $\mu_{H,h}$. Then $HB_h > 0$ (home bias). Moreover, if gravity variables (distance, lack of common language) increase the signal noise for foreign returns but not for home returns, then $\partial\alpha^*_{F,h}/\partial d_{ij} < 0$ and $\partial HB_h/\partial d_{ij} > 0$: greater distance both strengthens home bias and reduces cross–country similarity in the foreign share.*

Proof. 1. Deriving optimal shares under information asymmetry: In a Bayesian mean-variance framework, noisier information about the foreign asset return mean $\mu_{F,h}$ (due to larger prior variance $Var(\mu_{F,h})$) effectively increases its total perceived risk compared to the domestic asset. This allows us to replace the true variance $\sigma^2_{F,h}$ with an effectively higher perceived variance $\tilde{\sigma}^2_{F,h}$ in Lemma 5, such that $\tilde{\sigma}^2_{F,h} > \sigma^2_{H,h}$.

Setting $\mu_{F,h} = \mu_{H,h} = \mu$ and $\sigma_{FH,h} = 0$, the determinant is $\Delta_h = \tilde{\sigma}^2_{F,h}\sigma^2_{H,h}$. The optimal shares from Equations (25) and (26) (Lemma 5) simplify to:

$$\alpha^*_{F,h} = \frac{1}{\gamma\Delta_h}(\mu\sigma^2_{H,h} - \mu\cdot 0) = \frac{\mu\sigma^2_{H,h}}{\gamma(\tilde{\sigma}^2_{F,h}\sigma^2_{H,h})} = \frac{\mu}{\gamma\tilde{\sigma}^2_{F,h}}$$
$$\alpha^*_{H,h} = \frac{1}{\gamma\Delta_h}(\mu\tilde{\sigma}^2_{F,h} - \mu\cdot 0) = \frac{\mu\tilde{\sigma}^2_{F,h}}{\gamma(\tilde{\sigma}^2_{F,h}\sigma^2_{H,h})} = \frac{\mu}{\gamma\sigma^2_{H,h}}$$

2. Establishing home bias ($HB_h > 0$): Home bias is defined as the excess weight on the domestic asset: $HB_h := \alpha^*_{H,h} - \alpha^*_{F,h}$.

$$HB_h = \frac{\mu}{\gamma\sigma^2_{H,h}} - \frac{\mu}{\gamma\tilde{\sigma}^2_{F,h}}$$

Since we assume noisier foreign information implies $\tilde{\sigma}^2_{F,h} > \sigma^2_{H,h}$, it follows that $\frac{1}{\sigma^2_{H,h}} > \frac{1}{\tilde{\sigma}^2_{F,h}}$. Therefore, $\alpha^*_{H,h} > \alpha^*_{F,h}$, and thus $HB_h > 0$. The investor overweights the asset for which they have higher informational precision (the domestic asset).

3. Comparative statics with respect to distance (d_{ij}): Gravity variables, such as distance d_{ij}, increase the signal noise variance for foreign returns $\sigma^2_{u,h}$ (Equation (8)), which raises the perceived variance of foreign returns $\tilde{\sigma}^2_{F,h}$. Thus, we have the assumption $\frac{\partial\tilde{\sigma}^2_{F,h}}{\partial d_{ij}} > 0$.

We differentiate the optimal foreign share $\alpha^*_{F,h}$ with respect to d_{ij}:

$$\frac{\partial \alpha^*_{F,h}}{\partial d_{ij}} = \frac{\partial}{\partial d_{ij}} \left(\frac{\mu}{\gamma \tilde{\sigma}^2_{F,h}} \right)$$

Using the chain rule, $\frac{d}{dx}\frac{1}{u} = -\frac{1}{u^2}\frac{du}{dx}$:

$$\frac{\partial \alpha^*_{F,h}}{\partial d_{ij}} = \frac{\mu}{\gamma} \left(-\frac{1}{(\tilde{\sigma}^2_{F,h})^2} \frac{\partial \tilde{\sigma}^2_{F,h}}{\partial d_{ij}} \right)$$

Since $\mu > 0$, $\gamma > 0$, $\tilde{\sigma}^2_{F,h} > 0$, and we assume $\frac{\partial \tilde{\sigma}^2_{F,h}}{\partial d_{ij}} > 0$:

$$\frac{\partial \alpha^*_{F,h}}{\partial d_{ij}} < 0$$

Conclusion 1: Greater distance strictly reduces the optimal foreign asset share $\alpha^*_{F,h}$.

4. Effect on home bias: Since $\alpha^*_{H,h}$ is independent of d_{ij} in this simplified model, the effect of distance on home bias $HB_h = \alpha^*_{H,h} - \alpha^*_{F,h}$ is given by:

$$\frac{\partial HB_h}{\partial d_{ij}} = \frac{\partial \alpha^*_{H,h}}{\partial d_{ij}} - \frac{\partial \alpha^*_{F,h}}{\partial d_{ij}} = 0 - \frac{\partial \alpha^*_{F,h}}{\partial d_{ij}}$$

Since $\frac{\partial \alpha^*_{F,h}}{\partial d_{ij}} < 0$, it follows that:

$$\frac{\partial HB_h}{\partial d_{ij}} > 0$$

Conclusion 2: Greater distance strengthens home bias (HB_h increases).

An analogous argument holds for common language L_{ij}: $L_{ij} \uparrow \implies \sigma^2_{u,h} \downarrow \implies \tilde{\sigma}^2_{F,h} \downarrow$. This increases $\alpha^*_{F,h}$ and decreases HB_h. ■

Proposition 3.3 shows that the same gravity variables that drive cross–country similarity in foreign portfolios in the baseline model also shape home bias and the relative composition of domestic versus foreign assets. This provides an integrated theoretical framework for interpreting the empirical role of distance and other bilateral frictions in both dimensions.

4 Conclusion

This paper develops a unified theoretical framework to explain why investors located in different countries choose systematically different foreign portfolio allocations despite having access to the same global asset menu. By embedding gravity-type information frictions and bilateral exchange-rate stabilization into an analytically tractable Bayesian portfolio-choice environment, we derive a set of gravity-consistent predictions for the cross-country similarity of foreign-asset shares. Two central mechanisms emerge. The first is a risk-reduction channel: stronger bilateral exchange-rate pegs reduce idiosyncratic currency risk, align the stochastic environments faced by investors, and thereby compress the distance between optimal portfolio choices. The second is an information channel: gravity variables such as distance, common language, and historical ties affect signal precision and correlation, generating persistent heterogeneity in posterior beliefs and portfolio weights.

We extend the baseline environment in several important directions. The multivariate extension shows that the entire portfolio vector—not only the allocation to a single foreign asset—responds systematically to the covariance structure of returns and to bilateral information frictions. The dynamic Bayesian learning model demonstrates that the long-run variance of portfolio differences retains the comparative statics of the static environment, while introducing transitional dynamics that can be calibrated to time-series patterns in portfolio similarity. Finally, the welfare-based extension endogenizes the exchange-rate regime: when governments trade off monetary-policy independence against the risk-sharing value of greater portfolio convergence, the optimal peg strength depends directly on the magnitude of gravity frictions. Countries facing larger informational or geographic barriers optimally select deeper bilateral monetary integration.

The model yields several policy-relevant insights. First, the interaction of gravity frictions and exchange-rate stabilization provides a theoretical foundation for why regional monetary arrangements—such as currency unions, hard pegs, or tightly managed bilateral bands—produce more homogeneous cross-border investment portfolios. Stabilization policies that reduce idiosyncratic currency risk can be viewed as an instrument for increasing bilateral risk sharing. Second, the model implies that financial integration policies may be more effective when targeted toward country pairs with preexisting informational proximity: removing barriers to information flows can substantially amplify the portfolio-convergence effects of monetary integration. Third, by endogenizing the optimal degree of exchange-rate stabilization, the framework provides a welfare-based rationale for deeper bilateral monetary cooperation among countries with strong real and informational linkages. Conversely, in country pairs characterized by severe information frictions or large geographic barriers, looser exchange-rate arrangements may be optimal despite the resulting divergence in portfolio positions.

Overall, the theory developed in this paper offers a tractable microfoundation for understanding bilateral patterns in foreign portfolio composition and their interaction with monetary regimes and structural frictions. Future work may extend the framework to incorporate heterogeneous investors, intermediary constraints, or global-cycle shocks, allowing the gravity-consistent portfolio framework to inform broader debates on international risk sharing and the design of monetary arrangements.

References

Bekaert, G., Hodrick, R.J. & Zhang, X. (2009). International stock return comovements. *Journal of Finance*, 64(6), 2591–2626.

Brennan, M.J. & Cao, H.H. (1997). International portfolio investment flows. *Journal of Finance,* 52(5), 1851–1880.

Coeurdacier, N. & Gourinchas, P.-O. (2016). When bonds matter: Home bias in goods and assets. *Journal of Monetary Economics*, 82, 119-137.

Devereux, M.B. & Engel, C. (2002). Exchange rate pass-through, exchange rate volatility, and exchange rate disconnect. *Journal of Monetary Economics*, 49(5), 913–940.

Exchange rates and interest parity. In G. Gopinath, E. Helpman & K. Rogoff (Eds.), *Handbook of International Economics*, 4, 453–522.

Forbes, K. J. & Warnock, F. E. (2012). Capital flow waves: Surges, stops, flight, and retrenchment. *Journal of International Economics*, 88(2), 235–251.

Gehrig, T. (1993). An information based explanation of the domestic bias in international equity investment. *Scandinavian Journal of Economics*, 95(1), 97-109.

Gopinath, G. & Itskhoki, O. (2011). In search of real rigidities. *NBER Macroeconomics Annual 2010*, 25, 195-263.

Lane, P. R. & Milesi-Ferretti, G. M. (2008). The drivers of financial globalization. *American Economic Review: Papers & Proceedings*, 98(2), 327-332.

Martin, P. & Rey, H. (2004). Financial super-markets: Size matters for asset trade. *Journal of International Economics*, 64(2), 335-361.

Obstfeld, M. & Rogoff, K. (1995). The intertemporal approach to the current account. *Handbook of International Economics*, Vol. 3, 1731–1799.

Okawa, Y. & van Wincoop, E. (2012). Gravity in international finance. Journal of International Economics, 87(2), 205-215.

Pan, L., Hu, R. & Du, Q. (2022). Foreign portfolio investment patterns: evidence from a gravity model. *Empirical Economics*, 63, 391-415.

Portes, R. & Rey, H. (2005). The determinants of cross-border equity flows. *Journal of International Economics*, 65(2), 269–296.

Van Nieuwerburgh, S. & Veldkamp, L. (2009). Information immobility and the home bias puzzle. *Journal of Finance*, 64(3), 1187-1215.