International Investing: Diversification and Beyond *

Soohun Kim[†] Robert A. Korajczyk[‡] Andreas Neuhierl[§]

November 18, 2025

*We thank Sanjiv Das, Nicolae Gârleanu, Zhiyu Fu, Gustavo Schwenkler, Sergio Rebelo, Josef Zechner and seminar and conference participants at Queen Mary University of London, Santa Clara University, Sungkyunkwan University, Washington University in St. Louis, Wirtschaftsuniversität Wien and Fifth International Workshop in Financial Econometrics at Brazil, Exploring the Frontiers of Financial Econometrics in the Big Data Era at Seoul, and 4th Frontiers of Factor Investing 2024 for useful comments.

[†]College of Business, Korea Advanced Institute of Science and Technology
[‡]Kellogg School of Management, Northwestern University
[§]Mitch Daniels School of Business, Purdue University

International Investing Beyond Diversification

Abstract

We develop a framework for analyzing individual stocks across two countries to uncover international investment gains arising from (i) differences in global factor risk premia and (ii) country-specific factors unavailable domestically. Applying the method to individual stocks in 27 countries from 1966 to 2022, we find large and robust opportunities. Investing in country-specific factor portfolios more than triple the Sharpe ratio of a US market-index investor, and strategies that exploit cross-country differences in common factor premia are similarly profitable – even among G7 markets – challenging the view of completely integrated global equity markets.

JEL classification codes: G12, G15, C58, C38

Keywords: arbitrage, market integration, characteristic based model, factor pricing

1 Introduction

Introductory textbooks teach that investors should diversify internationally. Holding assets across countries is supposed to reduce portfolio risk and improve risk-adjusted returns because national economies do not move in perfect lockstep. This advice has guided both practitioners and academics for decades and is supported by early empirical evidence showing low cross-country correlations in stock returns. For example, Heston and Rouwenhorst (1994) finds that "the low correlation between country indices is almost completely due to country-specific sources of return variation." Yet, as global markets have become increasingly integrated and many companies engage in foreign direct investments, the classical case for international diversification has become less clear. When information, capital, and risk flow freely across borders, can investors still improve their portfolios by looking abroad – or has globalization erased most of these gains?

This paper develops a framework to examine these questions systematically. Consider a baseline scenario where identical systematic factors drive individual asset returns in both countries, and factor pricing holds. In such a case, investing abroad does not offer additional benefits asymptotically, i.e., when the investor already holds a very large and well-diversified portfolio of domestic assets.

Our framework departs from this baseline in two key respects: (1) while some systematic factors may be common across countries, others may be country-specific; and (2) the compensation for common factors may vary across borders. We demonstrate how to detect these deviations and translate them into actionable investment strategies.

In a globalized economy, common factors across countries are almost inevitable. For instance, over our sample period, the correlation between the US market index and the UK (Canada) market index is 0.69 (0.72). As a result, mean-variance enhancement is rarely achieved by simply combining country-level indices. Thus, isolating country-specific risks is essential to expand the investment opportunity set.

To address this, we extend the setup of the large-N, short-T economy of Kim et al. (2021) to a two-country model, enabling full exploitation of a large number of individual stocks in each country. We further assume that firm characteristics contain valuable information about factor loadings, allowing for efficient estimation with short panel data. We begin by applying Principal Component Analysis (PCA) to each country to extract systematic factors, which initially represent a mix of common and country-specific factors. To disentangle these, we leverage the fact that common factors should appear in both sets of systematic factors. Specifically, we employ Canonical Correlation Analysis (CCA) to identify combinations of systematic factors that maximize cross-country co-movement, and we theoretically establish that this method consistently recovers the common factors and the country-specific factors. For a given pair of countries, we identify factors that are uncorrelated to the common factors so that they can be interpreted as country-specific sources of risk that are only accessible by investing in a particular country.

Additionally, we develop a feasible portfolio strategy to exploit differences in the risk premia of common factors across countries if markets are not completely integrated. For example, suppose the US and Canadian market indices perfectly comove. If the US market index has a higher risk premium, then regression of the US index on a constant and the Canadian index should yield a positive intercept with an R^2 of 1. This positive intercept reflects the difference in risk premia between the US and Canada and can be exploited through a long-short strategy: long the US market index and short the Canadian index. Similarly, we regress common factors identified from one country's assets on a constant and the corresponding common factors from another country to detect disparities in risk premia. If such a strategy were to deliver consistent profits, the economic implications are profound. It would indicate market segmentation in terms of the pricing of common risk. We view this as a stronger challenge to market integration than the mere existence of country-specific factors.

We apply our method to individual assets across 27 countries, using up to 153 firm characteristics over the sample period from 1966 to 2022. We use the code from Jensen et al. (2023) (available at https://jkpfactors.com) to construct our dataset. We analyze 26 country pairs, each consisting of the US and one other country.

As a baseline, we use rolling estimation windows of 60 months, assume seven systematic factors, and separate the common factors from the country-specific factors using the bootstrapping method of Gonçalves et al. (2025). In each estimation window, we construct two types of portfolios: (1) country-specific portfolios, which bet on country-specific factors, and (2) a segmentation portfolio, which exploits differences in risk premia while hedging out common factor risks. These portfolios are implemented in an out-of-sample manner, with positions held for one month following each estimation window.

Our analysis shows that country-specific portfolios deliver impressive performance, with annualized Sharpe ratios ranging from 0.50 (Brazil) to 1.31 (Belgium). For most of 26 pairs, these portfolios generate highly profitable returns at the 1% statistical significance level. Moreover, they exhibit remarkably low correlations with the US market and with each other, offering substantial diversification benefits compared to traditional passive diversification using market indices. While combining country-level market portfolios yields little improvement, mixing the US market with country-specific portfolios can nearly quadruple the Sharpe ratio relative to the US market portfolio.

Segmentation portfolios also deliver strong performance, with returns quantitatively comparable to those of country-specific portfolios. Notably, we find high profitability in segmentation portfolios even between the United States and other G7 countries, casting serious doubt on the conventional view of tight financial market integration among developed economies. Furthermore, we show that the profitability of the segmentation portfolio is not driven by country-level or global exchange rate movements.

Naive international diversification over our sample period yields Sharpe ratios between 0.48 and 0.56. Our approach to exploiting country-specific risk premia yields Sharpe ratios from 1.44 to 2.32. Our approach to exploiting pairwise cross-country differences in the pricing of common risks yields Sharpe ratios from 0.73 to 1.41 for G7 countries and from 0.57 to 1.33 for non-G7 countries.

Related Literature

A central motivation for international investment has long been the potential for diversification benefits arising from low correlations across national markets. Early studies such as Grubel (1968) and Solnik (1974) documented substantial gains from holding internationally diversified portfolios, attributing the benefits to the low correlations of financial markets across countries and the presence of country-specific economic cycles. As this insight spread and old barriers to capital flows slowly crumbled, global investment flows surged. This shift contributed not only to the growth of cross-border financial investment but also to the economic development of capital-scarce countries, as Bekaert and Harvey (1995) and Bekaert et al. (2005).

However, asset behavior changes once markets become integrated into the global financial system. For example, Karolyi and Stulz (2003) show that the pricing of cross-listed firms is strongly influenced by global factors, indicating that even firms based in relatively independent domestic economies are subject to international pricing dynamics. Similarly, Bekaert and Harvey (1995) and Carrieri et al. (2007) find that the covariance between national markets and the global market accounts for a significant portion of expected returns, particularly in developed markets. As a result, several studies, including Sinquefield (1996) and Quinn and Voth (2008), question the diversification benefits of international investment within developed markets. In response, Berger et al. (2011) suggest turning to frontier markets as a more promising source of international diversification gains.

We find that skepticism regarding the benefits of international investment is, at least in part, due to a lack of attention to individual assets in foreign markets. Cho et al. (1986) use individual assets to study market integration and find that the data are not consistent with the joint hypothesis of integration and exact factor pricing. Chaieb et al. (2021) also analyze the behavior of individual foreign stocks and highlight the importance of local factors that can be overlooked in standard country indices. Taking a different approach, Eun et al. (2017) demonstrate how selecting stocks with low exposure to global factors can yield meaningful diversification benefits. Our paper is also related to the works of Errunza et al. (1999) and Bae et al. (2019), which explore hidden investment opportunities by examining granular assets within the U.S. or other developed markets. We contribute to this literature by proposing a systematic approach to isolate country-specific factors, that are independent of global influences, from a broad set of individual assets and, therefore, enhance diversification potential.

We propose a novel method to assess the parity of risk premia on global factors across countries, advancing the understanding of international market equilibrium. Cho et al. (1986), Korajczyk and Viallet (1989), and Korajczyk and Viallet (1992) apply the Arbitrage Pricing Theory (APT) framework to international markets, while Harvey (1991) extends the Capital Asset Pricing Model (CAPM) to the global setting - a framework further developed by Dumas and Solnik (1995) to incorporate currency risk. Our approach builds on a country-level factor structure, which enables the application of APT on a country-by-country basis. By allowing risk premia to differ across countries, we uncover deviations from market integration and equal compensation for common risk between markets. In a similar vein, Patton and Weller (2022) explore cross-country heterogeneity in factor compensation using clustering techniques. This concept of evaluating equilibrium through risk premia parity builds on foundational work by Roll and Ross (1980) and Brown and Weinstein (1983).

Recent efforts to model the factor structure of global financial markets have been advanced by Choi and Kim (2023) and Linton et al. (2025). Furthermore, Sandulescu and Schneider (2021) and Sandulescu et al. (2021) provide evidence of disequilibrium in global asset markets, suggesting the presence of unexploited investment opportunities in international settings. We contribute to this line of research by proposing a factor pricing model based on large panels of individual asset data. The core of our method is Projected-PCA, developed by Fan et al. (2016) and extended by Kim et al. (2021) to identify mispricing in a large cross-sectional economy. We expand upon the single-market analysis in Kim et al. (2021) by extending it to a two-market framework. Additionally, we apply factor grouping techniques from Andreou et al. (2019) and Andreou et al. (2020), utilizing canonical correlation analysis in a short T setup. More broadly, our work draws on the literature that estimates risk premia and systematic factors in short panel data settings (Connor and Korajczyk (1988), Kim and Skoulakis (2018), Raponi et al. (2020), Kim and Korajczyk (2024), Zaffaroni (2025), and Fortin et al. (2025)), as well as on the literature that infers risk premia through firm characteristics (Freyberger et al. (2020), Freyberger et al. (2024), and Kelly et al. (2019)).

2 The Model

2.1 Motivating Example

We first describe a simple version of our setup, in which we abstract from estimation problems, and only consider a one-factor economy. The intuition will carry over to the general case, but the general development requires more notation. Throughout, we will consider an investor with a home country (country 1) who is considering investing in another country, country 2. Excess returns follow a factor model for both countries, and more specifically for country $j = 1, 2$, we have:

$$R_{jit} = \beta_{ji} \underbrace{\left(\lambda_2^c + f_t\right)}_{\text{common risk}} + \delta_{ji} \underbrace{\left(\lambda_j^s + g_{jt}\right)}_{\text{country specific}} + e_{jit}, \; j = 1, 2; i = 1, \cdots, N. \tag{2.1}$$

That is, individual asset returns follow a three-factor model. The first factor, f_t, captures risks that are common to both countries. We denote its risk premium by λ_2^c. The second and third factors, g_{jt}, represent country-specific risks, i.e. shocks to which firms in a country are exposed; its risk premium is denoted by λ_j^s. To further simplify this example, we assume that the factor loadings β_{ji} and δ_{ji} are known and they are cross-sectionally orthogonal, i.e., $\beta_{ji} \perp \delta_{ji}$.

We now outline the portfolios of interest in this paper. The first one is called the "country-specific" portfolio. It seeks to gain exposure to sources of systematic risk that are only present in the foreign country (country 2). Consider the portfolio, with weights (for assets in country 2), $w_i^s = \frac{\delta_{2i}}{N}$:

$$\begin{aligned}
\sum_{i=1}^{N} w_i^s R_{2it} &= \sum_{i=1}^{N} \frac{\delta_{2i}\beta_{2i}}{N} \left(\lambda_2^c + f_t\right) + \sum_{i=1}^{N} \frac{\delta_{2i}^2}{N} \left(\lambda_2^s + g_{2t}\right) + \sum_{i=1}^{N} \frac{\delta_{2i} e_{2it}}{N} \\
&\to^P \left(\lim_{N \to \infty} \sum_{i=1}^{N} \frac{\delta_{2i}^2}{N} \right) \left(\lambda_2^s + g_{2t}\right),
\end{aligned}$$

where the limit follows from the orthogonality between β_{2i} and δ_{2i} (eliminating the first term) and from the diversification of the idiosyncratic component (eliminating the last term). This portfolio expands the investment opportunity set for our country 1 investor.

In addition, our investor can attempt to exploit pricing differences in the pricing of the common factor across the two countries. In this setting consider the following portfolio, which we dub the "segmentation portfolio". Portfolio weights $w_i^c = \frac{\beta_{2i}}{N}$ will isolate the common factor with the associated premia in country 2:

$$\sum_{i=1}^{N} w_i^c R_{2it} = \sum_{i=1}^{N} \frac{\beta_{2i}^2}{N} (\lambda_2^c + f_t) + \sum_{i=1}^{N} \frac{\beta_{2i}\delta_{2i}}{N} (\lambda_2^s + g_{2t}) + \sum_{i=1}^{N} \frac{\beta_{2i}e_{2it}}{N}$$
$$\rightarrow^P \left(\lim_{N\to\infty} \sum_{i=1}^{N} \frac{\beta_{2i}^2}{N} \right) (\lambda_2^c + f_t)$$

because the factor loading for the common factor β_{2i} is orthogonal to that for the country specific factor δ_{2i}, and the residuals are diversified away. An analogous portfolio in country 1 recovers the equivalent quantity, $\rightarrow^P \left(\lim_{N\to\infty} \sum_{i=1}^{N} \frac{\beta_{1i}^1}{N} \right) (\lambda_1^c + f_t)$. The difference in these two portfolios converges to the difference in common risk premia, i.e. λ_1^c - λ_2^c. The sign of the premia difference allows a portfolios manager to devise a strategy to exploit any non-zero difference.

This simple approach relies on (at least) two unrealistic assumptions. First, it assumes that investors know the true values of factor loadings, as well as common and country-specific factors. Second, it assumes that common and country-specific factors can be easily identified through cross-sectional orthogonality of factor loadings. Both assumptions are very likely to be violated in practice. Instead, the investor faces a difficult estimation problem. The main goal of this paper is to develop feasible estimators for these quantities so that the "country specific" and "segmentation portfolio" can be implemented.

2.2 General Model

Consider a pair of two countries, indexed by $g = 1, 2$.[1] We assume that there exists a large number of securities in each country and the return generating processes for those individual securities are stable over a relatively short horizon $t = 1, \cdots, T$.

We specify the return generating process of individual securities in each country g. The excess returns of individual stocks in country g follow a K_g-factor model in which the factors are unobservable, latent factors. Among K_g factors, K^c factors are common across the two countries and K_g^s factors are country-specific. Thus, it holds that $K_g = K^c + K_g^s$. In particular, the excess return of i-th asset in country at time t

[1]These two countries are generic. In the empirical application in Section 4, we construct various pairs using 27 countries.

is generated by the following model: for $g = 1, 2,\ i = 1, \cdots, N_g$ and $t = 1, \cdots, T$,

$$R_{git} = \boldsymbol{\beta}'_{gi} \left(\boldsymbol{\lambda}^c_g + \mathbf{f}_t \right) + \boldsymbol{\delta}'_{gi} \left(\boldsymbol{\lambda}^s_g + \mathbf{g}_{gt} \right) + e_{git}, \tag{2.2}$$

where $\boldsymbol{\beta}'_{gi} = [\beta_{gi1} \ \cdots \ \beta_{giK^c}]'$ is the $(K^c \times 1)$ factor loadings of the i-th asset to the common factors, $\boldsymbol{\lambda}^c_g = \left[\lambda^c_{g1} \ \cdots \ \lambda^c_{gK^c}\right]'$ is the $(K^c \times 1)$ vector of risk premia in country g for exposure to common factors, $\mathbf{f}_t = [f_{g1} \ \cdots \ f_{gK^c}]'$ is the $(K^c \times 1)$ systematic zero-mean common factor realization in period t, $\boldsymbol{\delta}'_{gi} = \left[\delta_{gi1} \ \cdots \ \delta_{giK^s_g}\right]'$ is the $\left(K^s_g \times 1\right)$ factor loadings of the i-th asset to the country-specific factors in country g, $\boldsymbol{\lambda}^s_g = \left[\lambda^s_{g1} \ \cdots \ \lambda^s_{gK^s_g}\right]'$ is the $\left(K^s_g \times 1\right)$ vector of risk premium in country g on the exposure to country-specific systematic factors, $\mathbf{g}_{gt} = \left[g_{g1} \ \cdots \ g_{gK^s_g}\right]'$ is the $\left(K^s_g \times 1\right)$ systematic zero-mean country-specific factor realization in period t, and e_{git} is the zero-mean idiosyncratic return of asset i at time t.

Throughout, we use $\mathbf{0}_m$, $\mathbf{1}_m$, and $\mathbf{0}_{m \times l}$ denote the $(m \times 1)$ vectors of zeros and ones and the $(m \times l)$ matrix of zeros, respectively. The return generating process of (2.2) is expressed compactly in matrices: for $g = 1, 2$,

$$\mathbf{R}_g = \underbrace{\mathbf{B}_g \left(\boldsymbol{\lambda}^c_g \mathbf{1}'_T + \mathbf{F}' \right)}_{\text{common factors}} + \underbrace{\mathbf{D}_g \left(\boldsymbol{\lambda}^s_g \mathbf{1}'_T + \mathbf{G}'_g \right)}_{\text{country-specific factors}} + \mathbf{E}_g, \tag{2.3}$$

where the (i, t) element of the $(N_g \times T)$ matrix $\mathbf{R}_g$ is R_{git}, the i-th row of the $(N_g \times K^c)$ matrix $\mathbf{B}_g$ is $\boldsymbol{\beta}'_{gi}$, the t-th row of the $(T \times K^c)$ matrix $\mathbf{F}$ is $\mathbf{f}'_t$, the i-th row of the $\left(N_g \times K^s_g\right)$ matrix $\mathbf{D}_g$ is $\boldsymbol{\delta}'_{gi}$, the t-th row of the $\left(T \times K^s_g\right)$ matrix $\mathbf{G}_g$ is $\mathbf{g}'_{gt}$, and the (i, t) element of the $(N_g \times T)$ matrix $\mathbf{E}_g$ is e_{git}.

The goal of this paper is to demonstrate how an investor can systematically uncover and exploit investment opportunities in a foreign country. The first two terms on the right-hand side of (2.3) provide distinct avenues for expanding the investment opportunity set. To illustrate, let us consider an investor based in country 1 seeking investment opportunities in country 2. Hence, we set $g = 2$ in (2.3) and dissect the avenues available to her as follows.

1. Common factors:

   The first term relates to common risk factors, i.e. those shared between countries 1 and 2. At first glance, leveraging these factors in country 2 might seem redundant, as she already has exposure to them in her domestic market. However, when the pricing of these factors differs between the two countries, the investor may attempt to take advantage of this pricing differential. If $\boldsymbol{\lambda}_1^c < \boldsymbol{\lambda}_2^c$, investing in the common factor via country 2 is more attractive than via country 1. She might further hedge her exposure to the common factors by shorting the common factor in her domestic market. Conversely, if $\boldsymbol{\lambda}_1^c > \boldsymbol{\lambda}_2^c$, she may invest in the common factor via country 1 and still short the common factors in country 2 for the purpose of hedging. This strategy, however, induces capital flows across borders, potentially equalizing risk premia of common factors over time. Thus, the viability of this strategy hinges on timely detection of cross-country disparities or the presence of frictions that inhibit capital mobility.

2. Country-specific factors:
   The second term of RHS in (2.3) highlights the country-specific factors within the country 2. This channel represents the traditional justification for international investing: expanding the mean-variance frontier. By including foreign assets, she gains exposure to factors that are not available in her domestic market. We aim to move beyond the conventional approach. This paper proposes a systematic framework to isolate and precisely target investments driven purely by country-specific factors. This approach not only enhances the investor's ability to capitalize on international opportunities but also underscores the potential for more refined and targeted global strategies.

While the conceptual pathways are clear, implementing these strategies in practice poses challenges. Central to these challenges is the lack of precise knowledge regarding the exposures to the systematic factors, $[\mathbf{B}_g \; \mathbf{D}_g]$. Once investors obtain these information, the implementation is relatively clear. To illustrate, consider a simplified scenario where mispricing does not exist, and only the common factors are present without country-specific factors. Then, if $\mathbf{B}_g$ were known, one could estimate $\boldsymbol{\lambda}_g^c \mathbf{1}_T' + \mathbf{F}'$ through straightforward cross-sectional regressions in each country under reasonable assumptions.[2] Then, the difference between $\boldsymbol{\lambda}_1^c$ and $\boldsymbol{\lambda}_2^c$ could be derived from the difference in the estimated $\boldsymbol{\lambda}_g^c \mathbf{1}_T' + \mathbf{F}'$ because $\mathbf{F}'$, the random realization of common factors,

[2]The object of $\boldsymbol{\lambda}_g^c \mathbf{1}_T' + \mathbf{F}'$ is closely related to the concept of expost risk premia discussed in Shanken (1992) and Kim and Skoulakis (2018)

would cancel out. Armed with this information, the investor could systematically exploit the second pathway, strategically allocating long and short positions in common factors across the two markets.

The bulk of this section illustrates the structure and assumptions necessary for investors to acquire reasonable knowledge of exposures to the systematic factors, $[\mathbf{B}_g \ \mathbf{D}_g]$. A common approach involves using time-series regressions of asset returns on factor realizations. However, the literature has long recognized that beta estimates derived from short time horizons are plagued by substantial errors (see Fama and MacBeth (1973) and Shanken (1992)). Moreover, increasing the cross-sectional sample size does little to mitigate these errors.

To address these limitations, we extend the methodologies of Fan et al. (2016) and Kim et al. (2021) to a two country setup. This approach leverages the large cross section of asset characteristics to consistently estimate mispricing and factor loadings, even when time-series data is limited. Unlike Kim et al. (2021), whose focused on analyzing a single-country market, our objective is broader. We seek to provide a comprehensive framework that not only identifies and exploits mispricing but also capitalizes on disparities in common factor premia and country-specific factors. This strategy unlocks a richer dimension of international investment, empowering investors to systematically harness the opportunities that arise in foreign domains.

The key insight lies in recognizing that asset characteristics are determinants of risks. In particular, we allow the systematic risk $[\mathbf{B}_g \ \mathbf{D}_g]$ to be functions of asset-specific characteristics in each country. Let $\mathbf{x}_{gi} = \begin{bmatrix} x_{gi1} & \cdots & x_{giL_g} \end{bmatrix}'$ be the $(L_g \times 1)$ vector of the characteristics associated with stock i in country g.[3] Then, define the $(N_g \times L_g)$ matrix of $\mathbf{X}_g$, the i-th row of which is $\mathbf{x}'_{gi}$. We assume the following structure for $\mathbf{B}_g$ and $\mathbf{D}_g$:

$$\mathbf{B}_g = \mathbf{X}_g \Theta_g^c + \Gamma_g^c \tag{2.4}$$

and

$$\mathbf{D}_g = \mathbf{X}_g \Theta_g^s + \Gamma_g^s, \tag{2.5}$$

where Θ_g^c is the $(L_g \times K^c)$ matrix, Θ_g^s is the $\left(L_g \times K_g^s\right)$ matrix, and the $(N_g \times K^c)$ matrix, Γ_g^c and the $\left(N_g \times K_g^s\right)$ matrix, Γ_g^s are cross-sectionally orthogonal to the characteristic space of $\mathbf{X}_g$.

[3] Note that the number of characteristics can differ across countries.

We call the two matrices of Θ_g^c and Θ_g^s factor loading matrices because they relate characteristics to factor loadings to the common and country-specific systematic factors, respectively. The terms of Γ_g^c and Γ_g^s represent the sources of risk that are not attributable to the characteristics, including omitted characteristics. As will be shown later, while the factor loading matrices, Θ_g^c and Θ_g^s can be consistently estimated in the large N/small T setting used here, consistent estimates of Γ_g^c and Γ_g^s cannot be estimated consistently in a small T setting. Therefore, our procedure does not attempt to exploit the gammas, just their orthogonality to the characteristics in each country. Furthermore, although we restrict the relation between factor loadings and characteristics to be linear, there are various approaches to incorporate non-linearity. For example, we would have chosen $\mathbf{X}_g$ to be a large set of characteristics, possibly containing suitable polynomials of some underlying characteristics, $\mathbf{X}_g^*$, as in Kim et al. (2021).

Incorporating (2.4) and (2.5) into (2.3), we have that:

$$\mathbf{R}_g = \left(\mathbf{X}_g \Theta_g^c + \Gamma_g^c\right)\left(\boldsymbol{\lambda}_g^c \mathbf{1}_T' + \mathbf{F}'\right) + \left(\mathbf{X}_g \Theta_g^s + \Gamma_g^s\right)\left(\boldsymbol{\lambda}_g^s \mathbf{1}_T' + \mathbf{G}_g'\right) + \mathbf{E}_g. \tag{2.6}$$

Using the above specification, we provide a powerful framework to extract a wide cross-section of betas, surpassing the limitations of conventional time-series regression estimates. Firstly, by instrumenting characteristics of $\mathbf{X}_g$, we can extract the information on the betas (via $\mathbf{X}_g\Theta_g^c$ and $\mathbf{X}_g\Theta_g^s$) even when data are relatively infrequently observed (such as monthly) over a horizon less than a decade. This is a strong advantage over other factor loadings extraction methods requiring long time series or high frequency observations. Second, because we can set T to be small, the process in (2.6) can be treated as a local approximation of a conditional model over a long horizon. In particular, our rolling estimation of (2.6) enables us to study the *temporal* relation of characteristics to risk. Many empirical approaches (e.g. Kelly et al. (2019), Ferson and Harvey (1999), Ghysels (1998)) construct conditional models by allowing the characteristics to change period-by-period but holding the cross-sectional relation between characteristics and mispricing or risk *constant*. By contrast, our method explicitly captures the dynamics of Θ_g^c and Θ_g^s. Lastly, we do not need to necessarily have all important characteristics for mispricings or risks in (2.6). Because any information in the missing characteristics is captured by Γ_g^c and Γ_g^s, our model already incorporates the possibility of misspecifying the set of characteristics. Hence, if some important characteristics are missing, we may lose some precision but will not generate spurious results.

Before formalizing the two-country economy, let us first consider how our framework provides a novel perspective on evaluating market integration between countries. According to the Arbitrage Pricing Theory (APT; Ross (1976)), in a world free of frictions and arbitrage opportunities, compensation for common risks should be identical across all assets in both economies, implying $\boldsymbol{\lambda}_1^c = \boldsymbol{\lambda}_2^c$. However, our framework allows the possibility that $\boldsymbol{\lambda}_1^c$ may differ from $\boldsymbol{\lambda}_2^c$. Investors can construct strategies that exploit these differences, and the profitability of such strategies becomes a tangible measure of market integration. High profitability points to significant segmentation between the markets, while negligible profits suggest a greater degree of integration

We formalize the two-country economy. First, we assume the standard regularity conditions on the characteristics and residual returns.

Assumption 1. *In each country $g = 1, 2$, as $N_g \to \infty$, it holds that*
(i) $\frac{\mathbf{R}_g'\mathbf{R}_g}{N_g} \xrightarrow{p} \mathbf{V}_{R_g}$ *and* $\frac{\mathbf{X}_g'\mathbf{X}_g}{N_g} \to \mathbf{V}_{X_g}$, *where* $\mathbf{V}_{R_g}$, $\mathbf{V}_{X_g}$ *are positive definite matrices,*
(ii) $\frac{\mathbf{X}_g'\Gamma_g^c}{N_g} \xrightarrow{p} \mathbf{0}_{L_g \times K^c}$, $\frac{\mathbf{X}_g'\Gamma_g^s}{N_g} \xrightarrow{p} \mathbf{0}_{L_g \times K_g^s}$, *and* $\frac{\mathbf{X}_g'\mathbf{E}_g}{N_g} \xrightarrow{p} \mathbf{0}_{L_g \times T}$.

Condition (i) simply states that the cross section of returns and characteristics are not redundant but well-spread over individual stocks in each country. Condition (ii) imposes the cross-sectional orthogonality conditions between the characteristics, $\mathbf{X}_g$, and factor loading residuals, Γ_g^c and Γ_g^s, and residual returns, $\mathbf{E}_g$, in each country.

Next, we assume mild restrictions to separately identify factor loading matrices Θ_g^c and Θ_g^s. We introduce the symmetric $(T \times T)$ matrix $\mathbf{J}_T = \mathbf{I}_T - \frac{1}{T}\mathbf{1}_T\mathbf{1}_T'$, which corresponds to time-series demeaning. We define Θ_g and $\mathbf{H}_g$ as $\left[\Theta_g^c \; \Theta_g^s\right]$ and $\left[\mathbf{F} \; \mathbf{G}_g^s\right]$, respectively.

Assumption 2. *In each country $g = 1, 2$, as $N_g \to \infty$, it holds that*
(i) $\frac{\Theta_g'\mathbf{X}_g'\mathbf{X}_g\Theta_g}{N_g} = \begin{bmatrix} \frac{\Theta_g^{c\prime}\mathbf{X}_g'\mathbf{X}_g\Theta_g^c}{N_g} & \frac{\Theta_g^{c\prime}\mathbf{X}_g'\mathbf{X}_g\Theta_g^s}{N_g} \\ \frac{\Theta_g^{s\prime}\mathbf{X}_g'\mathbf{X}_g\Theta_g^c}{N_g} & \frac{\Theta_g^{s\prime}\mathbf{X}_g'\mathbf{X}_g\Theta_g^s}{N_g} \end{bmatrix} \to \begin{bmatrix} \Theta_g^{c\prime}\mathbf{V}_{X_g}\Theta_g^c & \Theta_g^{c\prime}\mathbf{V}_{X_g}\Theta_g^s \\ \Theta_g^{s\prime}\mathbf{V}_{X_g}\Theta_g^c & \Theta_g^{s\prime}\mathbf{V}_{X_g}\Theta_g^s \end{bmatrix} = \mathbf{V}_{\Theta_g}$, *where* $\mathbf{V}_{\Theta_g}$ *is a $(K_g \times K_g)$ positive definite matrix,*
(ii) $\frac{\mathbf{H}_g'\mathbf{J}_T\mathbf{H}_g}{T} = \begin{bmatrix} \frac{\mathbf{F}'\mathbf{J}_T\mathbf{F}}{T} & \frac{\mathbf{F}'\mathbf{J}_T\mathbf{G}_g}{T} \\ \frac{\mathbf{G}_g'\mathbf{J}_T\mathbf{F}}{T} & \frac{\mathbf{G}_g'\mathbf{J}_T\mathbf{G}_g}{T} \end{bmatrix} = \begin{bmatrix} \Sigma_F & \mathbf{0}_{K^c \times K_g^s} \\ \mathbf{0}_{K_g^s \times K^c} & \Sigma_{G_g} \end{bmatrix} = \Sigma_{H_g}$, *where* Σ_{H_g} *is a $(K_g \times K_g)$ positive definite matrix,*
(iii) $\begin{bmatrix} \frac{\mathbf{G}_1'\mathbf{J}_T\mathbf{G}_1}{T} & \frac{\mathbf{G}_1'\mathbf{J}_T\mathbf{G}_2}{T} \\ \frac{\mathbf{G}_2'\mathbf{J}_T\mathbf{G}_1}{T} & \frac{\mathbf{G}_2'\mathbf{J}_T\mathbf{G}_2}{T} \end{bmatrix}$ *is a $(K_1^s + K_2^s) \times (K_1^s + K_2^s)$ positive definite matrix.*

Condition (i) implies that each column of $\mathbf{X}_g \Theta_g$ provides non-redundant information.[4] Condition (ii) ensures that factor realizations are not redundant in each country. Furthermore, the common factors and country-specific factors are orthogonal over the sample period. Note that these restrictions are without loss of generality because we can regress any correlated country-specific factors on the common factors and reconstruct the orthogonal country-specific factors from the residuals. Condition (iii) imposes minimal restrictions on the behavior of country-specific factors, requiring that they cannot be replicated across countries. Otherwise, the country-specific factors would lose their country-specific features and would instead be classified as common factors.

2.3 Methodology

We present our systematic approach that enables investors to uncover and capitalize on opportunities in foreign markets, structured into three steps. In the first step, we identify common and country-specific factors. Extracting systematic factors from a large cross-section of returns is relatively straightforward when factor pricing holds, as shown by Connor and Korajczyk (1986). Our method extends this approach by projecting returns onto the characteristic space, which improves the efficiency of the estimator. Once systematic factors are identified for each country, canonical correlation analysis is applied to separate common factors from country-specific ones.

In the second step, we assess how firm characteristics inform mispricing and factor loadings for the two types of factors identified in the first step. By examining how the interaction between asset characteristics and factors explains stock returns, we learn about the information in firm characteristics on factor loadings. Any residual explanatory power of firm characteristics on returns is then attributed to mispricing.

Finally, in the third step, we propose practical portfolio strategies to systematically harness foreign investment opportunities. The information on the common/country-specific factor loadings from the second step can be interpreted as portfolio weights, recovering common/country-specific factors accordingly. Then, we can aggregate portfolios generating country-specific factors into a single portfolio based on a valid criterion, such as mean-variance maximization. Furthermore, we can assess the premium difference on common factors by regressing the common factors of country 2 on a constant

[4]This assumption does not require that all characteristics contribute to the factor loadings. Any redundant characteristics will be accounted for by zeros in the corresponding entries of Θ_g.

and the common factors of country 1, and then construct a hedging portfolio accordingly.

Step 1: Estimation of common factors and country-specific factors The first step of our procedure involves estimating the common factor innovation, $\mathbf{F}'\mathbf{J}_T$, and the country-specific factor innovation, $\mathbf{G}_g'\mathbf{J}_T$, from returns in the two countries. Recall that the observed returns in (2.6) are driven by the risk premiums $\boldsymbol{\lambda}_g^c$ and $\boldsymbol{\lambda}_g^s$, as well as the realizations of $\mathbf{F}$ and $\mathbf{G}_g$. Because our objective in this step is to estimate the demeaned factor realization rather than the risk premiums, we eliminate the effects of the risk premium by demeaning the observed returns by post-multiplying $\mathbf{J}_T \left(= \mathbf{I}_T - \frac{1}{T}\ \mathbf{1}_T\mathbf{1}_T\right)$:

$$\begin{aligned}\mathbf{R}_g\mathbf{J}_T &= \left(\mathbf{X}_g\Theta_g^c + \Gamma_g^c\right)\left(\boldsymbol{\lambda}_g^c\mathbf{1}_T' + \mathbf{F}'\right)\mathbf{J}_T + \left(\mathbf{X}_g\Theta_g^s + \Gamma_g^s\right)\left(\boldsymbol{\lambda}_g^s\mathbf{1}_T' + \mathbf{G}_g'\right)\mathbf{J}_T + \mathbf{E}_g\mathbf{J}_T \\ &= \left(\mathbf{X}_g\Theta_g^c + \Gamma_g^c\right)\mathbf{F}'\mathbf{J}_T + \left(\mathbf{X}_g\Theta_g^s + \Gamma_g^s\right)\mathbf{G}_g'\mathbf{J}_T + \mathbf{E}_g\mathbf{J}_T, \qquad (2.7)\end{aligned}$$

where the last equality is from the property of $\mathbf{1}_T'\mathbf{J}_T = \mathbf{1}_T'\left(\mathbf{I}_T - \frac{1}{T}\ \mathbf{1}_T\mathbf{1}_T'\right) = \mathbf{1}_T' - \frac{T}{T}\mathbf{1}_T' = \mathbf{0}_T'$. To further isolate $\mathbf{X}_g\Theta_g^c$ and $\mathbf{X}_g\Theta_g^s$, we project the demeaned returns of (2.7) on the (linear) span of $\mathbf{X}_g$ by premultiplying the projection matrix $\mathbf{P}_g = \mathbf{X}_g\left(\mathbf{X}_g'\mathbf{X}_g\right)^{-1}\mathbf{X}_g'$. This gives:

$$\begin{aligned}\widehat{\mathbf{R}}_g \equiv & \mathbf{P}_g\mathbf{R}_g\mathbf{J}_T \qquad (2.8) \\ &= \left(\mathbf{P}_g\mathbf{X}_g\Theta_g^c + \mathbf{P}_g\Gamma_g^c\right)\mathbf{F}'\mathbf{J}_T + \left(\mathbf{P}_g\mathbf{X}_g\Theta_g^s + \mathbf{P}_g\Gamma_g^s\right)\mathbf{G}_g'\mathbf{J}_T + \mathbf{P}_g\mathbf{E}_g\mathbf{J}_T \\ &= \left(\mathbf{X}_g\Theta_g^c + \mathbf{P}_g\Gamma_g^c\right)\mathbf{F}'\mathbf{J}_T + \left(\mathbf{X}_g\Theta_g^s + \mathbf{P}_g\Gamma_g^s\right)\mathbf{G}_g'\mathbf{J}_T + \mathbf{P}_g\mathbf{E}_g\mathbf{J}_T,\end{aligned}$$

where the last equality is from $\mathbf{P}_g\mathbf{X}_g = \mathbf{X}_g$. Furthermore, exploiting the orthogonality of Γ_g^c, Γ_g^s and $\mathbf{E}_g$ with respect to $\mathbf{X}_g$ from Assumption 1(ii), the terms of $\mathbf{P}_g\Gamma_g^c$, $\mathbf{P}_g\Gamma_g^s$, and $\mathbf{P}_g\mathbf{E}_g$ in the above equation will become negligible when N_g is large. Hence, with a large N_g, it follows that $\widehat{\mathbf{R}}_g = \mathbf{P}_g\mathbf{R}_g\mathbf{J}_T \approx \left[\mathbf{X}_g\Theta_g^c\ \mathbf{X}_g\Theta_g^s\right]\left[\mathbf{F}'\mathbf{J}_T\ \mathbf{G}_g'\mathbf{J}_T\right] = \mathbf{X}_g\Theta_g\mathbf{H}_g'\mathbf{J}_T$, where $\Theta_g = \left[\Theta_g^c\ \Theta_g^s\right]$ and $\mathbf{H}_g = \left[\mathbf{F}\ \mathbf{G}_g\right]$. Finally, we estimate $\mathbf{J}_T\mathbf{H}_g$ by applying standard principal component analysis to $\widehat{\mathbf{R}}_g$.

Theorem 2.1. *Let $\widehat{\mathbf{H}}_g$ denote the $(T \times K_g)$ matrix, the k-th column of which is the eigenvector of $\frac{\widehat{\mathbf{R}}_g'\widehat{\mathbf{R}}_g}{N_g}$ corresponding to the k-th largest eigenvalue of $\frac{\widehat{\mathbf{R}}_g'\widehat{\mathbf{R}}_g}{N_g}$, where $\widehat{\mathbf{R}}_g$ is given by (2.8). Under Assumptions 1 and 2, as N_g increases, $\widehat{\mathbf{H}}_g \xrightarrow{p} \mathbf{J}_T\mathbf{H}_g\mathcal{O}_g$, where the $(K_g \times K_g)$ matrix $\mathcal{O}_g$ is given in Lemma C.1.*

Note that we identify the (demeaned) factors only up to rotation, as is common in the latent factor literature.[5] To provide some intuition of the above theorem, recall that $\widehat{\mathbf{R}}_g$ converges (as $N_g \to \infty$) to $\mathbf{X}_g \Theta_g \mathbf{H}_g' \mathbf{J}_T = \mathbf{X}_g \Theta_g \mathcal{O}_g^{-1\prime} \mathcal{O}_g' \mathbf{H}_g' \mathbf{J}_T$. Therefore, $\frac{\widehat{\mathbf{R}}_g' \widehat{\mathbf{R}}_g}{N_g}$ converges to $\mathbf{J}_T \mathbf{H}_g \mathcal{O}_g \left(\mathcal{O}_g^{-1} \frac{\Theta_g' \mathbf{X}_g \mathbf{X}_g \Theta_g}{N_g} \mathcal{O}_g^{-1\prime} \right) \mathcal{O}_g' \mathbf{H}_g' \mathbf{J}_T$. From Lemma C.1, $\mathcal{O}_g$ features that $\mathcal{O}_g' \mathbf{H}_g' \mathbf{J}_T \mathbf{H}_g \mathcal{O}_g = \mathbf{I}_{K_g}$ and, with Assumption 2(i), $\mathcal{O}_g^{-1} \frac{\Theta_g' \mathbf{X}_g \mathbf{X}_g \Theta_g}{N_g} \mathcal{O}_g^{-1\prime}$ converges to a diagonal matrix. Hence, with a large N_g, each column of $\mathbf{J}_T \mathbf{H}_g \mathcal{O}_g$ and each diagonal element of $\mathcal{O}_g^{-1} \frac{\Theta_g' \mathbf{X}_g \mathbf{X}_g \Theta_g}{N_g} \mathcal{O}_g^{-1\prime}$ can be interpreted as an eigenvector and an eigenvalue of $\frac{\widehat{\mathbf{R}}_g' \widehat{\mathbf{R}}_g}{N_g}$, respectively. Resorting to these observations, we attempt to recover $\mathbf{J}_T \mathbf{H}_g \mathcal{O}_g$ as stated in Theorem 2.1.

Then, we proceed to identify common factors and and country specific factors. The main analytical tool for this purpose is Canonical Correlation Analysis (CCA). Define the $(K_1 \times K_1)$ matrix $\widehat{\Sigma}_{H_1}$, the $(K_2 \times K_2)$ matrix $\widehat{\Sigma}_{H_2}$, and the $(K_1 \times K_2)$ matrix $\widehat{\Sigma}_{H_{12}}$ as $\widehat{\Sigma}_{H_1} = \frac{\widehat{\mathbf{H}}_1' \widehat{\mathbf{H}}_1}{T}$, $\widehat{\Sigma}_{H_2} = \frac{\widehat{\mathbf{H}}_2' \widehat{\mathbf{H}}_2}{T}$, and $\widehat{\Sigma}_{H_{12}} = \frac{\widehat{\mathbf{H}}_1' \widehat{\mathbf{H}}_2}{T}$, respectively.

Because the setup is symmetric, we demonstrate the identification of common factors and country-specific factors using the estimated systematic factors from country 1. We propose a common factor estimator, $\widehat{\mathbf{F}}_1$, and a country-specific factor estimator, $\widehat{\mathbf{G}}_1$, as follows. Let $\widehat{\mathbf{W}}_1$ be the $(K_1 \times K_1)$ matrix, the k-th column of which is the k-th canonical direction (or, the eigenvector of $\widehat{\Sigma}_{H_1}^{-1} \widehat{\Sigma}_{H_{12}} \widehat{\Sigma}_{H_2}^{-1} \widehat{\Sigma}_{H_{12}}'$, corresponding to the k-th largest eigenvalue), such that $\widehat{\mathbf{W}}_1' \widehat{\Sigma}_{H_1} \widehat{\mathbf{W}}_1 = \mathbf{I}_{K_1}$. We then partition $\widehat{\mathbf{W}}_1$ as $\begin{bmatrix} \widehat{\mathbf{W}}_1^c & \widehat{\mathbf{W}}_1^s \end{bmatrix}$, where $\widehat{\mathbf{W}}_1^c$ be the $(K_1 \times K^c)$ matrix and $\widehat{\mathbf{W}}_1^s$ be the $(K_1 \times K_1^s)$ matrix. Finally, we propose the estimators:

$$\widehat{\mathbf{F}}_1 = \widehat{\mathbf{H}}_1 \widehat{\mathbf{W}}_1^c, \quad \widehat{\mathbf{G}}_1 = \widehat{\mathbf{H}}_1 \widehat{\mathbf{W}}_1^s.$$

Canonical Correlation Analysis (CCA) is designed to find linear combinations of two sets of variables that maximize their correlation. In our two-country setting, because both $\widehat{\mathbf{H}}_1$ and $\widehat{\mathbf{H}}_2$ contain the common factors, CCA will extract the common factors, which generate correlations of one, and then proceed to the country-specific factors, which generate correlations strictly less than one. Accordingly, we recover the common factors from $\widehat{\mathbf{F}}_1$ using the first K^c cannonical directions and the country-specific factors from $\widehat{\mathbf{G}}_1$ using the next K_1^s canonical directions. The following theorem formalizes this result.

[5] We choose the matrix $\mathcal{O}_g$ so that $\mathcal{O}_g' \Sigma_{H_g} \mathcal{O}_g$ is an identity matrix and $\mathcal{O}_g^{-1} \mathbf{V}_{\Theta_g} \mathcal{O}_g^{-1\prime}$ is a diagonal matrix.

Theorem 2.2. *Under Assumptions 1 and 2, it holds that* $\widehat{\mathbf{F}}_g \xrightarrow{p} \mathbf{J}_T\mathbf{F}\mathcal{S}^c$ *and* $\widehat{\mathbf{G}}_g \xrightarrow{p} \mathbf{J}_T\mathbf{G}_g\mathcal{S}_g^s$ *for some* $\mathcal{S}^c$ *such that* $\mathcal{S}^c\mathcal{S}^{c\prime} = \Sigma_F^{-1}$ *and some* $\mathcal{S}_g^s$ *such that* $\mathcal{S}_g^s\mathcal{S}_g^{s\prime} = \Sigma_{G_g}^{-1}$ *for* $g = 1, 2$.

Step 2: Estimation of factor loading matrices on common and country-specific factors Next, we show how to estimate Recall that in Step 1, we introduce $\widehat{\mathbf{R}}_g = \mathbf{P}_g\mathbf{R}_g\mathbf{J}_T$ and show the following property of $\widehat{\mathbf{R}}_g$:

$$\widehat{\mathbf{R}}_g \approx \mathbf{X}_g\Theta_g^c\mathbf{F}'\mathbf{J}_T + \mathbf{X}_g\Theta_g^s\mathbf{G}_g'\mathbf{J}_T.$$

Because we already obtain consistent estimators $\widehat{\mathbf{F}}_g' \xrightarrow{p} \mathcal{S}^{c\prime}\mathbf{F}'\mathbf{J}_T$ and $\widehat{\mathbf{G}}_g' \xrightarrow{p} \mathcal{S}_g^{s\prime}\mathbf{G}'\mathbf{J}_T$ from Step 1, Θ_g^c and Θ_g^s can be recovered by regressing $\widehat{\mathbf{R}}_g$ on the interaction between $\mathbf{X}_g$ and $\widehat{\mathbf{F}}'$. The following theorem formalizes this procedure.

Theorem 2.3. *Let* $\widehat{\Theta}_g^c$ *(a* $(L_g \times K^c)$ *matrix) and* $\widehat{\Theta}_g^s$ *(a* $\left(L_g \times K_g^s\right)$ *matrix) solve the following optimization problem:*

$$\left(\widehat{\Theta}_g^c, \widehat{\Theta}_g^s\right) = \arg\min_{\left(\Theta_g^c, \Theta_g^s\right)} \|\widehat{\mathbf{R}}_g - \left(\mathbf{X}_g\Theta_g^c\right)\widehat{\mathbf{F}}_g' - \left(\mathbf{X}_g\Theta_g^s\right)\widehat{\mathbf{G}}_g'\|. \tag{2.9}$$

Under Assumptions 1 and 2, it holds that $\widehat{\Theta}_g^c \xrightarrow{p} \Theta_g^c\mathcal{S}^{c\prime-1}$ *and* $\widehat{\Theta}_g^s \xrightarrow{p} \Theta_g^s\mathcal{S}_g^{s\prime-1}$ *for some* $\mathcal{S}^c$ *such that* $\mathcal{S}^c\mathcal{S}^{c\prime} = \Sigma_F^{-1}$ *some* $\mathcal{S}_g^s$ *such that* $\mathcal{S}_g^s\mathcal{S}_g^{s\prime} = \Sigma_{G_g}^{-1}$ *for* $g = 1, 2$.

The optimization problem in the above theorem can be transformed into a conventional ordinary least square problem.

Step 3: Two pathways of portfolio formation We proceed to construct portfolios which would systematically exploit investment opportunities in foreign markets. It is convenient to rewrite (2.6) as follows:

$$\mathbf{R}_g = \left(\mathbf{X}_g\Theta_g^c\right)\left(\boldsymbol{\lambda}_g^c\mathbf{1}_T' + \mathbf{F}'\right) + \left(\mathbf{X}_g\Theta_g^s\right)\left(\boldsymbol{\lambda}_g^s\mathbf{1}_T' + \mathbf{G}_g'\right) + \mathbf{U}_g, \tag{2.10}$$

where $\mathbf{U}_g = \Gamma_g^c\left(\boldsymbol{\lambda}_g^c\mathbf{1}_T' + \mathbf{F}'\right) + \Gamma_g^s\left(\boldsymbol{\lambda}_g^s\mathbf{1}_T' + \mathbf{G}_g'\right) + \mathbf{E}_g$.

Consider $(N_g \times K_g)$ infeasible portfolios $\mathbf{w}_g^b = \frac{\mathbf{X}_g}{N}\left[\Theta_g^c\ \Theta_g^s\right]\left(\left[\Theta_g^c\ \Theta_g^s\right]'\mathbf{V}_{X_g}\left[\Theta_g^c\ \Theta_g^s\right]\right)^{-1}$. These portfolios deliver common factors and country-specific factors along with their associated premium as follows:

$$\mathbf{w}_g^{b\prime}\mathbf{R}_g = \left(\left[\Theta_g^c\ \Theta_g^s\right]'\mathbf{V}_{X_g}\left[\Theta_g^c\ \Theta_g^s\right]\right)^{-1}\left[\Theta_g^c\ \Theta_g^s\right]'\frac{\mathbf{X}_g'\mathbf{X}_g}{N}\left[\Theta_g^c\ \Theta_g^s\right]\begin{bmatrix}\boldsymbol{\lambda}_g^c\mathbf{1}_T'+\mathbf{F}'\\ \boldsymbol{\lambda}_g^s\mathbf{1}_T'+\mathbf{G}_g'\end{bmatrix} \tag{2.11}$$

$$+\left(\left[\Theta_g^c\ \Theta_g^s\right]'\mathbf{V}_{X_g}\left[\Theta_g^c\ \Theta_g^s\right]\right)^{-1}\left[\Theta_g^c\ \Theta_g^s\right]'\frac{\mathbf{X}_g'\mathbf{U}_g}{N} \tag{2.12}$$

$$\xrightarrow{p}\begin{bmatrix}\boldsymbol{\lambda}_g^c\mathbf{1}_T'+\mathbf{F}'\\ \boldsymbol{\lambda}_g^s\mathbf{1}_T'+\mathbf{G}_g'\end{bmatrix}, \tag{2.13}$$

where the last limit follows from Assumptions 1(ii).

The next theorem shows that the feasible analogue of

$$\widehat{\mathbf{w}}_g^b = \frac{\mathbf{X}_g}{N_g}\left[\widehat{\Theta}_g^c\ \widehat{\Theta}_g^s\right]\left(\left[\widehat{\Theta}_g^c\ \widehat{\Theta}_g^s\right]'\frac{\mathbf{X}_g'\mathbf{X}_g}{N_g}\left[\widehat{\Theta}_g^c\ \widehat{\Theta}_g^s\right]\right)^{-1}$$

constructed using the Step 2 results also converges to the same limits, up to a rotation.

Theorem 2.4. *Under Assumptions 1 and 2, it holds that* $\widehat{\mathbf{w}}_g^{b\prime}\mathbf{R}_g \xrightarrow{p} \begin{bmatrix}\mathcal{S}^{c\prime}\left(\boldsymbol{\lambda}_g^c\mathbf{1}_T'+\mathbf{F}'\right)\\ \mathcal{S}_g^{s\prime}\left(\boldsymbol{\lambda}_g^s\mathbf{1}_T'+\mathbf{G}_g'\right)\end{bmatrix}$, *for some* $\mathcal{S}^c$ *such that* $\mathcal{S}^c\mathcal{S}^{c\prime}=\Sigma_F^{-1}$ *and some* $\mathcal{S}_g^s$ *such that* $\mathcal{S}_g^s\mathcal{S}_g^{s\prime}=\Sigma_{G_g}^{-1}$ *for* $g=1,2$.

The implication of the above theorem follows. If we partition $\widehat{\mathbf{w}}_g^b$ into $\left[\widehat{\mathbf{w}}_g^c\ \widehat{\mathbf{w}}_g^s\right]$ where $\widehat{\mathbf{w}}_g^c$ is an $(N_g\times K^c)$ matrix and $\widehat{\mathbf{w}}_g^s$ is an $\left(N_g\times K_g^s\right)$ matrix, investors can use the portfolios $\widehat{\mathbf{w}}_g^s$ to gain exposure to the country specific factors in country g.

Furthermore, it holds that $\widehat{\mathbf{w}}_1^{c\prime}\mathbf{R}_1 \xrightarrow{p} \mathcal{S}^{c\prime}\left(\boldsymbol{\lambda}_1^c\mathbf{1}_T'+\mathbf{F}'\right)$ and $\widehat{\mathbf{w}}_2^{c\prime}\mathbf{R}_2 \xrightarrow{p} \mathcal{S}^{c\prime}\left(\boldsymbol{\lambda}_2^c\mathbf{1}_T'+\mathbf{F}'\right)$, implying that investors can evaluate the difference in risk premia on the common factors across two countries. In particular, we can measure the difference through time-series regression as shown in the following theorem.

Theorem 2.5. *Regress* $\widehat{\mathbf{w}}_2^{c\prime}\mathbf{R}_2$ *on a constant and* $\widehat{\mathbf{w}}_1^{c\prime}\mathbf{R}_1$*. Under Assumptions 1 and 2, the* $(K^c\times 1)$ *vector intercept converges to* $\mathcal{S}^{c\prime}(\boldsymbol{\lambda}_2^c-\boldsymbol{\lambda}_1^c)$ *for some* $\mathcal{S}^c$ *such that* $\mathcal{S}^c\mathcal{S}^{c\prime}=\Sigma_F^{-1}$.

The above theorem can be translated into practical common-factor hedging strategies. For example, if an investor finds that the element in the intercept vector is substantially positive, she can construct a hedging portfolio of buying the portfolio of $\widehat{\mathbf{w}}_2^{c\prime}\mathbf{R}_2$ while taking an appropriate short position on the portfolio of $\widehat{\mathbf{w}}_1^{c\prime}\mathbf{R}_1$.

3 Simulation

This section presents simulation evidence to validate our method, which enables investors in Country 1 to diversify into Country 2-specific factors and capitalize on differences in risk premia between the two countries.

As a representative country pair, we consider the USA for Country 1 and Canada for Country 2 to calibrate the return-generating process of individual asset returns. For the factor structure, we consider six systematic factors per country, comprising three common factors shared across both countries and three country-specific factors. Applying this specification to the most recent 60-month window from 2018 to 2022, we estimate betas, residual variance, common factors, and country-specific factors, which are treated as the true parameters in the simulation.

One of the key assumptions of our method is that we know the true number of systematic factors for each country and the number of common factors across the two countries before applying our portfolio construction method. The main purpose of the simulation exercise is to investigate whether our method works under the correct specification and to further assess how our results respond when this assumption deviates.

3.1 Effect of the Estimated Total Number of Systematic Factors

We explore the impact of over- or underestimating the total number of systematic factors on the correct classification of common and country-specific factors while the true number of factors is six in the true return generating process for both countries. In Table 1, we report the mean of estimated canonical correlations over 10,000 repetitions from K estimated systematic factors from the simulated panel returns of both countries.

Table 1: Canonical Correlations for Different Numbers of Estimated Systematic Factors
This table reports the average canonical correlations across 10,000 repetitions for varying numbers of estimated systematic factors (K), comparing cases of underestimation ($K = 5$), correct estimation ($K = 6$), and overestimation ($K = 7$).

Canonical Correlations	1	2	3	4	5	6	7
$K = 5$, Underestimation	0.9958	0.9368	0.7722	0.2330	0.0698	-	-
$K = 6$, Correct	0.9963	0.9641	0.9389	0.3342	0.1808	0.0566	-
$K = 7$, Overestimation	0.9964	0.9659	0.9418	0.4118	0.2720	0.1540	0.0481

For the correct case ($K = 6$, the middle row), we find that the first three canonical correlations are quite high, 0.9963, 0.9641, and 0.9389, and suddenly drop to 0.3342 for the fourth canonical correlation, consistent with the theoretical prediction, given that the true number of common factors is three. Such a sudden drop from the third to the fourth canonical correlation is also evident in both the underestimation case ($K = 5$, the first row), where the correlations drop from 0.7722 to 0.2330, and the overestimation case ($K = 7$, the last row), where they drop from 0.9418 to 0.4118. This consistent drop from the third to the fourth canonical correlation highlights the presence of the three common factors in the economy. However, the underestimation case ($K = 5$) appears to impair the accurate recovery of these common factors, as reflected in the lower value of the third canonical correlation (0.7722), compared to 0.9389 in the correctly specified case ($K = 6$) and 0.9418 in the overestimation case ($K = 7$). Thus, overestimation appears to be relatively less harmful than underestimation, supporting our baseline choice of seven as a number of systematic factors in the empirical application.

3.2 Portfolio Performance

We evaluate whether country-specific portfolios or segmentation portfolios effectively capture their intended profits. To this end, we consider four scenarios:

(A) no segmentation in common factor premia and no risk premia for country-specific factors: neither country-specific nor segmentation portfolios should deliver positive outcomes

(B) no segmentation in common factor premia but risk premia exist for country-specific factors: country-specific portfolios should exhibit positive performance, while segmentation portfolios should not

(C) segmentation in common factor premia but no risk premia for country-specific factors: segmentation portfolios should perform well, while country-specific portfolios should show no benefits

(D) both segmentation in common factor premia and risk premia for country-specific factors: both country-specific and segmentation portfolios are expected to perform well.

Table 2 reports the performance of the two portfolios across different numbers of estimated common factors (K^c) and country-specific factors (K_s). We set $K^c = K_s = 3$ in the true return-generating process for both countries. The four panels (A, B, C, D) correspond to Scenarios A, B, C, and D, respectively.

Table 2: Sharpe Ratios of Segmentation and Country-Specific Portfolios for Varying Common and Country-Specific Factors

This table presents the annualized Sharpe Ratios (SR) for segmentation and country-specific portfolios, estimated for different combinations of common factors (K^c) and country-specific factors (K_s), with the total number of factors fixed at six. The results are based on 10,000 repetitions. Panel A corresponds to the case with neither segmentation nor country-specific risk premia. Panel B corresponds to the case with country-specific risk premia in Country 2 only. Panel C corresponds to the case with segmentation in risk premia between Country 1 and Country 2 only. Panel D allows for both country-specific premia and segmentation.

K^c (Common Factors) K_s (Country-Specific Factors)	1 5	2 4	3 3	4 2	5 1
Panel A: Neither					
Segmentation Portfolio SR	-0.0136	-0.0154	-0.0424	-0.0460	-0.0462
Country-Specific Portfolio SR	0.0125	0.0099	0.0091	0.0049	-0.0105
Panel B: Country-Specific Premia Only					
Segmentation Portfolio SR	0.0116	0.0475	0.0419	0.2911	0.4610
Country-Specific Portfolio SR	0.8652	0.8723	0.8735	0.6783	0.4383
Panel C: Segmentation in Common Factor Premia Only					
Segmentation Portfolio SR	0.5518	0.6853	0.7903	0.7822	0.7800
Country-Specific Portfolio SR	0.2333	0.1405	0.0106	0.0065	-0.0232
Panel D: Both					
Segmentation Portfolio SR	0.7050	0.8594	0.9598	1.0583	1.1278
Country-Specific Portfolio SR	0.9197	0.8967	0.8766	0.6822	0.4364

From Panel A (neither segmentation nor country-specific premia), neither the segmentation portfolio nor the country-specific portfolio exhibits positive performance, as expected, with negative or near-zero Sharpe ratios.

From Panel B (only country-specific premia), the segmentation portfolio shows minimal performance when the number of common factors is set for 1, 2, or 3, consistent with the absence of segmentation in risk premia for the common factors. However, as K^c increases over 3, the true number of common factors, the segmentation portfolio shows some profitability (e.g., 0.2911 for $K^c = 4$ and 0.4610 for $K^c = 5$), because country-specific premia are partially misclassified into market segmentation. For country-specific portfolios, performance increases as K_s increases from 1 (last column) to 3 (middle column), from 0.4383 to 0.8735. The performance of the country-specific portfolio remains robust for $K_s = 4, 5$ (e.g., 0.8723 and 0.8652).

From Panel C (only segmentation in common factor premia), we observe the flipped result of Panel B, as theory predicts. The segmentation portfolio performs strongly as the number of estimated common factors increases from 1 (first column) up to the true value of 3 (middle column) (0.5518 for $K^c = 1$ to 0.7903 for $K^c = 3$), and remains robust for $K^c = 4$ and $K^c = 5$ (0.7822 and 0.7800). Conversely, the country-specific portfolio performs poorly when K_s is set to 1, 2, or 3. However, as K_s increases to 4, 5 (first two columns), performance improves (e.g., 0.2333 for $K_s = 5$), because the segmentation in risk premia on the common factor is incorrectly attributed to the risk premia of country-specific factors.

In Panel D (covering both segmentation and country-specific premia), both portfolios exhibit robust performance. Consistent with Panels B and C, the segmentation (country-specific) portfolio's performance improves as the number of identified common (country-specific) factors rises from 1 to the true value of 3. However, Panel D shows a distinct pattern: overestimating common factors elevates the performance of segmentation portfolio (e.g., 1.0583 for $K^c = 4$ and 1.1278 for $K^c = 5$), as country-specific factors are erroneously assigned to common factors, generating profit opportunities for the segmentation portfolio while diminishing the country-specific portfolio's performance. Similarly, overestimating country-specific factors enhances the country-specific portfolio's results (e.g., 0.9197 for $K_s = 5$), as common factors are wrongly categorized as country-specific, boosting profits for the country-specific portfolio at the expense of the segmentation portfolio.

4 Empirical Analysis

4.1 Data

Our dataset builds on Jensen et al. (2023) and covers equity markets across 27 countries: Argentina (ARG), Australia (AUS), Austria (AUT), Belgium (BEL), Brazil (BRA), Canada (CAN), Chile (CHL), China (CHN), Denmark (DNK), Germany (DEU), Spain (ESP), France (FRA), United Kingdom (GBR), Hong Kong (HKG), India (IND), Italy (ITA), Japan (JPN), South Korea (KOR), Mexico (MEX), Norway (NOR), Poland (POL), Singapore (SGP), Sweden (SWE), Thailand (THA), Turkey (TUR), Taiwan (TWN), and the United States (USA).[6] Return data are from CRSP for the United States and from Compustat for all other countries All accounting data are from Compustat. For international data, all variables are measured in US dollars The sample periods vary by country, starting as early as January 1966 for the USA and extending to December 2022 for all countries, as detailed in Table 3.

To ensure data quality, we exclude the smallest 20% of firms in each country, based on market capitalization in local currency, to mitigate the impact of illiquid or micro-cap stocks.[7]

We follow methodology of Jensen et al. (2023) for constructing up to 153 firm characteristics.[8] Missing firm characteristics are imputed following the methodology of Freyberger et al. (2024), which employs a robust framework to preserve missing data points. Additionally, all firm-level characteristics are rank-normalized to lie within the $[0, 1]$ interval in each period, preserving cross-sectional information while mitigating the influence of outliers.

[6] We retain securities listed on their primary exchange only, thereby avoiding cross-listings.

[7] Our market cap filter is based on the market capitalization in local currency at the start of the estimation window, avoiding look-ahead bias. Unlike many studies, we do not winsorize returns. Our results remain robust to the conventional winsorization method proposed by Ince and Porter (2006).

[8] We thank the authors for sharing their code.

Table 3: Summary statistics. This table reports the coverage period, number of firms, market capitalization (in million USD), cross-sectional sample size within each estimation window, and the average excess returns of individual stocks for each of twenty-seven countries, listed in alphabetical order. Market capitalization columns include the 25th percentile (q_{25}), average (Avg.), and 75th percentile (q_{75}). Cross-sectional columns report the minimum (N_{min}), average ($\bar{N}$), and maximum (N_{max}) number of firms used over the sample period. Excess return columns show the average excess return ($\mu(R_M)$) and the standard deviation ($\sigma(R_M)$) of value-weighted market excess returns, expressed annually.

				Market Cap.			Cross Section			Excess Return	
Country	Beg.	End.	# Firms	q_{25}	Avg.	q_{75}	N_{min}	$\bar{N}$	N_{max}	$\mu(R_M)$	$\sigma(R_M)$
arg	Jul-2005	Dec-2022	95	21.55	907.08	479.11	46	54.88	69	0.08	0.39
aus	Feb-1986	Dec-2022	3721	8.93	650.58	125.35	117	842.87	1583	0.09	0.22
aut	Nov-1999	Dec-2022	153	44.49	1226.49	1226.17	44	61.02	79	0.08	0.23
bel	Jan-1991	Dec-2022	286	35.52	1590.81	585.27	44	103.57	133	0.07	0.19
bra	Dec-2006	Dec-2022	336	323.76	2419.44	2223.03	61	143.95	216	0.07	0.31
can	Jan-1983	Dec-2022	3191	4.02	505.00	81.45	179	751.27	1392	0.09	0.23
chl	Feb-1997	Dec-2022	229	153.78	1623.10	1414.77	52	111.80	144	0.19	0.70
chn	Jan-1995	Dec-2022	4883	335.38	1255.28	1074.47	40	1430.51	3793	0.08	0.29
deu	Jan-1987	Dec-2022	1774	6.96	1220.62	195.40	75	502.61	830	0.05	0.20
dnk	Jan-1994	Dec-2022	379	19.62	813.37	221.08	65	126.32	171	0.11	0.18
esp	May-1988	Dec-2022	417	175.73	5091.85	2978.30	36	115.99	180	0.06	0.22
fra	Jan-1987	Dec-2022	1814	19.91	2386.31	453.87	97	480.59	687	0.07	0.20
gbr	Feb-1986	Dec-2022	5548	12.77	1382.21	254.38	258	1225.09	1819	0.06	0.18
hkg	Apr-1987	Dec-2022	2872	54.34	1504.17	608.74	38	814.20	1963	0.09	0.26
ind	May-1989	Dec-2022	5096	11.94	658.76	194.24	41	1298.94	3299	0.11	0.32
ita	Jan-1987	Dec-2022	781	72.85	2296.99	1092.67	43	192.64	310	0.04	0.23
jpn	Jan-1987	Dec-2022	5648	30.43	945.28	369.38	1107	2537.70	3242	0.02	0.20
kor	Mar-1989	Dec-2022	3309	22.52	442.11	121.12	91	1012.31	2027	0.04	0.32
mex	Mar-2001	Dec-2022	195	207.32	3095.61	2506.78	56	82.48	105	0.11	0.22
nor	Jun-1993	Dec-2022	664	39.91	1100.65	563.90	54	151.54	269	0.10	0.25
pol	Mar-1996	Dec-2022	1176	11.26	349.57	126.11	51	345.52	684	0.06	0.31
sgp	Feb-1989	Dec-2022	1143	21.55	620.93	187.04	42	370.89	660	0.07	0.22
swe	Jan-1992	Dec-2022	1336	9.41	956.02	265.10	52	300.46	653	0.10	0.24
tha	Oct-1988	Dec-2022	1121	14.93	327.42	139.11	42	373.80	713	0.08	0.31
tur	Dec-1993	Dec-2022	606	30.43	834.56	436.42	50	232.79	419	0.17	0.47
twn	Jan-1989	Dec-2022	2549	35.50	461.45	248.47	46	846.08	1691	0.05	0.30
usa	Jan-1966	Dec-2022	25603	62.35	3010.13	1269.86	1579	4008.30	6448	0.06	0.16

Returns are calculated as excess returns from the perspective of an US investor, converted to US dollars (USD) to account for exchange rate effects. Table 3 provides key statistics for each market, including the number of firms, market capitalization quartiles (in USD millions), cross-sectional sample sizes (minimum, average, and maximum number of firms in each esimtation window), and excess return characteristics (mean and standard deviation). For example, the USA has the largest sample with total

25,603 firms over the sample period and an average market capitalization of \$3,010.13 million, while Argentina has the smallest with 95 firms and an average market capitalization of \$907.08 million. Value-weighted market excess returns range from 0.02 (Japan) to 0.19 (Chile) on average, with standard deviations reflecting varying levels of volatility—highest in Chile (0.70) and lowest in the USA (0.16).

4.2 Implementation

This section outlines the implementation details of our investment strategies. A first critical choice concerns the length of the estimation window used to apply the theoretical results developed in Section 2.2. We adopt a rolling estimation window of 60 months, and all portfolio returns are computed in an out-of-sample fashion, which will be elaborated later.

Another important modeling decision is the number of systematic factors. We set the total number of factors to six ($K_1 = K_2 = 7$), which provides sufficient flexibility to capture systematic sources of return.[9] Given this specification, we allocate the seven systematic factors between common and country-specific components.

Figure 4 illustrates the distribution of canonical correlations between the estimated factor spaces of Country 1 (USA) and Country 2 across different estimation windows. The top-left panel displays results when Country 2 is Canada. For instance, the distribution of the first canonical correlation exhibits a vase-shaped pattern clustered right below the perfect correlation. The red dots indicate the average correlation. The top-right panel shows results for G7 countries, the bottom-left for non-G7 countries, and the bottom-right for all 26 countries in our sample.

The canonical correlation between USA and Canada tends to be slightly higher than in other cases. Nonetheless, the overall pattern is remarkably consistent across all countries: the first two canonical correlations are high, typically exceeding 0.7, indicating substantial commonality, whereas the sixth and seventh canonical correlations are consistently low, below 0.2, highlighting the country-specific nature of the final few factors.

In selecting the number of common factors, we follow the criterion proposed by Gonçalves et al. (2025), which employs a bootstrap method to evaluate the number of

[9]The qualitative results remain similar as long as the number of systematic factors exceeds five.

common factors across two groups.[10]

We now turn to the out-of-sample implementation of the strategies implied by Theorems 2.4 and 2.5. Let t denote the last period of the estimation window. Next, we explain how to implement our strategies in Theorems 2.4 and 2.5 in a out of sample manner. Let t be the last time period of the estimation window. At time t, we have the $\left(N_g \times K_g^s\right)$ matrix of the country-specific portfolio weight $\widehat{\mathbf{w}}_{g,t}^s$ and the $(N_g \times K^s)$ matrix of the common factor portfolio weight $\widehat{\mathbf{w}}_{g,t}^c$.

Note that we take the perspective of an investor based in Country 1 (USA), seeking to exploit investment opportunities in Country 2. For the country-specific portfolios, we consolidate them into a single portfolio as follows. Let $\widehat{\mu}_t^s$ and $\widehat{\Sigma}_t^s$ be the sample mean and the sample variance of the K_2^s country-specific portfolios over the estimation window. denote the sample mean and covariance matrix of the K_2^s country-specific portfolios over the estimation window. The optimized strategy then holds:

$$r_{t+1}^{CS} = \widehat{\mu}_t^{s\prime} \left(\widehat{\Sigma}_t^s\right)^{-1} \left(\widehat{\mathbf{w}}_{2,t}^{s\prime} \mathbf{r}_{2,t+1}\right).$$

Then, we implement the strategy from Theorem 2.5 to exploit differences in common factor risk premia between Country 1 and Country 2. Over the estimation window, we estimate the following vector regression: $\widehat{\mathbf{w}}_{2,t}^{c\prime} \mathbf{r}_{2,\tau} = \Delta_\lambda + \Psi \widehat{\mathbf{w}}_{1,t}^{c\prime} \mathbf{r}_{1,\tau} + \epsilon_\tau$, where τ is a period in the estimation window, Δ_λ is a $(K^c \times 1)$ vector and Ψ is a $(K^c \times K^c)$ matrix. Using ordinary least squares, we obtain estimates of $\widehat{\Delta}_\lambda$, $\widehat{\Psi}$ and the variance of $\widehat{\Delta}_\lambda$, denoted by $\widehat{\Sigma}_{\Delta_\lambda}$. The resulting strategy holds:[11]

$$r_{t+1}^{\Delta\lambda} = \widehat{\Delta}_\lambda' \widehat{\Sigma}_{\Delta_\lambda}^{-1} \left(\widehat{\mathbf{w}}_{2,t}^{c\prime} \mathbf{r}_{2,t+1} - \widehat{\Psi} \widehat{\mathbf{w}}_{1,t}^{c\prime} \mathbf{r}_{1,t+1}\right).$$

For the types of portfolios, we normalize the in-sample standard deviation to 20% annualized to facilitate meaningful comparisons. We report the performance of these strategies in the next section.

[10] For robustness, we also consider a fixed specification with $K_1^c = K_2^c = 4$, and the qualitative interpretations remain robust.

[11] This approach is equivalent to the mean-variance optimization using the sample moments of $\widehat{\mathbf{w}}_{2,t}^{c\prime} \mathbf{r}_{2,\tau} - \widehat{\Psi} \widehat{\mathbf{w}}_{1,t}^{c\prime} \mathbf{r}_{1,\tau}$ over the estimation window.

4.3 Country Specific Portfolios

In this section, we explore the properties of country-specific portfolios. Table 4 reports the performance metrics of country-specific portfolios, starting with G7 countries (Panel A) and extending to non-G7 countries (Panel B). For G7 countries, annualized mean returns range from 13.24% (Canada) to 21.41% (United Kingdom), with 1% statistical significance for all countries. Standard deviations are approximately 20% annually, a calibrated level for in-sample fitting over the estimation window. Sharpe ratios, indicating risk-adjusted returns, peak at 1.07 for Germany and dip to 0.69 for Japan, revealing sufficient profitability for country-specific portfolios, even among G7 countries. Furthermore, correlations with the US market are low, ranging from -0.08 (France) to 0.09 (Canada), the implications of which for an investor holding the US market portfolio will be elaborated on later.

Table 4: Country-specific portfolio returns.

This table presents key financial performance metrics of country-specific portfolio returns, using G7 countries (Panel A) and non-G7 countries (Panel B) as Country 2. The metrics include annualized mean returns (%), annualized standard deviations (%), t-statistics of the average returns, Sharpe ratios, maximum drawdowns, worst and best monthly returns, and correlations with the US market (ρ_{US}). The mean returns, standard deviations and Sharpe ratios are annualized.

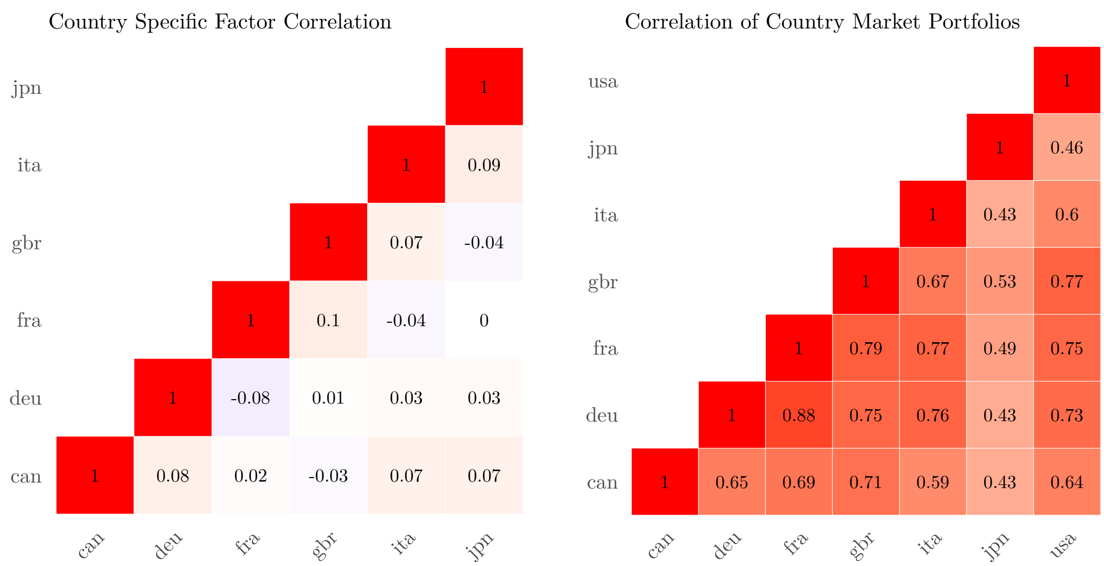
Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)	ρ_{US}
Panel A: G7 Countries as Country 2								
can	13.24	18.93	4.11	0.70	41.47	-18.94	19.11	0.09
deu	20.04	18.81	5.86	1.07	38.71	-11.93	18.04	-0.08
fra	14.33	19.00	4.02	0.75	33.04	-16.35	25.17	0.01
gbr	21.41	20.15	5.42	1.06	38.59	-19.09	23.06	0.01
ita	19.84	19.83	4.95	1.00	38.70	-23.01	18.46	-0.02
jpn	13.54	19.65	4.45	0.69	43.77	-22.01	15.63	0.07
Panel B: Non-G7 Countries as Country 2								
Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)	ρ_{US}
arg	20.14	19.07	3.54	1.06	22.96	-10.47	23.99	0.02
aus	16.66	20.96	4.57	0.79	29.42	-16.23	40.68	0.05
aut	19.54	18.93	3.81	1.03	36.14	-13.09	18.61	-0.14
bel	27.49	20.93	7.23	1.31	33.22	-14.78	24.83	-0.08
bra	10.68	21.33	1.63	0.50	32.95	-17.79	18.59	0.10
chl	19.78	21.75	4.65	0.91	50.88	-14.36	27.19	-0.03
chn	17.14	21.38	3.26	0.80	47.28	-19.60	22.59	-0.04
dnk	26.71	22.06	5.87	1.21	29.91	-18.43	21.02	-0.02
esp	18.79	20.11	5.39	0.93	34.48	-16.84	24.38	0.05
hkg	21.87	20.40	6.94	1.07	32.45	-23.01	26.09	0.10
ind	18.01	20.02	4.06	0.90	40.30	-16.09	21.62	-0.07
kor	11.39	19.68	2.62	0.58	50.19	-15.89	25.05	0.02
mex	22.92	21.03	4.19	1.09	39.66	-21.93	20.86	0.03
nor	18.14	22.34	3.88	0.81	60.58	-22.76	25.27	-0.01
pol	19.54	21.67	3.64	0.90	37.27	-22.24	23.41	-0.07
sgp	24.22	20.84	5.40	1.16	34.21	-13.83	29.82	0.09
swe	20.18	22.10	4.56	0.91	40.78	-28.07	37.55	-0.03
tha	12.47	19.23	3.62	0.65	44.87	-25.35	25.65	-0.02
tur	22.67	20.81	4.71	1.09	32.10	-18.28	29.24	-0.03
twn	12.10	19.32	3.14	0.63	49.63	-12.60	17.14	0.02

The results for non-G7 countries in Panel B convey a similar message. Annualized mean returns range from 11.39% (Korea) to 27.49% (Belgium), with Sharpe ratios peaking at 1.31 for Belgium. Correlations with the US market remain low, typically near zero or slightly negative, much like those for G7 countries.

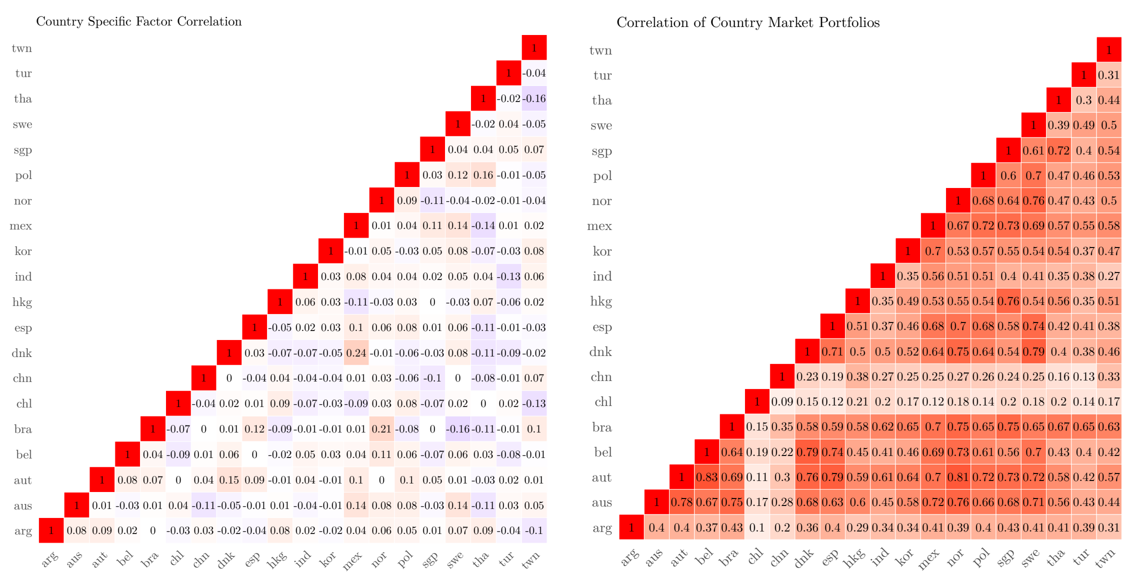
Not only the strong performance of country-specific portfolios, an investor holding the US market portfolio would find also the low correlation with the US market, as reported in the last column of Table 4, particularly attractive. Additionally, we examine correlations among country-specific portfolios. The top-left plot of Figure 1 presents a heatmap of correlations for country-specific portfolios of G7 countries, revealing very low correlations among them. In contrast, the top-right plot shows substantially higher correlations among market portfolios of G7 countries, including USA. Consistent results are found for non-G7 countries: the bottom-left plot indicates low correlations among country-specific portfolios, while the bottom-right plot shows high correlations among market portfolios of non-G7 countries. The high performance and low correlations with the US market and among country-specific portfolios create an ideal environment for diversification benefits, which we explore further.

Figure 1: Correlation Heatmaps of Country-Specific and Market Portfolios.
This figure displays heatmaps of correlations across different portfolios. The top-left panel shows correlations among country-specific portfolios of G7 countries. The top-right panel presents correlations among market portfolios of G7 countries, including USA. The bottom-left panel illustrates correlations among country-specific portfolios of non-G7 countries. The bottom-right panel depicts correlations among market portfolios of non-G7 countries.

G7 Countries

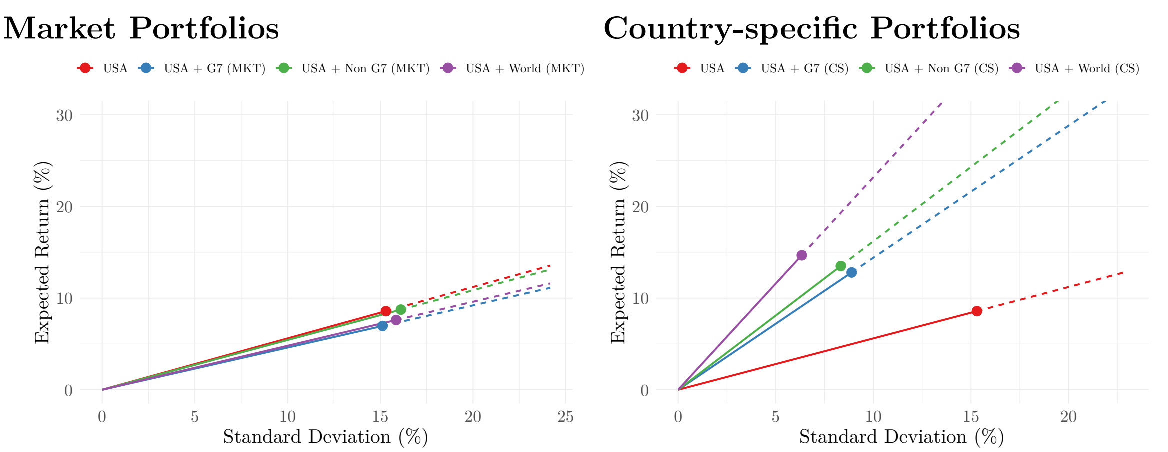


Non-G7 Countries



Firstly, we report the mean-variance improvement when the US market portfolio is combined with other country-level market portfolios in the top plot of Figure 2. The blue line, USA+G7(MKT), represents an equal-weighted investment in the US market portfolio and the average of the market portfolios of G7 countries. The green line, USA+Non-G7(MKT), represents an equal-weighted investment in the US market portfolio and the average of the market portfolios of non-G7 countries. The purple line, USA+World(MKT), represents an equal-weighted investment in three assets: the US market portfolio, the average G7 market portfolio, and the average non-G7 market portfolio. Across all three cases, combining the US market with other country-level market portfolios yields little improvement or, in some instances, deteriorates performance. These findings align with skeptical views on the benefits of international diversification (see Carrieri et al. (2007), Quinn and Voth (2008), Pukthuanthong and Roll (2009)).

Figure 2: Diversification Benefits.
This figure illustrates the mean-variance expansion of the US market portfolio through investments in international assets. The left panel shows the effect on the mean-variance efficiency from including country-level market portfolios: The blue line, USA + G7 (MKT), represents an equal-weighted investment in the US market portfolio and the average of the market portfolios of G7 countries. The green line, USA + Non-G7 (MKT), represents an equal-weighted investment in the US market portfolio and the average of the market portfolios of non-G7 countries. The purple line, USA + World (MKT), represents an equal-weighted investment in three assets: the US market portfolio, the average G7 market portfolio, and the average non-G7 market portfolio. The right panel repeats the same exercise, replacing the foreign countries' market portfolios with country-specific portfolios (CS) and reports the results. Returns are in excess of the risk-free rate and annualized.



However, when we replace country-level market portfolios with country-specific (CS) portfolios, we observe a substantial improvement of mean-variance efficiency, as shown in the bottom panel of Figure 2. Combining US market with country-specific portfolios yields higher expected returns with lower standard deviations. Table 5 reports the numerical results of Figure 2. For an investor holding half US market and half the average of country-specific portfolios of G7 countries (USA+G7(CS)), the mean excess return is 12.80% with a standard deviation of 8.89%, significantly better than the 8.58% mean excess return and 15.30% standard deviation of US market alone. The substantial reduction in standard deviation can be explained by the low correlations with US market and among country-specific portfolios. Furthermore, if an investor holds one-third US market, one-third the average of country-specific portfolios of G7 countries, and one-third that of non-G7 countries (USA+World(CS)), the Sharpe ratio nearly quadruples compared to holding only US market portfolio.

Table 5: Diversification Benefits.

This table presents the mean return, standard deviation, and Sharpe ratio for the US market portfolio, its combination with other country-level market portfolios, and its combination with country-specific portfolios. The USA + G7 (MKT) denotes an equal-weighted portfolio of the US market and the average G7 countries' market portfolios. The USA + Non-G7 (MKT) denotes an equal-weighted portfolio of the US market and the average non-G7 countries' market portfolios. The USA + World (MKT) denotes an equal-weighted portfolio comprising the US market, the average G7 market portfolio, and the average non-G7 market portfolio. The table also presents results where foreign market portfolios are replaced by country-specific portfolios (CS). All returns are in excess of the risk-free rate, and all statistics are annualized.

Portfolio	Mean (%)	Standard Deviation (%)	Sharpe Ratio
USA	8.58	15.30	0.56
G7 (MKT)	5.34	16.31	0.33
Panel A: Passive diversification			
USA + G7 (MKT)	6.96	15.12	0.46
USA + Non G7 (MKT)	8.74	16.11	0.54
USA + World (MKT)	7.61	15.85	0.48
Panel B: Factor-based diversification			
USA + G7 (CS)	12.80	8.89	1.44
USA + Non G7 (CS)	13.50	8.33	1.62
USA + World (CS)	14.67	6.33	2.32

Overall, we find that the individual performance of country-specific portfolios is superb. Furthermore, country-specific portfolios provide significant diversification benefits due to their low correlations, enhancing mean-variance efficiency for investors holding US market.

4.4 Segmentation Portfolios

Next, we examine the properties of segmentation portfolios. In contrast to country-specific portfolios, which are grounded in the classical mean-variance framework for efficiency enhancement as proposed by Markowitz (1952), segmentation portfolios exploit the segmentation of financial markets between two countries. In our framework, this segmentation is manifested through differences in risk premia across markets. Furthermore, the profitability of segmentation portfolios can be interpreted as empirical evidence of market segmentation – an issue central to global financial market equilibrium.

Table 6 presents a detailed analysis of segmentation portfolio returns, divided into two panels: Panel A for G7 countries and Panel B for non-G7 countries. In Panel A, Japan stands out with the highest mean return (31.30%), while other G7 countries also exhibit notably high returns. Standard deviations are approximately 20% annually, close to the calibration. As expected, segmentation portfolios show very low correlations with the US market, since common risks are intentionally designed to be cancelled out. Overall, the strong performance of these portfolios challenges the widely held assumption of tight integration among developed markets.

The same assessment holds for non-G7 countries. In Panel B, Singapore leads with a mean return of 30.55% and a Sharpe ratio of 1.33, while all other non-G7 countries generate statistically significant returns. The strong performance of segmentation portfolios provides further evidence of market segmentation between the US and non-G7 countries.

Next, we investigate whether segmentation portfolio returns are driven by global forces or are independent across different choices of Country 2. To this end, Figure 5 reports correlations among segmentation portfolios. The top-left plot presents a heatmap of correlations for segmentation portfolios of G7 countries, revealing very low correlations among them. For comparison, the top-right plot shows substantially higher correlations among market portfolios of G7 countries, including the USA. Consistent results are found for non-G7 countries: the bottom-left plot indicates low correlations among segmentation portfolios, while the bottom-right plot shows high correlations among market portfolios of non-G7 countries. These results suggest that the drivers of segmentation portfolio returns may not stem from global impediments to market integration.

It is worth noting that the independence of segmentation portfolios is not driven by the independence of common factors across different choices of Country 2. Rather, segmentation portfolios are constructed to hedge out exposures to common factors and to take positions on differences in the risk premia of those factors. In fact, we find strong commonality among the common factors across different choices of Country 2.

Figure 6 illustrates the distribution of average canonical correlations across two sets of common factors. For each pair of countries selected as candidates for Country 2, we identify three common factors with USA (Country 1), compute the canonical correlations between the two sets of three common factors over a short window, and then average them over time. Wider sections indicate higher density, and red dots represent the mean level. The red color denotes pairs of countries (candidates for Country 2) both drawn from non-G7 countries, whereas the blue color denotes pairs drawn from G7 countries. The first canonical correlation remains persistently high–mostly above 0.8–indicating that at least one factor is global in nature. In contrast, the third canonical correlation tends to be below 0.3, suggesting that the number of global factors across the common factor spaces is likely fewer than three.

Table 6: Segmentation portfolio returns

This table presents key financial performance metrics of segmentation portfolio returns, using G7 countries (Panel A) and non-G7 countries (Panel B) as Country 2. The metrics include annualized mean returns (%), annualized standard deviations (%), t-statistics of the average returns, Sharpe ratios, maximum drawdowns, worst and best monthly returns, and correlations with the US market (ρ_{US}). The mean returns, standard deviations and Sharpe ratios are annualized.

Panel A: G7 Countries as Country 2

Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)	ρ_{US}
can	16.83	23.20	4.76	0.73	37.76	-24.67	20.83	0.01
deu	24.41	22.29	6.29	1.09	38.54	-16.78	22.59	-0.03
fra	19.72	22.70	4.66	0.87	42.49	-22.84	29.12	0.02
gbr	23.70	26.19	4.98	0.90	60.44	-54.64	24.26	-0.06
ita	24.15	24.29	5.08	0.99	47.79	-18.26	22.03	-0.05
jpn	31.30	22.15	7.52	1.41	41.08	-14.32	31.56	0.08

Panel B: Non-G7 Countries as Country 2

Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)	ρ_{US}
arg	13.37	23.41	1.83	0.57	55.62	-12.99	38.68	0.03
aus	20.51	27.43	3.96	0.75	46.52	-15.14	93.11	-0.04
aut	12.94	21.76	2.19	0.59	45.91	-22.11	18.16	0.10
bel	15.96	22.27	3.49	0.72	65.37	-16.97	29.67	0.09
bra	10.81	21.19	1.50	0.51	33.75	-12.86	18.93	0.14
chl	16.29	21.29	2.68	0.77	52.18	-19.25	21.53	-0.06
chn	16.30	24.27	2.84	0.67	44.92	-21.20	27.77	-0.01
dnk	20.15	23.92	4.09	0.84	40.87	-23.28	29.77	0.02
esp	19.40	23.20	4.30	0.84	43.04	-24.53	22.99	0.03
hkg	16.34	22.19	3.76	0.74	36.73	-19.48	29.92	0.05
ind	18.08	23.61	4.25	0.77	49.09	-26.59	22.86	-0.02
kor	17.13	21.85	3.59	0.78	73.19	-24.84	28.73	-0.01
mex	12.08	20.80	2.11	0.58	50.88	-16.47	14.61	0.05
nor	17.92	21.82	3.82	0.82	47.38	-21.90	26.51	0.00
pol	20.59	22.96	3.70	0.90	39.45	-21.43	40.48	0.09
sgp	30.55	22.99	6.46	1.33	32.28	-22.26	26.97	0.02
swe	24.72	23.17	5.29	1.07	40.77	-32.50	23.93	-0.04
tha	21.88	23.87	4.42	0.92	57.51	-27.18	22.88	0.05
tur	14.50	20.01	3.58	0.72	35.65	-12.62	26.10	0.00
twn	21.89	21.35	5.19	1.03	36.64	-31.84	18.77	-0.02

5 Additional Empirical Analysis

Other Base Currencies: GBP and JPY Recall that our interest lies in uncovering hidden investment opportunities in international markets for US investors focusing on the US market portfolio. Accordingly, we translate all international stock returns into USD and apply our theoretical procedures.

Yet, in a world dominated by the USD, the return patterns we observe may reflect more than market integration structure-they may be shaped by the hegemony of a USD currency. Because our theoretical framework should not depend on the choice of currency, our results are supposed to be robust to the currency in which returns are denominated. To test this, Tables 13 and 14 report the performance of country-specific and segmentation portfolios in GBP, while Tables 15 and 16 report the corresponding results in JPY.

We find that both the quantitative and qualitative results are comparable to the USD-denominated results reported in Tables 4 and 6, confirming the robustness of our findings and suggesting that the underlying investment opportunities exist independently of the arbitrary dominance of any single currency.

Exchange Rate Risk We conduct additional risk-based analyses in the international market. First, we examine the exposure of our portfolios to changes in exchange rates by estimating

$$r_{ct} = \alpha_c + \beta \times \Delta(\text{Exchange Rate vs. USD})_{ct} + u_{ct},$$

where c indexes country pairs. Table 7 shows that the coefficient on exchange rate changes is not economically significant for either the country-specific portfolios (Panel A) or the segmentation portfolios (Panel B).

Following Verdelhan (2018), who identify two systematic risks in currency markets, we also estimate

$$r_{ct} = \alpha + \beta_1 \times \text{Carry}_t + \beta_2 \times \text{Dollar}_t + u_{ct},$$

where Carry_t denotes the difference in exchange rate changes between high- and low-interest-rate countries, and Dollar_t represents the average change in exchange rates. The results, reported in Table 7, are again not significant for either portfolio type, suggesting that the profitability of our portfolios does not arise from systematic exposure to currency risk factors.

Table 7: Exchange Rate Risk Analysis

This table reports the estimated coefficients from the following regressions: $r_{ct} = \alpha_c + \beta \times \Delta(\text{Exchange Rate vs. USD})_{ct} + u_{ct}$ and $r_{ct} = \alpha_c + \beta_1 \times \text{Carry}_t + \beta_2 \times \text{Dollar}_t + u_{ct}$, where c indexes the 26 country pairs. Carry_t denotes the difference in exchange rate changes between high- and low-interest-rate countries, and Dollar_t represents the average change in exchange rates. Standard errors are clustered by time and reported in parentheses.

	Panel A: Country-specific		Panel B: Segmentation	
	(1)	(2)	(3)	(4)
Δ(Exchange Rate vs. USD)	0.009		0.023	
	(0.035)		(0.040)	
Carry		0.042		0.015
		(0.041)		(0.056)
Dollar		0.004		0.042
		(0.055)		(0.075)

Fixed Number of Common Factors A key parameter in our application is the number of common factors. Although we typically select this number using the bootstrap method of Gonçalves et al. (2025), here we intentionally fix it at four-an arbitrary choice-to test robustness. The results remain largely unchanged, as shown in Tables 11 and 12, which closely mirror the main findings in Tables 4 and 6.

Down-Market Exposure We have emphasized that our country-specific and segmentation portfolios offer distinctive value to US investors, largely owing to their near independence from the US market index. This independence suggests a rare opportunity for diversification in an increasingly synchronized global economy. Yet, as with many perceived advantages, it is natural to question whether this benefit endures when the stock market turns unfavourable to US investors during US market downturns. To examine this, we estimate

$$r_{ct} = \alpha_c + \beta_1 \times \text{Mkt}_t + \beta_2 \times \text{Mkt}_t \times 1(\text{Mkt}_t < \text{threshold}) + u_{ct},$$

where r_{ct} denotes the return of either a country-specific or segmentation portfolio over 26 pairs, and Mkt_t represents the U.S. market excess return. The coefficient β_2 captures the incremental sensitivity of the portfolio to adverse US market conditions. We consider two thresholds for defining such conditions: 0 percent and -5 percent market returns.

Table 8 presents the results. For both country-specific portfolios (Panel A) and segmentation portfolios (Panel B), we find no evidence that returns become more vulnerable during US market downturns. This resilience reinforces our central argument: even when the US market contracts, these portfolios remain to provide benefits to US investors.

Table 8: Resilience to Down-Market.

This table reports the estimated coefficients from the regression $r_{ct} = \alpha_c + \beta_1 \times \text{Mkt}_t + \beta_2 \times \text{Mkt}_t \times 1(\text{Mkt}_t < \text{threshold}) + u_{ct}$, where r_{ct} denotes either a country-specific portfolio or a segmentation portfolio, and Mkt_t represents the U.S. market excess return. The coefficient β_2 captures the additional sensitivity of portfolio returns to the U.S. market when the market return falls below the specified threshold. Thresholds of 0 percent and –5 percent are considered. Standard errors are clustered by time and reported in parentheses.

	Panel A: Country-specific		Panel B: Segmentation	
	(1)	(2)	(3)	(4)
Mkt_t	0.010	0.010	−0.020	−0.012
	(0.030)	(0.022)	(0.040)	(0.029)
$\text{Mkt} \times 1\,(\text{Mkt} < 0)$	−0.013		0.067	
	(0.049)		(0.067)	
$\text{Mkt} \times 1\,(\text{Mkt} < -5\%)$		−0.017		0.063
		(0.036)		(0.050)

Profitability Persistence Our study is motivated by a recurring concern in the international finance literature that the increasing comovement among country-level equity indices in recent decades signals the erosion of international diversification benefits. We examine whether similar concerns - the erosion of benefits over time - are relevant for our proposed portfolio strategies. Our sample, which begins in the late 1980s or 1990s depending on country pairs, enables us to revisit this issue. Specifically, we investigate whether the profitability of the country-specific and segmentation portfolios has declined in more recent periods.

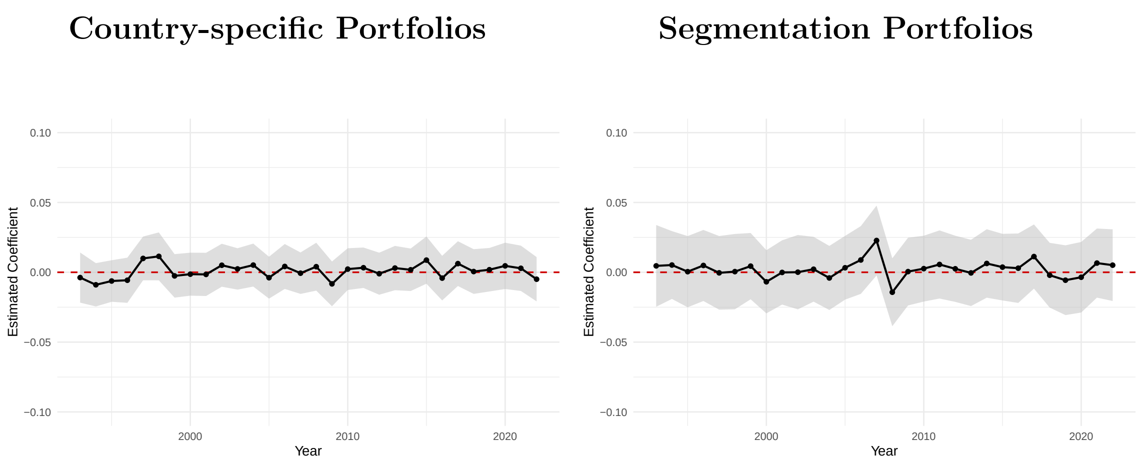
We estimate

$$r_{ct} = \alpha_c + \sum_{\text{year}=1992}^{2022} \beta_{\text{year}} \times 1(t = \text{year}) + u_{ct},$$

where r_{ct} denotes either a country-specific or segmentation portfolio return. The coefficient β_{year} would reveal any specific differences in the realized returns and

Figure 3 presents the estimation results. For both the country-specific (left) and segmented (right) portfolios, the estimated coefficients for the year dummies are not statistically significant at the 5% level in any case. Additionally, there is no clear trend suggesting a decline in profitability in recent years.

Figure 3: Profitability Persistence Over Time.
This table reports the estimated coefficients β_{year} from the regression
$r_{ct} = \alpha_c + \sum_{\text{year}=1992}^{2022} \beta_{\text{year}} \times 1(t = \text{year}) + u_{ct}$, where r_{ct} denotes either a country-specific (left) or segmentation (right) portfolio return. The solid line represents the estimated coefficients, and the gray area indicates the 95% confidence intervals, computed using time-clustered standard errors.



6 Conclusion

This paper proposes a framework for examining pairs of large cross-sectional markets. We apply this framework to international financial markets and further propose alternative approaches for exploiting overseas investment opportunities. Our findings reveal substantial potential for enhancing investment opportunities through international markets. In contrast to the country-level market indexes, which do not improve performance significantly when mixed with the US market, country-specific factors yield significant improvements in mean-variance efficiency, driven by high performance and low correlations with the US market. Naive international diversification over our sample period yields Sharpe ratios between 0.48 and 0.56. Our approach to exploiting country-specific risk premia yields Sharpe ratios from 1.44 to 2.32.

Furthermore, we identify opportunities to capitalize on differences in compensation for common risk factors across countries. Our approach to exploiting pairwise cross-country differences in the pricing of common risks yields Sharpe ratios from 0.73 to 1.41 for G7 countries and from 0.57 to 1.33 for non-G7 countries.

A Additional Tables

Table 9: Country-specific portfolio returns. ($K^c = K_1^s = K_2^s = 1$) This table presents key financial performance metrics of country-specific portfolio returns, using G7 countries (Panel A) and non-G7 countries (Panel B) as Country 2. The metrics include annualized mean returns (%), annualized standard deviations (%), t-statistics of the average returns, Sharpe ratios, maximum drawdowns, worst and best monthly returns, and correlations with the US market (ρ_{US}). The mean returns, standard deviations and Sharpe ratios are annualized.

Panel A: G7 Countries as Country 2								
Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)	ρ_{US}
can	5.08	21.02	1.37	0.24	66.03	-25.50	38.69	0.07
deu	10.76	19.02	2.71	0.57	59.95	-21.80	17.44	-0.09
fra	2.97	20.95	0.72	0.14	80.37	-28.11	19.08	-0.08
gbr	8.17	22.84	1.67	0.36	73.33	-37.25	20.74	-0.02
ita	6.82	17.92	2.26	0.38	50.03	-31.68	20.24	-0.01
jpn	9.86	18.68	2.88	0.53	40.37	-14.80	17.19	0.05
Panel B: Non-G7 Countries as Country 2								
Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)	ρ_{US}
arg	1.13	19.87	0.26	0.06	49.46	-19.91	29.47	0.03
aus	8.01	20.56	2.06	0.39	65.98	-16.68	44.56	-0.06
aut	11.09	18.05	2.30	0.61	46.86	-15.62	17.97	0.01
bel	10.97	21.26	2.62	0.52	52.11	-26.49	30.25	0.02
bra	9.08	22.92	1.13	0.40	40.20	-19.27	35.81	-0.05
chl	10.86	24.22	1.92	0.45	64.15	-19.29	36.73	-0.02
chn	4.42	20.74	1.07	0.21	54.95	-23.81	21.46	-0.03
dnk	12.85	23.78	2.35	0.54	66.36	-22.26	23.15	0.01
esp	14.31	22.22	3.47	0.64	51.86	-29.77	38.19	-0.04
hkg	4.28	21.85	1.07	0.20	76.38	-24.62	31.56	0.03
ind	2.05	22.47	0.50	0.09	87.36	-35.51	20.23	-0.04
kor	12.54	19.79	2.83	0.63	61.35	-20.63	21.74	-0.04
mex	11.24	21.77	2.05	0.52	41.56	-19.73	30.68	-0.09
nor	9.65	19.88	2.07	0.49	58.43	-15.92	22.93	0.02
pol	15.80	20.74	3.01	0.76	42.08	-12.24	38.87	0.08
sgp	9.04	20.44	2.52	0.44	49.35	-20.23	34.87	0.03
swe	2.96	19.58	0.63	0.15	70.16	-27.04	18.81	0.05
tha	11.58	20.04	2.70	0.58	50.42	-20.36	27.08	0.11
tur	13.00	21.29	2.98	0.61	31.42	-18.65	39.26	0.12
twn	4.47	18.71	1.44	0.24	60.71	-16.93	25.97	0.07

Table 10: Segmentation portfolio returns. ($K^c = K_1^s = K_2^s = 1$) This table presents key financial performance metrics of segmentation portfolio returns, using G7 countries (Panel A) and non-G7 countries (Panel B) as Country 2. The metrics include annualized mean returns (%), annualized standard deviations (%), t-statistics of the average returns, Sharpe ratios, maximum drawdowns, worst and best monthly returns, and correlations with the US market (ρ_{US}). The mean returns, standard deviations and Sharpe ratios are annualized.

Panel A: G7 Countries as Country 2								
Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)	ρ_{US}
can	4.52	20.61	1.29	0.22	54.86	-16.68	17.70	0.02
deu	12.36	20.64	3.36	0.60	39.34	-18.23	20.49	0.08
fra	3.02	20.58	0.70	0.15	79.38	-21.25	19.81	0.05
gbr	-1.44	20.19	-0.39	-0.07	67.51	-18.21	19.44	0.03
ita	7.27	21.54	1.87	0.34	60.54	-27.18	23.75	0.07
jpn	3.85	18.49	1.27	0.21	56.57	-14.16	30.02	0.11
Panel B: Non-G7 Countries as Country 2								
Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)	ρ_{US}
arg	9.77	21.20	1.59	0.46	45.94	-18.85	19.05	0.16
aus	8.30	21.35	2.27	0.39	39.93	-14.29	48.57	0.07
aut	9.78	20.85	1.89	0.47	47.79	-19.07	24.72	0.09
bel	7.11	20.87	1.48	0.34	57.79	-18.39	21.71	0.09
bra	9.31	21.93	1.48	0.42	35.74	-19.25	17.62	-0.05
chl	11.19	19.47	2.58	0.57	61.17	-15.67	19.27	0.09
chn	7.06	20.12	1.63	0.35	48.86	-14.85	36.04	0.12
dnk	7.55	20.51	1.66	0.37	52.53	-16.56	19.04	0.02
esp	8.63	18.33	2.59	0.47	54.97	-17.25	18.18	0.03
hkg	14.25	20.19	3.71	0.71	41.88	-19.57	24.61	0.01
ind	3.65	19.78	0.93	0.18	59.51	-18.38	26.09	0.04
kor	10.18	21.73	2.35	0.47	40.88	-21.09	25.95	0.03
mex	10.03	20.31	2.00	0.49	46.49	-14.19	16.65	0.08
nor	9.36	19.99	2.15	0.47	42.06	-17.02	16.45	0.00
pol	5.30	19.81	1.40	0.27	40.54	-17.69	15.79	0.12
sgp	6.19	20.51	1.41	0.30	56.93	-26.05	28.63	0.05
swe	3.94	20.01	1.05	0.20	68.61	-14.69	16.42	0.05
tha	5.10	19.89	1.48	0.26	46.36	-26.01	17.84	0.05
tur	13.63	20.32	3.04	0.67	33.29	-15.53	32.05	0.01
twn	4.31	18.14	2.08	0.24	33.84	-16.32	16.71	-0.03

Table 11: Country-specific portfolio returns. ($K^c = 4$, $K_1^s = K_2^s = 3$) This table presents key financial performance metrics of country-specific portfolio returns, using G7 countries (Panel A) and non-G7 countries (Panel B) as Country 2. The metrics include annualized mean returns (%), annualized standard deviations (%), t-statistics of the average returns, Sharpe ratios, maximum drawdowns, worst and best monthly returns, and correlations with the US market (ρ_{US}). The mean returns, standard deviations and Sharpe ratios are annualized.

Panel A: G7 Countries as Country 2								
Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)	ρ_{US}
can	13.24	18.93	4.11	0.70	41.47	-18.94	19.11	0.09
deu	18.37	18.74	5.42	0.98	38.71	-11.93	18.04	-0.09
fra	14.33	19.00	4.02	0.75	33.04	-16.35	25.17	0.01
gbr	21.41	20.15	5.42	1.06	38.59	-19.09	23.06	0.01
ita	17.96	19.75	4.51	0.91	38.70	-23.01	16.86	-0.03
jpn	13.54	19.65	4.45	0.69	43.77	-22.01	15.63	0.07

Panel B: Non-G7 Countries as Country 2								
Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)	ρ_{US}
arg	13.21	20.55	2.18	0.64	34.15	-22.25	24.63	-0.04
aus	15.11	20.68	4.19	0.73	30.44	-16.23	40.68	0.07
aut	13.41	19.37	2.75	0.69	35.17	-15.52	20.70	-0.06
bel	21.74	20.72	5.97	1.05	33.22	-14.78	25.93	-0.09
bra	7.35	21.20	1.27	0.35	27.38	-17.79	18.59	0.08
chl	10.12	21.11	2.37	0.48	58.90	-14.36	27.19	-0.02
chn	17.46	20.77	3.46	0.84	47.28	-19.60	21.76	-0.02
dnk	24.67	21.98	5.30	1.12	32.41	-18.43	16.90	-0.05
esp	13.64	21.45	3.80	0.64	44.63	-22.17	24.19	0.06
hkg	16.57	19.94	5.13	0.83	59.41	-23.37	26.09	0.07
ind	15.47	19.32	3.75	0.80	50.16	-16.09	21.62	-0.07
kor	11.77	19.84	2.62	0.59	51.05	-15.89	20.20	0.01
mex	19.38	20.26	4.14	0.96	33.11	-21.93	20.86	-0.01
nor	16.59	21.43	3.89	0.77	60.58	-22.76	18.04	-0.01
pol	16.49	21.05	3.16	0.78	43.99	-22.24	23.41	-0.09
sgp	22.00	19.75	5.25	1.11	38.58	-13.83	29.82	0.09
swe	18.57	23.10	4.02	0.80	49.14	-28.07	37.55	-0.04
tha	12.90	18.93	3.74	0.68	40.64	-25.35	24.03	-0.02
tur	16.97	20.00	3.86	0.85	32.10	-18.28	22.87	-0.05
twn	10.27	19.50	2.59	0.53	49.63	-19.48	16.50	0.00

Table 12: Segmentation portfolio returns. ($K^c = 4,\ K_1^s = K_2^s = 3$) This table presents key financial performance metrics of segmentation portfolio returns, using G7 countries (Panel A) and non-G7 countries (Panel B) as Country 2. The metrics include annualized mean returns (%), annualized standard deviations (%), t-statistics of the average returns, Sharpe ratios, maximum drawdowns, worst and best monthly returns, and correlations with the US market (ρ_{US}). The mean returns, standard deviations and Sharpe ratios are annualized.

Panel A: G7 Countries as Country 2								
Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)	ρ_{US}
can	16.83	23.20	4.76	0.73	37.76	-24.67	20.83	0.01
deu	25.61	22.30	6.37	1.15	38.54	-16.78	22.59	-0.02
fra	19.72	22.70	4.66	0.87	42.49	-22.84	29.12	0.02
gbr	23.70	26.19	4.98	0.90	60.44	-54.64	24.26	-0.06
ita	25.58	24.14	4.63	1.06	72.65	-33.08	22.03	-0.05
jpn	31.30	22.15	7.52	1.41	41.08	-14.32	31.56	0.08
Panel B: Non-G7 Countries as Country 2								
Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)	ρ_{US}
arg	24.92	20.07	4.09	1.24	24.60	-11.22	20.24	0.06
aus	22.62	27.12	4.49	0.83	46.52	-15.16	93.11	-0.04
aut	26.06	18.08	5.51	1.44	28.11	-12.99	18.62	0.00
bel	26.72	21.38	6.88	1.25	27.96	-17.60	29.67	-0.01
bra	16.52	21.93	2.35	0.75	28.62	-14.66	18.93	0.14
chl	29.35	21.81	5.70	1.35	36.07	-16.90	23.04	-0.05
chn	19.49	22.83	3.89	0.85	37.70	-22.58	27.77	0.00
dnk	23.78	23.99	4.76	0.99	33.82	-23.28	29.77	0.00
esp	27.18	21.27	6.34	1.28	44.94	-24.66	18.91	0.04
hkg	16.45	21.41	4.11	0.77	36.40	-19.84	26.17	0.11
ind	27.26	22.58	6.74	1.21	34.01	-20.92	27.03	-0.06
kor	20.25	21.69	4.49	0.93	51.07	-12.74	28.73	0.02
mex	24.09	20.15	5.02	1.20	21.48	-16.11	17.78	0.06
nor	20.06	21.18	4.56	0.95	42.86	-14.22	26.51	0.00
pol	23.97	22.85	4.62	1.05	33.16	-21.43	40.48	0.08
sgp	33.70	23.06	6.73	1.46	34.75	-18.27	31.24	0.01
swe	27.47	23.09	5.75	1.19	40.77	-32.50	25.22	-0.03
tha	24.30	24.74	4.90	0.98	61.97	-27.18	46.23	0.09
tur	20.84	21.62	4.48	0.96	33.36	-12.27	30.98	0.07
twn	25.39	21.65	6.08	1.17	34.15	-31.84	20.14	0.03

B Additional Figures

Figure 4: Canonical correlations. The vase-shaped figure illustrates the distribution of canonical correlations between the systematic factors of Country 1 (USA) and those of Country 2. Wider sections indicate higher density and red dots represent the average correlation. The top-left panel presents results with Canada as Country 2. The top-right panel shows results for G7 countries, the bottom-left for non-G7 countries, and the bottom-right for all 26 countries in the sample.

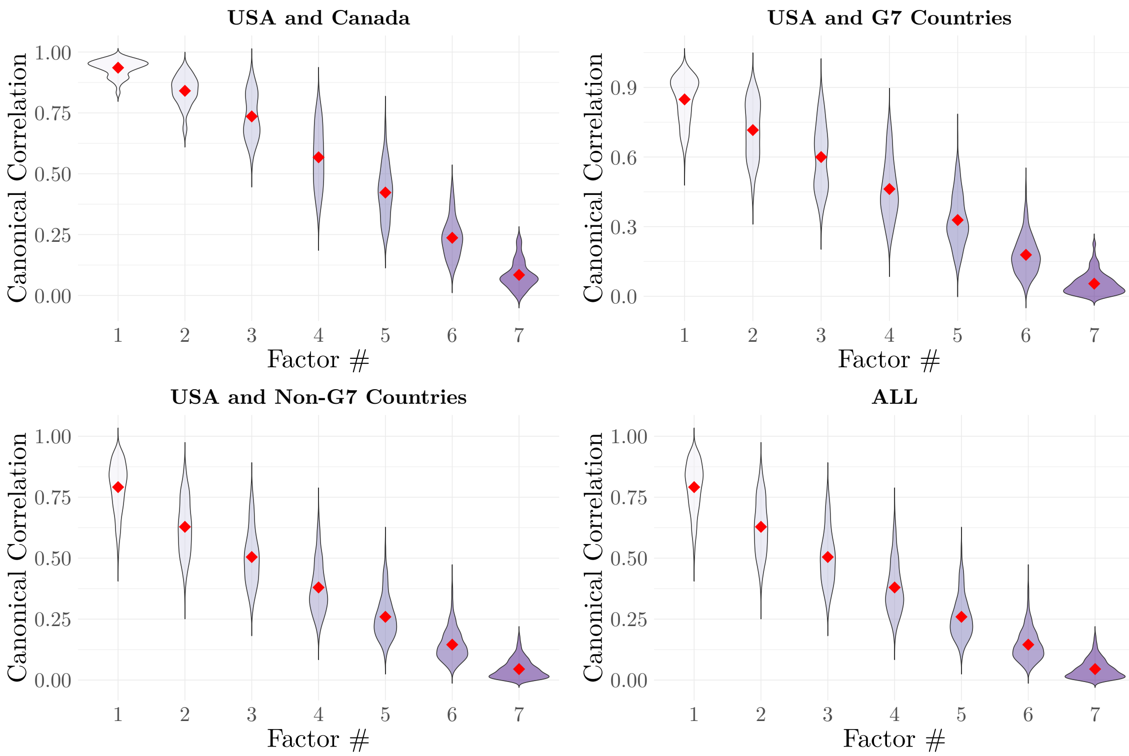


Figure 5: Correlation Heatmaps of Segmentation and Market Portfolios.

This figure displays heatmaps of correlations across different portfolios. The top-left panel shows correlations among country-specific portfolios of G7 countries. The top-right panel presents correlations among market portfolios of G7 countries, including USA. The bottom-left panel illustrates correlations among country-specific portfolios of non-G7 countries. The bottom-right panel depicts correlations among market portfolios of non-G7 countries.

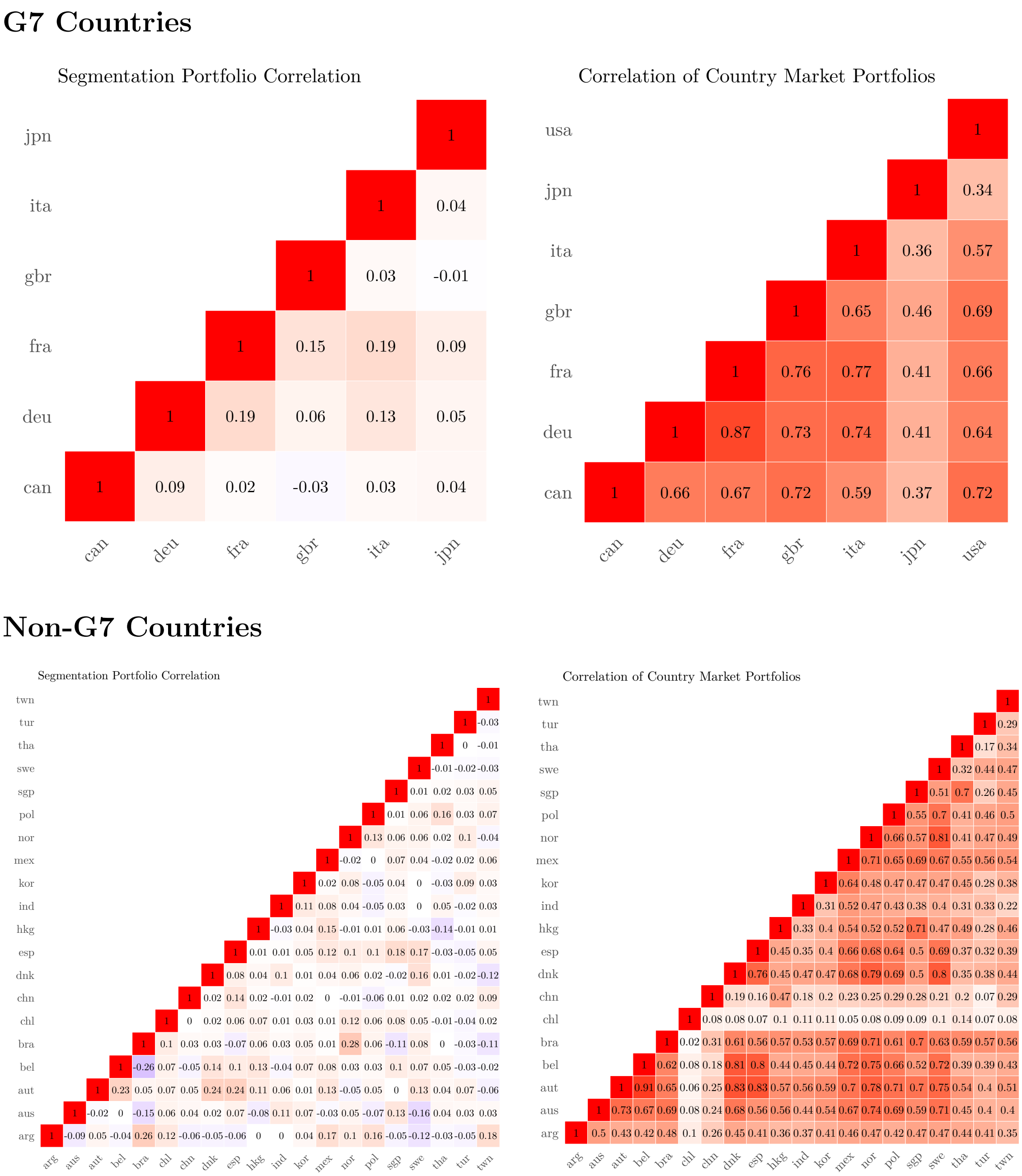
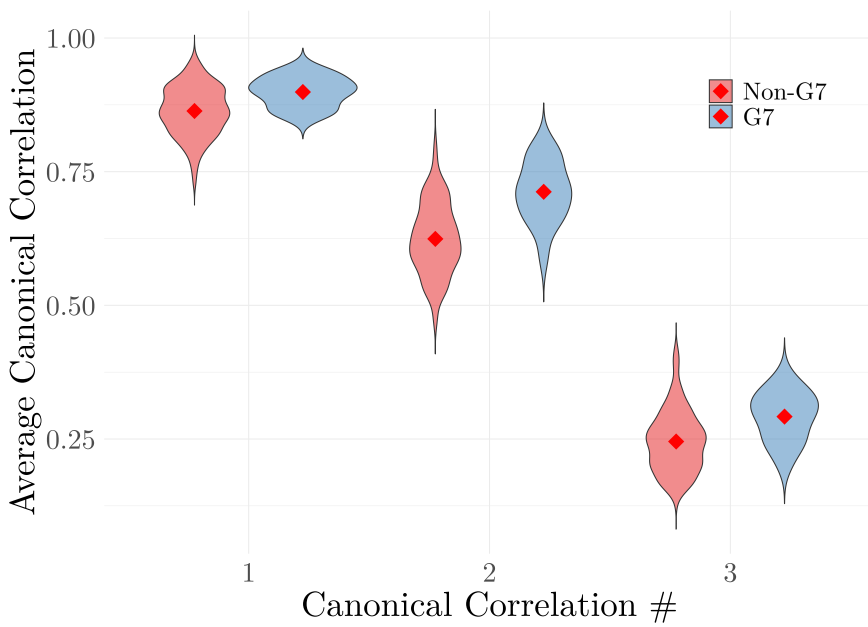


Figure 6: Canonical Correlation across Common Factors.

The vase-shaped figure illustrates the distribution of average canonical correlations across two sets of common factors. We select two countries as candidates for Country 2 and identify three common factors with USA (Country 1) for each of them. Then, we compute the canonical correlations between the two sets of three common factors over a short window and take their average over time. Wider sections indicate higher density, and red dots represent the mean level. The red color denotes pairs of countries (candidates for Country 2) both drawn from non-G7 countries, whereas the blue color denotes pairs drawn from G7 countries.



C Proofs

Lemma C.1. *Let* $\mathbf{L}_g$ *be a lower triangular matrix such that* $\mathbf{H}_g'\mathbf{J}_T\mathbf{H}_g = \mathbf{L}_g\mathbf{L}_g'$*. From the eigendecompotion of* $\mathbf{L}_g'\mathbf{V}_{\Theta_g}\mathbf{L}_g$*, find* $\mathbf{U}_g$ *such that* $\mathbf{L}_g'\mathbf{V}_{\Theta_g}\mathbf{L}_g = \mathbf{U}_g\mathbf{D}\mathbf{U}_g'$*, where* $\mathbf{D}$ *is a diagonal matrix and* $\mathbf{U}_g'\mathbf{U}_g = \mathbf{I}_{K_g}$*. Define* $\mathcal{O}_g$ *as* $\mathbf{L}_g'^{-1}\mathbf{U}_g$*. Then, it holds*
(i) $\mathcal{O}_g'\mathbf{H}_g'\mathbf{J}_T\mathbf{H}_g\mathcal{O}_g$ *is an identity matrix, and*
(ii) $\mathcal{O}_g^{-1}\mathbf{V}_{\Theta_g}\mathcal{O}_g^{-1\prime}$ *is a diagonal matrix.*

Proof First, we show (i) $\mathcal{O}_g'\mathbf{H}_g'\mathbf{J}_T\mathbf{H}_g\mathcal{O}_g$ is an identity matrix. Note that

$$\mathcal{O}_g'\mathbf{H}_g'\mathbf{J}_T\mathbf{H}_g\mathcal{O}_g = \mathbf{U}_g'\mathbf{L}_g^{-1}\mathbf{L}_g\mathbf{L}_g'\mathbf{L}_g'^{-1}\mathbf{U}_g = \mathbf{U}_g'\mathbf{U}_g = \mathbf{I}_{K_g},$$

where the last equality is from the property of $\mathbf{U}_g$.

We move to the next claim. Note that $\mathcal{O}_g^{-1} = \mathbf{U}_g^{-1}\mathbf{L}_g' = \mathbf{U}_g'\mathbf{L}_g'$, which in turn gives,

$$\mathcal{O}_g^{-1}\mathbf{V}_{\Theta_g}\mathcal{O}_g^{-1\prime} = \mathbf{U}_g'\mathbf{L}_g'\mathbf{V}_{\Theta_g}\mathbf{L}_g\mathbf{U}_g = \mathbf{U}_g'\mathbf{U}_g\mathbf{D}\mathbf{U}_g'\mathbf{U}_g = \mathbf{D},$$

where the second equality is from the eigendecompotion of $\mathbf{L}_g'\mathbf{V}_{\Theta_g}\mathbf{L}_g$ and the last equality is from the property of $\mathbf{U}_g$. This completes the proof of the lemma. □

Proof of Theorem 2.1 The following three steps complete the proof.

Step 1. $\frac{\widehat{\mathbf{R}}_g'\widehat{\mathbf{R}}_g}{N_g} \xrightarrow{p} \mathbf{J}_T\mathbf{H}_g\mathbf{V}_{\Theta_g}\mathbf{H}_g'\mathbf{J}_T$: From (2.8), we have that

$$\widehat{\mathbf{R}}_g = j_1 + j_2 + j_3 + j_4, \tag{C.1}$$

where $j_1 = \mathbf{X}_g\Theta_g\mathbf{H}_g'\mathbf{J}_T$, $j_2 = \mathbf{P}_g\Gamma_g^c\mathbf{F}'\mathbf{J}_T$, $j_3 = \mathbf{P}_g\Gamma_g^s\mathbf{G}_g'\mathbf{J}_T$, and $j_4 = \mathbf{P}_g\mathbf{E}_g\mathbf{J}_T$. Then,

$$\frac{\widehat{\mathbf{R}}_g'\widehat{\mathbf{R}}_g}{N_g} = \sum_{k,l=1,2,3,4}\frac{j_k'j_l}{N_g}. \tag{C.2}$$

Note that

$$\frac{j_1'j_1}{N_g} = \mathbf{J}_T\mathbf{H}_g\frac{\Theta_g'\mathbf{X}_g'\mathbf{X}_g\Theta_g}{N_g}\mathbf{H}_g'\mathbf{J}_T \xrightarrow{p} \mathbf{J}_T\mathbf{H}_g\mathbf{V}_{\Theta_g}\mathbf{H}_g'\mathbf{J}_T, \tag{C.3}$$

where the limit is from Assumption 2(i), and that

$$\frac{j_2'j_2}{N_g} = \frac{\mathbf{J}_T\mathbf{F}\Gamma_g^{c\prime}\mathbf{P}_g\Gamma_g^c\mathbf{F}'\mathbf{J}_T}{N_g} = \mathbf{J}_T\mathbf{F}\frac{\Gamma_g^{c\prime}\mathbf{X}_g}{N_g}\left(\frac{\mathbf{X}_g'\mathbf{X}_g}{N_g}\right)^{-1}\frac{\mathbf{X}_g'\Gamma_g^c}{N_g}\mathbf{F}'\mathbf{J}_T \xrightarrow{p} \mathbf{0}_{T\times T}, \tag{C.4}$$

where the limit is from Assumptions 1(i) and 1(ii), and that

$$\frac{j_3'j_3}{N_g} = \frac{\mathbf{J}_T\mathbf{G}_g\Gamma_g^{s\prime}\mathbf{P}_g\Gamma_g^s\mathbf{G}_g'\mathbf{J}_T}{N_g} = \mathbf{J}_T\mathbf{G}_g\frac{\Gamma_g^{s\prime}\mathbf{X}_g}{N_g}\left(\frac{\mathbf{X}_g'\mathbf{X}_g}{N_g}\right)^{-1}\frac{\mathbf{X}_g'\Gamma_g^s}{N_g}\mathbf{G}_g'\mathbf{J}_T \xrightarrow{p} \mathbf{0}_{T\times T}, \tag{C.5}$$

where the limit is from Assumptions 1(i) and 1(ii), and that

$$\frac{j_4'j_4}{N_g} = \frac{\mathbf{J}_T\mathbf{E}_g'\mathbf{P}_g\mathbf{E}_g\mathbf{J}_T}{N_g} = \mathbf{J}_T\frac{\mathbf{E}_g'\mathbf{X}_g}{N_g}\left(\frac{\mathbf{X}_g'\mathbf{X}_g}{N_g}\right)^{-1}\frac{\mathbf{X}_g'\mathbf{E}_g}{N_g}\mathbf{J}_T \xrightarrow{p} \mathbf{0}_{T\times T}, \tag{C.6}$$

where the limit is from Assumptions 1(i) and 1(ii). From (C.3)-(C.6) and the submultiplicativity of Frobenius norm, we have that

$$\|\frac{j_1'j_l}{N_g}\| \leq \|\frac{j_1}{\sqrt{N_g}}\|\|\frac{j_l}{\sqrt{N_g}}\| = \sqrt{\operatorname{tr}\left(\frac{j_1'j_1}{N_g}\right)\operatorname{tr}\left(\frac{j_l'j_l}{N_g}\right)} \xrightarrow{p} 0 \tag{C.7}$$

for $l = 2, 3, 4$ and that

$$\|\frac{j_k'j_l}{N_g}\| \leq \|\frac{j_k}{\sqrt{N_g}}\|\|\frac{j_l}{\sqrt{N_g}}\| = \sqrt{\operatorname{tr}\left(\frac{j_k'j_k}{N_g}\right)\operatorname{tr}\left(\frac{j_l'j_l}{N_g}\right)} \xrightarrow{p} 0 \tag{C.8}$$

for $k, l = 2, 3, 4$. By plugging (C.3)-(C.8) into (C.2), we confirm the claim of Step 1.

Step 2. The k-th column of $\mathbf{J}_T\mathbf{H}_g\mathcal{O}_g$, where $\mathcal{O}_g$ is given by Lemma C.1, is the k-th eigenvector of $\mathbf{J}_T\mathbf{H}_g\mathbf{V}_{\Theta_g}\mathbf{H}_g'\mathbf{J}_T$: Note that

$$\mathbf{J}_T\mathbf{H}_g\mathbf{V}_{\Theta_g}\mathbf{H}_g'\mathbf{J}_T = \mathbf{J}_T\mathbf{H}_g\mathcal{O}_g\left(\mathcal{O}_g^{-1}\mathbf{V}_{\Theta_g}\mathcal{O}_g^{-1\prime}\right)\mathcal{O}_g'\mathbf{H}_g'\mathbf{J}_T. \tag{C.9}$$

Given the properties of $\mathcal{O}_g$ in Lemma C.1, the claim directly follows. Step 3. $\widehat{\mathbf{H}}_g \xrightarrow{p} \mathbf{J}_T\mathbf{H}_g\mathcal{O}_g$: The claim holds due to the continuity of the eigendecomposition. This completes the proof of the theorem. □

Lemma C.2. *Define* $\widehat{\Sigma}_{H_g}$ *and* $\widehat{\Sigma}_{H_{12}}$ *as* $\frac{\widehat{\mathbf{H}}_g'\mathbf{J}_T\widehat{\mathbf{H}}_g}{T}$ *and* $\frac{\widehat{\mathbf{H}}_1'\mathbf{J}_T\widehat{\mathbf{H}}_2}{T}$*, respectively, where* $\widehat{\mathbf{H}}_g$ *is given in Theorem 2.1. Let* $\widehat{\mathbf{L}}_{H_g}$ *be a lower triangular matrix such that* $\widehat{\Sigma}_{H_g} = \widehat{\mathbf{L}}_{H_g}\widehat{\mathbf{L}}_{H_g}'$

for $g = 1, 2$. From the singular value decomposition of $\widehat{\mathbf{L}}_{H_1}^{-1}\widehat{\Sigma}_{H_{12}}\widehat{\mathbf{L}}_{H_2}^{\prime -1}$, we have $\widehat{\mathbf{V}}_1$, $\widehat{\mathbf{V}}_2$, and $\widehat{\Sigma}$ such that

$$\widehat{\mathbf{L}}_{H_1}^{-1}\widehat{\Sigma}_{H_{12}}\widehat{\mathbf{L}}_{H_2}^{\prime -1} = \widehat{\mathbf{V}}_1\widehat{\Sigma}\widehat{\mathbf{V}}_2^{\prime}, \tag{C.10}$$

where $\widehat{\mathbf{V}}_1^{\prime}\widehat{\mathbf{V}}_1 = \mathbf{I}_{K_1}$, $\widehat{\mathbf{V}}_2^{\prime}\widehat{\mathbf{V}}_2 = \mathbf{I}_{K_2}$, and $\widehat{\Sigma}$ is a rectangular diagonal matrix with nonnegative diagonal elements in a decreasing order. Let $\widehat{\Sigma}_{kk}$ be the k-th diagonal element of $\widehat{\Sigma}$. Then, it holds that
(i) $\widehat{\mathbf{W}}_1 = \widehat{\mathbf{L}}_{H_1}^{\prime -1}\widehat{\mathbf{V}}_1$, and the k-th largest eigenvalue of $\widehat{\Sigma}_{H_1}^{-1}\widehat{\Sigma}_{H_{12}}\widehat{\Sigma}_{H_2}^{-1}\widehat{\Sigma}_{H_{12}}^{\prime}$ is $\left(\widehat{\Sigma}_{kk}\right)^2$,
(ii) $\widehat{\mathbf{W}}_2 = \widehat{\mathbf{L}}_{H_2}^{\prime -1}\widehat{\mathbf{V}}_2$, and the k-th largest eigenvalue of $\widehat{\Sigma}_{H_2}^{-1}\widehat{\Sigma}_{H_{12}}^{\prime}\widehat{\Sigma}_{H_1}^{-1}\widehat{\Sigma}_{H_{12}}$ is $\left(\widehat{\Sigma}_{kk}\right)^2$,
(iii) the correlation between $\widehat{\mathbf{H}}_1\widehat{\mathbf{W}}_1^{(m)}$ and $\widehat{\mathbf{H}}_2\widehat{\mathbf{W}}_2^{(n)}$ is the (m, n)-th element of $\widehat{\Sigma}$, where $\widehat{\mathbf{W}}_1^{(m)}$ is the m-th column of $\widehat{\mathbf{W}}_1$ and $\widehat{\mathbf{W}}_2^{(n)}$ is the n-th column of $\widehat{\mathbf{W}}_2$.

Proof First, we prove the claim (i). Recall that $\widehat{\mathbf{W}}_1$ is defined by the two conditions: (a) each column of $\widehat{\mathbf{W}}_1$ is an eigenvector of $\widehat{\Sigma}_{H_1}^{-1}\widehat{\Sigma}_{H_{12}}\widehat{\Sigma}_{H_2}^{-1}\widehat{\Sigma}_{H_{12}}^{\prime}$ with the corresponding eigenvalues in the diagonal element of $\widehat{\Sigma}\widehat{\Sigma}^{\prime}$ and (b) $\widehat{\mathbf{W}}_1^{\prime}\widehat{\Sigma}_{H_1}\widehat{\mathbf{W}}_1 = \mathbf{I}_{K_1}$.

From $\widehat{\mathbf{L}}_{H_1}^{-1}\widehat{\Sigma}_{H_{12}}\widehat{\mathbf{L}}_{H_2}^{\prime -1} = \widehat{\mathbf{V}}_1\widehat{\Sigma}\widehat{\mathbf{V}}_2^{\prime}$, we have that

$$\left(\widehat{\mathbf{L}}_{H_1}^{-1}\widehat{\Sigma}_{H_{12}}\widehat{\mathbf{L}}_{H_2}^{\prime -1}\right)\left(\widehat{\mathbf{L}}_{H_1}^{-1}\widehat{\Sigma}_{H_{12}}\widehat{\mathbf{L}}_{H_2}^{\prime -1}\right)^{\prime} = \left(\widehat{\mathbf{V}}_1\widehat{\Sigma}\widehat{\mathbf{V}}_2^{\prime}\right)\left(\widehat{\mathbf{V}}_1\widehat{\Sigma}\widehat{\mathbf{V}}_2^{\prime}\right)^{\prime},$$

which, in conjunction with $\widehat{\mathbf{V}}_2^{\prime}\widehat{\mathbf{V}}_2 = \mathbf{I}_{K_2}$, gives

$$\widehat{\mathbf{L}}_{H_1}^{-1}\widehat{\Sigma}_{H_{12}}\widehat{\mathbf{L}}_{H_2}^{\prime -1}\widehat{\mathbf{L}}_{H_2}^{-1}\widehat{\Sigma}_{H_{12}}^{\prime}\widehat{\mathbf{L}}_{H_1}^{\prime -1} = \widehat{\mathbf{V}}_1\widehat{\Sigma}\widehat{\Sigma}^{\prime}\widehat{\mathbf{V}}_1^{\prime},$$

yielding, along with $\widehat{\mathbf{V}}_1^{\prime}\widehat{\mathbf{V}}_1 = \mathbf{I}_{K_1}$, that

$$\widehat{\Sigma}_{H_1}^{-1}\widehat{\Sigma}_{H_{12}}\widehat{\Sigma}_{H_2}^{-1}\widehat{\Sigma}_{H_{12}}^{\prime}\left(\widehat{\mathbf{L}}_{H_1}^{\prime -1}\widehat{\mathbf{V}}_1\right) = \widehat{\mathbf{L}}_{H_1}^{\prime -1}\widehat{\mathbf{V}}_1\widehat{\Sigma}\widehat{\Sigma}^{\prime}\widehat{\mathbf{V}}_1^{\prime}\widehat{\mathbf{V}}_1 = \left(\widehat{\mathbf{L}}_{H_1}^{\prime -1}\widehat{\mathbf{V}}_1\right)\widehat{\Sigma}\widehat{\Sigma}^{\prime},$$

This shows that the k-th column of $\widehat{\mathbf{L}}_{H_1}^{\prime -1}\widehat{\mathbf{V}}_1$ is the eigenvector of $\widehat{\Sigma}_{H_1}^{-1}\widehat{\Sigma}_{H_{12}}\widehat{\Sigma}_{H_2}^{-1}\widehat{\Sigma}_{H_{12}}^{\prime}$ corresponding to the eigenvalue in the k-th diagonal element of $\widehat{\Sigma}\widehat{\Sigma}^{\prime}$. This confirms the condition (a) for $\widehat{\mathbf{W}}$. The condition (b) is satisfied due to that $\widehat{\mathbf{V}}_1^{\prime}\widehat{\mathbf{L}}_{H_1}^{-1}\widehat{\Sigma}_{H_1}\widehat{\mathbf{L}}_{H_1}^{\prime -1}\widehat{\mathbf{V}}_1 = \widehat{\mathbf{V}}_1^{\prime}\widehat{\mathbf{L}}_{H_1}^{-1}\widehat{\mathbf{L}}_{H_1}\widehat{\mathbf{L}}_{H_1}^{\prime -1}\widehat{\mathbf{L}}_{H_1}^{\prime -1}\widehat{\mathbf{V}}_1 = \widehat{\mathbf{V}}_1^{\prime}\widehat{\mathbf{V}}_1 = \mathbf{I}_{K_1}$. Hence, the claim (i) holds. The claim (ii) follows due to symmetry.

Next, we move to the claim (iii). From the claims (i) and (ii), we showed that

$\widehat{\mathbf{W}}_g'\widehat{\Sigma}_{H_g}\widehat{\mathbf{W}}_g = \mathbf{I}_{K_g}$, implying the variances of $\widehat{\mathbf{H}}_1\widehat{\mathbf{W}}_1^{(m)}$ and $\widehat{\mathbf{H}}_2\widehat{\mathbf{W}}_2^{(n)}$ are 1. Hence, the correlation of the two would be the covariance of the two, $\widehat{\mathbf{W}}_1^{(m)\prime}\widehat{\Sigma}_{H_{12}}\widehat{\mathbf{W}}_2^{(n)}$. Note that

$$\widehat{\mathbf{W}}_1'\widehat{\Sigma}_{H_{12}}\widehat{\mathbf{W}}_2 = \widehat{\mathbf{V}}_1'\widehat{\mathbf{L}}_{H_1}^{-1}\widehat{\Sigma}_{H_{12}}\widehat{\mathbf{L}}_{H_2}^{\prime -1}\widehat{\mathbf{V}}_2 = \widehat{\Sigma}, \tag{C.11}$$

which confirms that the correlation of the two is the (m,n)-th element of $\widehat{\Sigma}$. This completes the proof of the lemma. □

The next lemma reveals the N-limit behavior of CCA in our setup.

Lemma C.3. *Let* $\mathbf{V}_g$, $\breve{\mathbf{L}}_{H_g}$, $\breve{\Sigma}_{H_{12}}$ *and* Σ *be the N-limits of* $\widehat{\mathbf{V}}_g$, $\widehat{\mathbf{L}}_{H_g}$, $\widehat{\Sigma}_{H_{12}}$ *and* $\widehat{\Sigma}$, *respectively, where* $\widehat{\mathbf{V}}_g$, $\widehat{\mathbf{L}}_{H_g}$, $\widehat{\Sigma}_{H_{12}}$ *and* $\widehat{\Sigma}$ *are defined in Lemma C.2. Let* $\mathbf{L}_F$ *and* $\mathbf{L}_{G_g}$ *be lower triangular matrices such that* $\frac{\mathbf{F}'\mathbf{J}_T\mathbf{F}}{T} = \mathbf{L}_F\mathbf{L}_F'$ *and* $\frac{\mathbf{G}_g'\mathbf{J}_T\mathbf{G}_g}{T} = \mathbf{L}_{G_g}\mathbf{L}_{G_g}'$. *From the singular value decomposition of* $\mathbf{L}_{G_1}^{-1}\frac{\mathbf{G}_1'\mathbf{J}_T\mathbf{G}_2}{T}\mathbf{L}_{G_2}^{\prime -1}$, *we have* $\mathbf{Z}_1$, $\mathbf{Z}_2$, *and* Δ *such that*

$$\mathbf{L}_{G_1}^{-1}\frac{\mathbf{G}_1'\mathbf{J}_T\mathbf{G}_2}{T}\mathbf{L}_{G_2}^{\prime -1} = \mathbf{Z}_1\Delta\mathbf{Z}_2', \tag{C.12}$$

where $\mathbf{Z}_1'\mathbf{Z}_1 = \mathbf{I}_{K_1^s}$, $\mathbf{Z}_2'\mathbf{Z}_2 = \mathbf{I}_{K_2^s}$, *and* Δ *is a rectangular diagonal matrix with nonnegative diagonal elements in decreasing order. Then, the followings hold:*
(i) Any (k,k)*-th element of* Δ *is less than 1,*
(ii) $\mathbf{V}_g = \breve{\mathbf{L}}_{H_g}^{-1}\mathcal{O}_g'\begin{bmatrix} \mathbf{L}_F\mathcal{R}' & \mathbf{0}_{K_c\times K_g^s} \\ \mathbf{0}_{K_g^s\times K_c} & \mathbf{L}_{G_g}\mathbf{Z}_g \end{bmatrix}$, *where* $\mathcal{R}\mathcal{R}' = \mathbf{I}_{K^c}$,
(iii) $\Sigma = \begin{bmatrix} \mathbf{I}_{K_c} & \mathbf{0}_{K_c\times K_2^s} \\ \mathbf{0}_{K_1^s\times K_c} & \Delta \end{bmatrix}$.

Proof We start with the claim (i). Given $\mathbf{Z}_g'\mathbf{L}_{G_g}^{-1}\frac{\mathbf{G}_g'\mathbf{J}_T\mathbf{G}_g}{T}\mathbf{L}_{G_g}^{\prime -1}\mathbf{Z}_g = \mathbf{I}_{K_g^s}$, any linear combination of $\mathbf{G}_g$, denoted as $\mathbf{G}_g\mathbf{x}_g$, will have unit variance if $\mathbf{x}_g = \mathbf{L}_{G_g}^{\prime -1}\mathbf{Z}_g\mathbf{y}_g$ where $\mathbf{y}_g'\mathbf{y}_g = 1$. Hence, set $\mathbf{y}_g'\mathbf{y}_g = 1$ for $g = 1, 2$. Then, the covariance of $\mathbf{G}_1\mathbf{x}_1$ and $\mathbf{G}_2\mathbf{x}_2$ would be their correlation. Note that

$$\begin{aligned}\mathbf{x}_1'\frac{\mathbf{G}_1'\mathbf{J}_T\mathbf{G}_2}{T}\mathbf{x}_2 &= \mathbf{y}_1'\mathbf{Z}_1'\mathbf{L}_{G_1}^{-1}\frac{\mathbf{G}_1'\mathbf{J}_T\mathbf{G}_2}{T}\mathbf{L}_{G_2}^{\prime -1}\mathbf{Z}_2\mathbf{y}_2 \\ &= \mathbf{y}_1'\mathbf{Z}_1'\mathbf{Z}_1\Delta\mathbf{Z}_2'\mathbf{Z}_2\mathbf{y}_2 = \mathbf{y}_1'\Delta\mathbf{y}_2,\end{aligned} \tag{C.13}$$

where the second equality is from (C.12). It follows that the maximum of $\mathbf{y}_1'\Delta\mathbf{y}_2$ would be the maximum element of Δ. Because $\mathbf{y}_1'\Delta\mathbf{y}_2$ is a correlation between $\mathbf{G}_1\mathbf{x}_1$ and $\mathbf{G}_2\mathbf{x}_2$, the maximum of $\mathbf{y}_1'\Delta\mathbf{y}_2$ would be less than one due to Assumption 2(iii), thus proving the claim (i).

Next, we show the claims (ii) and (iii). From Lemma C.2, recall $\widehat{\mathbf{L}}_{H_1}^{-1}\widehat{\Sigma}_{H_{12}}\widehat{\mathbf{L}}_{H_2}'^{-1} = \widehat{\mathbf{V}}_1\widehat{\Sigma}\widehat{\mathbf{V}}_2'$. Hence, it should hold that

$$\breve{\mathbf{L}}_{H_1}^{-1}\breve{\Sigma}_{H_{12}}\breve{\mathbf{L}}_{H_2}'^{-1} = \mathbf{V}_1\Sigma\mathbf{V}_2'. \tag{C.14}$$

We prove the claims (ii) and (iii) by verifying the equation above along with $\mathbf{V}_g'\mathbf{V}_g = \mathbf{I}_{Kg}$. First, note that

$$\begin{aligned}
\mathbf{V}_g'\mathbf{V}_g &= \begin{bmatrix} \mathcal{R}\mathbf{L}_F' & \mathbf{0}_{K_c\times K_g^s} \\ \mathbf{0}_{K_g^s\times K_c} & \mathbf{Z}_g'\mathbf{L}_{G_g}' \end{bmatrix} \mathcal{O}_g\breve{\mathbf{L}}_{H_g}^{-1\prime}\breve{\mathbf{L}}_{H_g}^{-1}\mathcal{O}_g' \begin{bmatrix} \mathbf{L}_F\mathcal{R}' & \mathbf{0}_{K_c\times K_g^s} \\ \mathbf{0}_{K_g^s\times K_c} & \mathbf{L}_{G_g}\mathbf{Z}_g \end{bmatrix} \\
&= \begin{bmatrix} \mathcal{R}\mathbf{L}_F' & \mathbf{0}_{K_c\times K_g^s} \\ \mathbf{0}_{K_g^s\times K_c} & \mathbf{Z}_g'\mathbf{L}_{G_g}' \end{bmatrix} \begin{bmatrix} (\mathbf{L}_F\mathbf{L}_F')^{-1} & \mathbf{0}_{K_c\times K_g^s} \\ \mathbf{0}_{K_g^s\times K_g} & (\mathbf{L}_{G_1}\mathbf{L}_{G_1}')^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{L}_F\mathcal{R}' & \mathbf{0}_{K_c\times K_g^s} \\ \mathbf{0}_{K_g^s\times K_c} & \mathbf{L}_{G_g}\mathbf{Z}_g \end{bmatrix} = \mathbf{I}_{Kg}.
\end{aligned} \tag{C.15}$$

Next, note that

$$\begin{aligned}
&\breve{\mathbf{L}}_{H_1}^{-1}\breve{\Sigma}_{H_{12}}\breve{\mathbf{L}}_{H_2}'^{-1} = \breve{\mathbf{L}}_{H_1}^{-1}\mathcal{O}_1'\frac{\mathbf{H}_1'\mathbf{J}_T\mathbf{H}_2}{T}\mathcal{O}_2\breve{\mathbf{L}}_{H_2}'^{-1} \\
=&\breve{\mathbf{L}}_{H_1}^{-1}\mathcal{O}_1' \begin{bmatrix} \frac{\mathbf{F}'\mathbf{J}_T\mathbf{F}}{T} & \mathbf{0}_{K_c\times K_2^s} \\ \mathbf{0}_{K_1^s\times K_c} & \frac{\mathbf{G}_1'\mathbf{J}_T\mathbf{G}_2}{T} \end{bmatrix} \mathcal{O}_2\breve{\mathbf{L}}_{H_2}'^{-1} \\
=&\breve{\mathbf{L}}_{H_1}^{-1}\mathcal{O}_1' \begin{bmatrix} \mathbf{L}_F\mathcal{R}' & \mathbf{0}_{K_c\times K_1^s} \\ \mathbf{0}_{K_1^s\times K_c} & \mathbf{L}_{G_1}\mathbf{Z}_1 \end{bmatrix} \begin{bmatrix} \mathbf{I}_{K_c} & \mathbf{0}_{K_c\times K_2^s} \\ \mathbf{0}_{K_1^s\times K_c} & \Delta \end{bmatrix} \begin{bmatrix} \mathcal{R}\mathbf{L}_F' & \mathbf{0}_{K_c\times K_2^s} \\ \mathbf{0}_{K_2^s\times K_c} & \mathbf{Z}_2'\mathbf{L}_{G_2}' \end{bmatrix} \mathcal{O}_2\breve{\mathbf{L}}_{H_2}'^{-1} \\
=&\mathbf{V}_1\Sigma\mathbf{V}_2',
\end{aligned}$$

which confirms (C.14), which along with (C.15) proves the claims (ii) and (iii). This completes the proof of the lemma.

Proof of Theorem 2.2 From Lemmas C.2 and C.3,

$$\widehat{\mathbf{W}}_g \xrightarrow{p} \breve{\mathbf{L}}_{H_g}^{\prime -1}\breve{\mathbf{L}}_{H_g}^{-1}\mathcal{O}_g^{\prime}\begin{bmatrix} \mathbf{L}_F\mathcal{R}' & \mathbf{0}_{K_c\times K_g^s} \\ \mathbf{0}_{K_g^s\times K_c} & \mathbf{L}_{G_g}\mathbf{Z}_g \end{bmatrix} = \mathcal{O}_g^{-1}\left(\frac{\mathbf{H}_g'\mathbf{J}_T\mathbf{H}_g}{T}\right)^{-1}\begin{bmatrix} \mathbf{L}_F\mathcal{R}' & \mathbf{0}_{K_c\times K_g^s} \\ \mathbf{0}_{K_g^s\times K_c} & \mathbf{L}_{G_g}\mathbf{Z}_g \end{bmatrix}$$
$$= \mathcal{O}_g^{-1}\begin{bmatrix} (\mathbf{L}_F\mathbf{L}_F')^{-1} & \mathbf{0}_{K_c\times K_2^s} \\ \mathbf{0}_{K_g^s\times K_c} & \left(\mathbf{L}_{G_g}\mathbf{L}_{G_g}'\right)^{-1} \end{bmatrix}\begin{bmatrix} \mathbf{L}_F\mathcal{R}' & \mathbf{0}_{K_c\times K_g^s} \\ \mathbf{0}_{K_g^s\times K_c} & \mathbf{L}_{G_g}\mathbf{Z}_g \end{bmatrix}$$
$$= \mathcal{O}_g^{-1}\begin{bmatrix} \mathbf{L}_F^{\prime -1}\mathcal{R}' & \mathbf{0}_{K_c\times K_g^s} \\ \mathbf{0}_{K_g^s\times K_c} & \mathbf{L}_{G_g}^{\prime -1}\mathbf{Z}_g \end{bmatrix},$$

which in conjunction with Theorem 2.1, gives

$$\widehat{\mathbf{H}}_g\widehat{\mathbf{W}}_g = \widehat{\mathbf{H}}_g\left[\widehat{\mathbf{W}}_g^c\ \widehat{\mathbf{W}}_g^s\right] \xrightarrow{p} \mathbf{J}_T\mathbf{H}_g\mathcal{O}_g\mathcal{O}_g^{-1}\begin{bmatrix} \mathbf{L}_F^{\prime -1}\mathcal{R}' & \mathbf{0}_{K_c\times K_g^s} \\ \mathbf{0}_{K_g^s\times K_c} & \mathbf{L}_{G_g}^{\prime -1}\mathbf{Z}_g \end{bmatrix}$$
$$= \left[\mathbf{J}_T\mathbf{F}\mathbf{L}_F^{\prime -1}\mathcal{R}'\ \mathbf{J}_T\mathbf{G}_g\mathbf{L}_{G_g}^{\prime -1}\mathbf{Z}_g\right].$$

Finally, set $\mathcal{S}^c = \mathbf{L}_F^{\prime -1}\mathcal{R}'$ and $\mathcal{S}_g^s = \mathbf{L}_{G_g}^{\prime -1}\mathbf{Z}_g$. Note that $\mathcal{S}^c\mathcal{S}^{c\prime} = \mathbf{L}_F^{\prime -1}\mathcal{R}'\mathcal{R}\mathbf{L}_F^{-1} = \mathbf{L}_F^{\prime -1}\mathbf{L}_F^{-1} = \Sigma_F^{-1}$ and that $\mathcal{S}_g^s = \mathbf{L}_{G_g}^{\prime -1}\mathbf{Z}_g\mathbf{Z}_g'\mathbf{L}_{G_g}^{-1} = \Sigma_{G_g}^{-1}$. This completes the proof of the theorem.

□

Proof of Theorem 2.3 Note that

$$\|\widehat{\mathbf{R}}_g - \left(\mathbf{X}_g\Theta_g^c\right)\widehat{\mathbf{F}}_g' - \left(\mathbf{X}_g\Theta_g^s\right)\widehat{\mathbf{G}}_g'\| = \|\widehat{\mathbf{R}}_g - \mathbf{X}_g\left[\Theta_g^c\ \Theta_g^s\right]\left[\widehat{\mathbf{F}}\ \widehat{\mathbf{G}}\right]'\|$$
$$=\|\mathrm{vec}\left(\widehat{\mathbf{R}}_g\right) - \left(\left[\widehat{\mathbf{F}}\ \widehat{\mathbf{G}}\right]\otimes\mathbf{X}_g\right)\mathrm{vec}\left(\left[\Theta_g^c\ \Theta_g^s\right]\right)\|,$$

where the second equality is from the property of vectorize operator. Hence, the solution $\left(\widehat{\Theta}_g^c, \widehat{\Theta}_g^s\right)$ will be determined by

$$\mathrm{vec}\left(\left[\widehat{\Theta}_g^c\ \widehat{\Theta}_g^s\right]\right) = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y}, \tag{C.16}$$

where $\mathbf{y} = \mathrm{vec}\left(\widehat{\mathbf{R}}_g\right)$ and $\mathcal{X} = \left(\left[\widehat{\mathbf{F}}\ \widehat{\mathbf{G}}\right]\otimes\mathbf{X}_g\right)$.

Furthermore, note that

$$
\begin{aligned}
\widehat{\mathbf{R}}_g &= \mathbf{P}_g \mathbf{R}_g \mathbf{J}_T \\
&= \mathbf{X}_g \Theta_g^c \mathcal{S}_g^{c\prime -1} \mathcal{S}_g^{c\prime} \mathbf{F}' \mathbf{J}_T + \mathbf{X}_g \Theta_g^s \mathcal{S}_g^{s\prime -1} \mathcal{S}_g^{s\prime} \mathbf{G}_g' \mathbf{J}_T + \mathbf{P}_g \Gamma_g^c \mathbf{F}' \mathbf{J}_T + \mathbf{P}_g \Gamma_g^s \mathbf{G}_g' \mathbf{J}_T + \mathbf{P}_g \mathbf{E}_g \mathbf{J}_T \\
&= \mathbf{X}_g \left[\Theta_g^c \mathcal{S}_g^{c\prime -1} \ \Theta_g^s \mathcal{S}_g^{s\prime -1} \right] \left[\widehat{\mathbf{F}} \ \widehat{\mathbf{G}} \right]' + \mathcal{U}_g,
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{U}_g = {} & \mathbf{X}_g \Theta_g^c \mathcal{S}_g^{c\prime -1} \left(\mathbf{J}_T \mathbf{F} \mathcal{S}_g^c - \widehat{\mathbf{F}} \right)' + \mathbf{X}_g \Theta_g^s \mathcal{S}_g^{s\prime -1} \left(\mathbf{J}_T \mathbf{G} \mathcal{S}_g^c - \widehat{\mathbf{G}} \right)' \\
& + \mathbf{P}_g \Gamma_g^c \mathbf{F}' \mathbf{J}_T + \mathbf{P}_g \Gamma_g^s \mathbf{G}_g' \mathbf{J}_T + \mathbf{P}_g \mathbf{E}_g \mathbf{J}_T.
\end{aligned}
$$

We rewrite $\mathbf{y}$ as follows:

$$
\begin{aligned}
\mathbf{y} = \mathrm{vec}\left(\widehat{\mathbf{R}}_g \right) &= \mathrm{vec}\left(\mathbf{X}_g \left[\Theta_g^c \mathcal{S}_g^{c\prime -1} \ \Theta_g^s \mathcal{S}_g^{s\prime -1} \right] \left[\widehat{\mathbf{F}} \ \widehat{\mathbf{G}} \right]' + \mathcal{U}_g \right) \\
&= \left(\left[\widehat{\mathbf{F}} \ \widehat{\mathbf{G}} \right] \otimes \mathbf{X}_g \right) \mathrm{vec}\left(\left[\Theta_g^c \mathcal{S}_g^{c\prime -1} \ \Theta_g^s \mathcal{S}_g^{s\prime -1} \right] \right) + \mathrm{vec}\left(\mathcal{U}_g \right) \\
&= \mathbf{X} \mathrm{vec}\left(\left[\Theta_g^c \mathcal{S}_g^{c\prime -1} \ \Theta_g^s \mathcal{S}_g^{s\prime -1} \right] \right) + \mathrm{vec}\left(\mathcal{U}_g \right),
\end{aligned}
$$

which in conjunction with (C.16) gives

$$
\mathrm{vec}\left(\left[\widehat{\Theta}_g^c \ \widehat{\Theta}_g^s \right] \right) = \mathrm{vec}\left(\left[\Theta_g^c \mathcal{S}_g^{c\prime -1} \ \Theta_g^s \mathcal{S}_g^{s\prime -1} \right] \right) + \left(\mathbf{X}'\mathbf{X} \right)^{-1} \mathbf{X}' \mathrm{vec}\left(\mathcal{U}_g \right). \tag{C.17}
$$

Hence, it suffices to show that $\left(\frac{\mathbf{X}'\mathbf{X}}{N_g} \right)^{-1} \frac{\mathbf{X}' \mathrm{vec}(\mathcal{U}_g)}{N_g} \xrightarrow{p} \mathbf{0}_{L_g}$. First, note that

$$
\begin{aligned}
\frac{\mathbf{X}'\mathbf{X}}{N_g} &= \frac{\left(\left[\widehat{\mathbf{F}} \ \widehat{\mathbf{G}} \right] \otimes \mathbf{X}_g \right)' \left(\left[\widehat{\mathbf{F}} \ \widehat{\mathbf{G}} \right] \otimes \mathbf{X}_g \right)}{N_g} \\
&= \left(\left[\widehat{\mathbf{F}} \ \widehat{\mathbf{G}} \right]' \left[\widehat{\mathbf{F}} \ \widehat{\mathbf{G}} \right] \right) \otimes \frac{\mathbf{X}_g' \mathbf{X}_g}{N_g} \\
&\xrightarrow{p} \left(\left[\mathbf{J}_T \mathbf{F} \mathcal{S}_g^c \ \mathbf{J}_T \mathbf{G} \mathcal{S}_g^s \right]' \left[\mathbf{J}_T \mathbf{F} \mathcal{S}_g^c \ \mathbf{J}_T \mathbf{G} \mathcal{S}_g^s \right] \right) \otimes \mathbf{V}_{X_g},
\end{aligned} \tag{C.18}
$$

where the last limit is from Assumption 1(i) and Theorem 2.2. Next, using the property of vectorize operator, we rewrite $\mathrm{vec}\left(\mathcal{U}_g\right)$ as follows:

$$\begin{aligned}\mathrm{vec}\left(\mathcal{U}_g\right) = & \left(\left[\left(\mathbf{J}_T\mathbf{F}\mathcal{S}_g^c - \widehat{\mathbf{F}}\right)\ \left(\mathbf{J}_T\mathbf{G}\mathcal{S}_g^c - \widehat{\mathbf{G}}\right)\right] \otimes \mathbf{X}_g\right) \mathrm{vec}\left(\left[\Theta_g^c\mathcal{S}_g^{c\prime-1}\ \Theta_g^s\mathcal{S}_g^{s\prime-1}\right]\right) \\ & + \left(\mathbf{I}_T \otimes \left(\mathbf{P}_g\Gamma_g^c\right)\right)\mathrm{vec}\left(\mathbf{F}'\mathbf{J}_T\right) + \left(\mathbf{I}_T \otimes \left(\mathbf{P}_g\Gamma_g^s\right)\right)\mathrm{vec}\left(\mathbf{G}_g'\mathbf{J}_T\right) \\ & + \left(\mathbf{I}_T \otimes \left(\mathbf{P}_g\mathbf{E}_g\right)\right)\mathrm{vec}\left(\mathbf{J}_T\right),\end{aligned}$$

which in turn gives

$$\begin{aligned}& \frac{\mathbf{X}'\mathrm{vec}\left(\mathcal{U}_g\right)}{N_g} \\ = & \left(\left(\left[\widehat{\mathbf{F}}\ \widehat{\mathbf{G}}\right]'\left[\left(\mathbf{J}_T\mathbf{F}\mathcal{S}_g^c - \widehat{\mathbf{F}}\right)\ \left(\mathbf{J}_T\mathbf{G}\mathcal{S}_g^c - \widehat{\mathbf{G}}\right)\right]\right) \otimes \frac{\mathbf{X}_g'\mathbf{X}_g}{N_g}\right)\mathrm{vec}\left(\left[\Theta_g^c\mathcal{S}_g^{c\prime-1}\ \Theta_g^s\mathcal{S}_g^{s\prime-1}\right]\right) \\ & + \left(\left[\widehat{\mathbf{F}}\ \widehat{\mathbf{G}}\right]' \otimes \left(\frac{\mathbf{X}_g'\mathbf{P}_g\Gamma_g^c}{N_g}\right)\right)\mathrm{vec}\left(\mathbf{F}'\mathbf{J}_T\right) + \left(\left[\widehat{\mathbf{F}}\ \widehat{\mathbf{G}}\right]' \otimes \left(\frac{\mathbf{X}_g'\mathbf{P}_g\Gamma_g^s}{N_g}\right)\right)\mathrm{vec}\left(\mathbf{G}_g'\mathbf{J}_T\right) \\ & + \left(\left[\widehat{\mathbf{F}}\ \widehat{\mathbf{G}}\right]' \otimes \left(\frac{\mathbf{X}_g'\mathbf{P}_g\mathbf{E}_g}{N_g}\right)\right)\mathrm{vec}\left(\mathbf{J}_T\right) \xrightarrow{p} \mathbf{0}_{L_g}, \qquad \text{(C.19)}\end{aligned}$$

where the last limit is from Assumptions 1(i) and 1(ii) and Theorem 2.2.

Lastly, applying (C.18) and (C.19) into (C.17), we have that

$$\mathrm{vec}\left(\left[\widehat{\Theta}_g^c\ \widehat{\Theta}_g^s\right]\right) \xrightarrow{p} \mathrm{vec}\left(\left[\Theta_g^c\mathcal{S}_g^{c\prime-1}\ \Theta_g^s\mathcal{S}_g^{s\prime-1}\right]\right),$$

which completes the proof of the theorem. □

Proof of Theorem 2.4 We rewrite $\mathbf{R}_g$ in (2.6) as follows:

$$\begin{aligned}\mathbf{R}_g = & \mathbf{X}_g\left[\Theta_g^c\mathcal{S}^{c\prime-1}\ \Theta_g^s\mathcal{S}_g^{s\prime-1}\right]\begin{bmatrix}\mathcal{S}^{c\prime}\left(\boldsymbol{\lambda}_g^c\mathbf{1}_T' + \mathbf{F}'\right) \\ \mathcal{S}_g^{s\prime}\left(\boldsymbol{\lambda}_g^s\mathbf{1}_T' + \mathbf{G}_g'\right)\end{bmatrix} \\ & + \Gamma_g^c\left(\boldsymbol{\lambda}_g^c\mathbf{1}_T' + \mathbf{F}'\right) + \Gamma_g^s\left(\boldsymbol{\lambda}_g^s\mathbf{1}_T' + \mathbf{G}_g'\right) + \mathbf{E}_g.\end{aligned}$$

Set $\Xi = \left(\left[\widehat{\Theta}_g^c \ \widehat{\Theta}_g^s\right]' \frac{\mathbf{X}_g'\mathbf{X}_g}{N}\left[\widehat{\Theta}_g^c \ \widehat{\Theta}_g^s\right]\right)^{-1}$, which gives $\widehat{\mathbf{w}}_g^b = \frac{\mathbf{X}_g}{N}\left[\widehat{\Theta}_g^c \ \widehat{\Theta}_g^s\right]\Xi$. Note that

$$\begin{aligned}
\widehat{\mathbf{w}}_g^{b\prime}\mathbf{R}_g &= \Xi\left[\widehat{\Theta}_g^c \ \widehat{\Theta}_g^s\right]' \frac{\mathbf{X}_g'\mathbf{R}_g}{N_g} \\
&= \Xi\left[\widehat{\Theta}_g^c \ \widehat{\Theta}_g^s\right]' \frac{\mathbf{X}_g'\mathbf{X}_g}{N_g}\left[\Theta_g^c\mathcal{S}^{c\prime-1} \ \Theta_g^s\mathcal{S}_g^{s\prime-1}\right]\left[\begin{array}{c}\mathcal{S}^{c\prime}\left(\boldsymbol{\lambda}_g^c\mathbf{1}_T'+\mathbf{F}'\right)\\ \mathcal{S}_g^{s\prime}\left(\boldsymbol{\lambda}_g^s\mathbf{1}_T'+\mathbf{G}_g'\right)\end{array}\right] \\
&\quad + \Xi\left[\widehat{\Theta}_g^c \ \widehat{\Theta}_g^s\right]'\left(\frac{\mathbf{X}_g'\Gamma_g^c}{N_g}\left(\boldsymbol{\lambda}_g^c\mathbf{1}_T'+\mathbf{F}'\right)+\frac{\mathbf{X}_g'\Gamma_g^s}{N_g}\left(\boldsymbol{\lambda}_g^s\mathbf{1}_T'+\mathbf{G}_g'\right)+\frac{\mathbf{X}_g'\mathbf{E}_g}{N_g}\right) \\
&\xrightarrow{p}\left[\begin{array}{c}\mathcal{S}^{c\prime}\left(\boldsymbol{\lambda}_g^c\mathbf{1}_T'+\mathbf{F}'\right)\\ \mathcal{S}_g^{s\prime}\left(\boldsymbol{\lambda}_g^s\mathbf{1}_T'+\mathbf{G}_g'\right)\end{array}\right],
\end{aligned}$$

where the last limit is from Assumptions 1(i) and 1(ii) and Theorem 2.3. This completes the proof of the theorem. □

Proof of Theorem 2.5 From Theorem 2.4, we have that $\widehat{\mathbf{w}}_1^{c\prime}\mathbf{R}_1 \xrightarrow{p} \mathcal{S}^{c\prime}\left(\boldsymbol{\lambda}_1^c\mathbf{1}_T'+\mathbf{F}'\right)$ and $\widehat{\mathbf{w}}_2^{c\prime}\mathbf{R}_2 \xrightarrow{p} \mathcal{S}^{c\prime}\left(\boldsymbol{\lambda}_2^c\mathbf{1}_T'+\mathbf{F}'\right)$. Hence, it holds that $\widehat{\mathbf{w}}_2^{c\prime}\mathbf{R}_2-\widehat{\mathbf{w}}_1^{c\prime}\mathbf{R}_1 \xrightarrow{p} \mathcal{S}^{c\prime}\left(\boldsymbol{\lambda}_2^c-\boldsymbol{\lambda}_1^c\right)\mathbf{1}_T'$, implying that regressing the $(T\times K^c)$ matrix $\mathbf{R}_2'\widehat{\mathbf{w}}_2^c$ on a constant vector and the $(T\times K^c)$ matrix $\mathbf{R}_1'\widehat{\mathbf{w}}_1^c$ yields the intercept of $\mathcal{S}^{c\prime}\left(\boldsymbol{\lambda}_2^c-\boldsymbol{\lambda}_1^c\right)$ and the coefficient on $\mathbf{R}_1'\widehat{\mathbf{w}}_1^c$ as an identity matrix. This completes the proof of the theorem. □

D Additional Figure and Tables

Table 13: Country-specific portfolio returns. (GBP) This table presents key financial performance metrics of country-specific portfolio returns, using G7 countries (Panel A) and non-G7 countries (Panel B) as Country 2. The metrics include annualized mean returns (%), annualized standard deviations (%), t-statistics of the average returns, Sharpe ratios, maximum drawdowns, worst and best monthly returns. The mean returns, standard deviations and Sharpe ratios are annualized.

Panel A: G7 Countries as Country 2							
Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)
can	14.58	19.69	4.33	0.74	45.15	-17.81	20.86
deu	20.16	19.36	5.19	1.04	36.33	-14.21	19.44
fra	13.09	20.27	3.42	0.65	34.47	-21.58	25.29
gbr	21.68	19.62	6.06	1.11	39.80	-19.84	20.69
ita	20.01	20.04	5.19	1.00	38.77	-26.16	21.70
jpn	14.22	19.95	4.21	0.71	31.73	-21.77	18.75
Panel B: Non-G7 Countries as Country 2							
Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)
arg	20.54	19.27	3.48	1.07	20.90	-13.05	23.05
aus	16.60	20.32	5.18	0.82	27.41	-13.61	42.36
aut	21.46	18.58	4.25	1.15	35.23	-11.21	19.35
bel	27.85	20.32	6.93	1.37	25.86	-15.24	21.84
bra	16.15	17.83	2.82	0.91	24.42	-13.89	12.50
chl	18.08	21.34	4.15	0.85	48.27	-14.39	30.78
chn	14.95	21.09	3.02	0.71	53.65	-20.93	24.03
dnk	27.12	22.55	5.64	1.20	28.14	-17.30	22.58
esp	19.71	21.40	4.79	0.92	42.25	-17.28	25.25
hkg	21.13	19.50	6.80	1.08	35.87	-22.83	21.29
ind	19.89	19.92	4.57	1.00	38.96	-13.92	19.27
kor	15.45	19.76	3.74	0.78	39.12	-14.69	24.69
mex	24.27	20.77	4.36	1.17	38.35	-18.63	19.40
nor	17.39	22.16	3.65	0.78	66.72	-22.62	24.23
pol	21.11	20.93	4.11	1.01	30.62	-20.56	25.01
sgp	25.55	20.94	5.25	1.22	31.17	-15.36	29.91
swe	19.97	23.40	4.46	0.85	47.45	-20.09	36.69
tha	13.38	17.67	4.14	0.76	35.93	-25.13	24.00
tur	20.41	20.41	4.43	1.00	33.93	-18.36	25.22
twn	11.71	19.16	2.98	0.61	55.36	-13.98	17.09

Table 14: Segmentation portfolio returns. (GBP) This table presents key financial performance metrics of segmentation portfolio returns, using G7 countries (Panel A) and non-G7 countries (Panel B) as Country 2. The metrics include annualized mean returns (%), annualized standard deviations (%), t-statistics of the average returns, Sharpe ratios, maximum drawdowns, worst and best monthly returns. The mean returns, standard deviations and Sharpe ratios are annualized.

Panel A: G7 Countries as Country 2							
Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)
can	15.15	23.34	4.08	0.65	52.17	-23.76	19.88
deu	28.39	23.07	6.55	1.23	31.51	-21.21	24.09
fra	19.99	21.74	4.95	0.92	37.65	-21.40	23.89
gbr	22.37	26.93	4.56	0.83	66.27	-53.69	23.35
ita	24.21	24.44	5.10	0.99	66.49	-20.30	22.00
jpn	29.20	21.54	7.00	1.36	32.27	-15.25	31.75

Panel B: Non-G7 Countries as Country 2							
Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)
arg	14.76	22.26	2.05	0.66	50.48	-13.87	38.65
aus	23.95	28.15	4.75	0.85	38.04	-16.95	99.35
aut	14.00	21.91	2.39	0.64	45.69	-23.43	17.39
bel	15.82	22.74	3.43	0.70	65.50	-18.62	30.32
bra	13.31	21.11	2.02	0.63	30.61	-13.69	17.35
chl	16.56	20.42	3.10	0.81	34.96	-16.57	21.21
chn	17.65	25.43	3.31	0.69	40.41	-25.93	32.22
dnk	20.25	23.66	3.87	0.86	45.13	-21.85	32.60
esp	22.78	22.59	5.42	1.01	43.70	-24.64	24.63
hkg	17.40	21.05	4.00	0.83	51.11	-20.82	29.79
ind	21.34	23.41	4.89	0.91	42.27	-25.66	22.84
kor	19.70	23.06	4.07	0.85	70.84	-24.91	30.02
mex	6.92	20.82	1.31	0.33	63.38	-15.87	15.10
nor	19.15	21.26	4.12	0.90	51.90	-19.88	19.47
pol	19.76	22.26	3.67	0.89	45.84	-17.62	37.99
sgp	29.79	22.78	6.00	1.31	35.67	-14.62	27.59
swe	23.99	22.55	5.00	1.06	46.96	-30.23	22.90
tha	20.58	23.78	4.36	0.87	49.61	-30.57	21.95
tur	17.56	21.25	3.78	0.83	41.14	-21.97	30.46
twn	24.44	20.54	6.04	1.19	32.00	-17.27	17.31

Table 15: Country-specific portfolio returns. (JPY) This table presents key financial performance metrics of country-specific portfolio returns, using G7 countries (Panel A) and non-G7 countries (Panel B) as Country 2. The metrics include annualized mean returns (%), annualized standard deviations (%), t-statistics of the average returns, Sharpe ratios, maximum drawdowns, worst and best monthly returns. The mean returns, standard deviations and Sharpe ratios are annualized.

Panel A: G7 Countries as Country 2							
Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)
can	15.53	19.55	4.49	0.79	45.29	-22.81	20.24
deu	19.98	18.53	5.42	1.08	41.81	-22.35	18.66
fra	16.12	19.32	4.89	0.83	30.29	-15.30	25.99
gbr	26.37	20.03	6.93	1.32	30.21	-19.03	20.13
ita	18.99	19.97	4.74	0.95	46.21	-25.11	21.46
jpn	11.38	19.30	3.79	0.59	36.03	-18.88	18.77

Panel B: Non-G7 Countries as Country 2							
Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)
arg	18.13	20.10	3.03	0.90	24.34	-12.89	24.64
aut	19.39	19.02	3.83	1.02	37.40	-13.54	18.05
bel	28.36	20.02	7.97	1.42	24.11	-15.17	22.87
bra	11.93	20.60	1.89	0.58	26.87	-14.80	16.85
chl	19.84	21.68	4.50	0.92	39.63	-16.13	28.66
chn	17.78	21.47	3.35	0.83	38.66	-21.15	23.56
dnk	27.92	22.26	6.06	1.25	24.13	-14.49	21.54
esp	22.68	21.40	6.12	1.06	37.10	-17.93	25.16
hkg	25.63	19.99	7.22	1.28	34.98	-23.30	19.92
ind	19.19	20.79	4.20	0.92	40.48	-26.27	22.58
kor	15.94	19.39	3.90	0.82	43.83	-14.76	26.75
mex	27.72	20.02	5.86	1.38	28.22	-22.71	18.42
nor	17.32	23.11	3.43	0.75	63.47	-22.92	24.98
pol	20.96	20.47	3.97	1.02	33.82	-22.17	26.31
sgp	22.04	20.98	5.13	1.05	37.76	-25.87	25.74
swe	27.09	21.18	6.21	1.28	29.63	-19.16	37.86
tha	13.13	18.61	3.72	0.71	37.17	-24.39	23.01
tur	21.34	20.69	4.55	1.03	30.01	-18.23	25.13
twn	10.49	19.00	2.76	0.55	52.65	-14.67	16.33

Table 16: Segmentation portfolio returns. (JPY) This table presents key financial performance metrics of segmentation portfolio returns, using G7 countries (Panel A) and non-G7 countries (Panel B) as Country 2. The metrics include annualized mean returns (%), annualized standard deviations (%), t-statistics of the average returns, Sharpe ratios, maximum drawdowns, worst and best monthly returns. The mean returns, standard deviations and Sharpe ratios are annualized.

Panel A: G7 Countries as Country 2							
Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)
can	16.65	22.74	4.04	0.73	52.79	-24.09	18.70
deu	27.83	22.34	7.06	1.25	29.61	-17.19	22.93
fra	19.33	22.12	4.45	0.87	34.03	-21.94	26.11
gbr	21.71	27.78	4.26	0.78	75.15	-54.25	22.98
ita	23.75	24.68	5.05	0.96	53.75	-17.33	22.81
jpn	30.66	21.67	7.52	1.41	35.90	-14.64	26.60
Panel B: Non-G7 Countries as Country 2							
Country	Mean (%)	Standard Deviation (%)	t	Sharpe Ratio	Maximum Drawdown (%)	Worst Month (%)	Best Month (%)
arg	19.52	22.52	2.65	0.87	44.24	-12.92	38.03
aut	16.19	22.36	2.96	0.72	46.90	-20.17	15.84
bel	14.97	22.15	3.13	0.68	68.76	-20.27	26.95
bra	15.56	21.47	2.18	0.72	30.80	-12.45	17.74
chl	18.22	20.91	3.39	0.87	39.22	-19.15	22.27
chn	16.90	26.57	3.03	0.64	46.18	-35.80	32.96
dnk	19.14	23.81	3.99	0.80	40.62	-24.24	27.68
esp	14.89	23.80	3.14	0.63	56.79	-23.48	24.72
hkg	19.74	21.78	4.39	0.91	35.73	-20.24	29.31
ind	23.51	21.85	5.46	1.08	44.29	-23.94	25.83
kor	19.07	22.23	4.11	0.86	69.98	-23.28	25.90
mex	10.89	20.50	1.99	0.53	52.42	-13.09	15.85
nor	22.32	22.08	5.28	1.01	41.86	-14.91	26.08
pol	20.73	22.10	3.75	0.94	35.40	-21.12	37.12
sgp	32.16	22.51	7.56	1.43	34.44	-14.21	28.31
swe	25.73	22.85	5.49	1.13	37.05	-26.91	22.78
tha	18.14	24.29	3.56	0.75	60.62	-26.30	20.11
tur	16.33	20.29	3.70	0.80	44.23	-12.93	26.41
twn	23.46	21.46	5.47	1.09	38.66	-18.36	19.44

E Bootstrap Method for the Number of Common Factors

For the reader's convenience, we outline the bootstrap inference procedure proposed by Gonçalves et al. (2025) to determine the number of common factors in our two-country model.

Model and Estimation

For expositional simplicity, we slightly abuse some notations; they do not exactly match those used in the main text. We consider two sets of time-series returns, $\mathbf{y}_{1t}$ $(N_1 \times 1)$ and $\mathbf{y}_{2t}$ $(N_2 \times 1)$, for $t = 1, \ldots, T$, modeled as

$$\mathbf{y}_{1t} = \mathbf{B}_1\mathbf{f}_t + \mathbf{D}_1\mathbf{g}_{1t} + \mathbf{e}_{1t}, \quad \mathbf{y}_{2t} = \mathbf{B}_2\mathbf{f}_t + \mathbf{D}_2\mathbf{g}_{2t} + \mathbf{e}_{2t},$$

where $\mathbf{f}_t$ denotes the K^c common factors, $\mathbf{g}_{1t}$ and $\mathbf{g}_{2t}$ represent country-specific factors of dimensions K_1^s and K_2^s, respectively, and $\mathbf{e}_{1t}$ and $\mathbf{e}_{2t}$ are errors. The total number of factors for each group is therefore $K_1 = K^c + K_1^s$ and $K_2 = K^c + K_2^s$.

Regarding the data $\widehat{\mathbf{R}}_g$ in (2.8), we define $\mathbf{Y}_g = [\mathbf{y}_{g1},\ \mathbf{y}_{g2},\ \cdots,\ \mathbf{y}_{gT}]' = \widehat{\mathbf{R}}_g'$, which is a $(T \times N_g)$ matrix for $g = 1, 2$. The factors for country g are estimated using principal component analysis (PCA) applied to the data matrix $\mathbf{Y}_g$. Specifically, the eigenvectors of the covariance matrix $\mathbf{Y}_g\mathbf{Y}_g'/N_g$ provide the estimated factor space $\widehat{\mathbf{H}}_g$.

Next, the canonical correlations $\widehat{\rho}_k$ for $k = 1, \ldots, K$ are computed as the square roots of the eigenvalues of the matrix $\left(\frac{\widehat{\mathbf{H}}_1'\widehat{\mathbf{H}}_1}{T}\right)^{-1}\left(\frac{\widehat{\mathbf{H}}_1'\widehat{\mathbf{H}}_2}{T}\right)\left(\frac{\widehat{\mathbf{H}}_2'\widehat{\mathbf{H}}_2}{T}\right)^{-1}\left(\frac{\widehat{\mathbf{H}}_2'\widehat{\mathbf{H}}_1}{T}\right)$. Using the canonical correlation analysis, the common factors $\widehat{\mathbf{f}}_t$ and the residuals $\widehat{\mathbf{e}}_{1t}$ and $\widehat{\mathbf{e}}_{2t}$ are obtained, along with the group-specific factors $\widehat{\mathbf{g}}_{1t}$ and $\widehat{\mathbf{g}}_{2t}$.

Test Statistic and Bootstrap Inference

We set the null hypothesis as $H_0 : K^c = K_0^c$. The test statistic is defined as

$$\widehat{\tau} = \sum_{k=1}^{K_0^c} \widehat{\rho}_k - K_0^c.$$

The number of common factors is determined through the following procedure:

1. As previously described, compute $\widehat{\tau}$ and fit the model

$$\mathbf{y}_{gt} = \widehat{\mathbf{B}}_g \widehat{\mathbf{f}}_t + \widehat{\mathbf{D}}_g \widehat{\mathbf{g}}_{gt} + \widehat{\mathbf{e}}_{gt}, \quad g = 1, 2,$$

   where $\widehat{\mathbf{e}}_t = [\widehat{\mathbf{e}}_{1t}', \widehat{\mathbf{e}}_{2t}']'$ denotes the stacked residuals.

2. For $b = 1, \ldots, B$:

   - Generate bootstrap residuals $\widehat{\mathbf{e}}_t^{*(b)} = [\widehat{\mathbf{e}}_{1t}^{*(b)\prime}, \widehat{\mathbf{e}}_{2t}^{*(b)\prime}]' = \widehat{\mathbf{e}}_t \odot \boldsymbol{\eta}_t^{(b)}$, where each element of $\boldsymbol{\eta}_t^{(b)}$ is independently drawn from $\mathcal{N}(0, 1)$.
   - Construct the bootstrap data:

$$\mathbf{y}_{gt}^{*(b)} = \widehat{\mathbf{B}}_g \widehat{\mathbf{f}}_t + \widehat{\mathbf{D}}_g \widehat{\mathbf{g}}_{gt} + \widehat{\mathbf{e}}_{gt}^{*(b)}, \quad g = 1, 2.$$

   - Recompute the test statistic $\widehat{\tau}^{*(b)}$ using the bootstrap data.

3. Calculate the bootstrap p-value as $p = \frac{1}{B} \sum_{b=1}^{B} \mathbf{1}\left(\widehat{\tau}^{*(b)} < \widehat{\tau}\right)$.

4. If $p < 0.05$, reject H_0 and increment K_0^c by one; otherwise, stop and accept the current K_0^c as the estimated number of common factors.

The test is performed annually, once every 12 months, using the most recent 60-month rolling window of data.

References

Andreou, E., P. Gagliardini, E. Ghysels, and M. Rubin (2019). Inference in group factor models with an application to mixed-frequency data. *Econometrica 87*(4), 1267–1305.

Andreou, E., P. Gagliardini, E. Ghysels, and M. Rubin (2020). Mixed-frequency macro–finance factor models: Theory and applications. *Journal of Financial Econometrics 18*(3), 585–628.

Bae, J. W., R. Elkamhi, and M. Simutin (2019). The best of both worlds: Accessing emerging economies via developed markets. *The Journal of Finance 74*(5), 2579–2617.

Bekaert, G. and C. R. Harvey (1995). Time-varying world market integration. *the Journal of Finance 50*(2), 403–444.

Bekaert, G., C. R. Harvey, and C. Lundblad (2005). Does financial liberalization spur growth? *Journal of Financial economics 77*(1), 3–55.

Berger, D., K. Pukthuanthong, and J. J. Yang (2011). International diversification with frontier markets. *Journal of Financial Economics 101*(1), 227–242.

Brown, S. J. and M. I. Weinstein (1983). A new approach to testing asset pricing models: The bilinear paradigm. *The Journal of Finance 38*(3), 711–743.

Carrieri, F., V. Errunza, and K. Hogan (2007). Characterizing world market integration through time. *Journal of Financial and Quantitative Analysis 42*(4), 915–940.

Chaieb, I., H. Langlois, and O. Scaillet (2021). Factors and risk premia in individual international stock returns. *Journal of Financial Economics 141*(2), 669–692.

Cho, D. C., C. S. Eun, and L. W. Senbet (1986). International arbitrage pricing theory: An empirical investigation. *The Journal of Finance 41*(2), 313–329.

Choi, S. H. and D. Kim (2023). Large volatility matrix analysis using global and national factor models. *Journal of Econometrics 235*(2), 1917–1933.

Connor, G. and R. A. Korajczyk (1986). Performance measurement with the arbitrage pricing theory: A new framework for analysis. *Journal of Financial Economics 15*(3), 373–394.

Connor, G. and R. A. Korajczyk (1988). Risk and return in an equilibrium apt: Application of a new test methodology. *Journal of Financial Economics 21*(2), 255–289.

Dumas, B. and B. Solnik (1995). The world price of foreign exchange risk. *The journal of finance 50*(2), 445–479.

Errunza, V., K. Hogan, and M.-W. Hung (1999). Can the gains from international diversification be achieved without trading abroad? *The Journal of Finance 54*(6), 2075–2107.

Eun, C. S., S. Kim, F. Wei, and T. Zhang (2017). Global diversification with local stocks: A road less traveled. *Georgia Tech Scheller College of Business Research Paper* (17-26).

Fama, E. F. and J. D. MacBeth (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy 81*(3), 607–636.

Fan, J., Y. Liao, and W. Wang (2016). Projected principal component analysis in factor models. *Annals of Statistics 44*, 219–254.

Ferson, W. E. and C. R. Harvey (1999). Conditioning variables and the cross section of stock returns. *Journal of Finance 54*(4), 1325–1360.

Fortin, A.-P., P. Gagliardini, and O. Scaillet (2025). Optimal maximin gmm tests for sphericity in latent factor analysis of short panels. *Swiss Finance Institute Research Paper* (25-27).

Freyberger, J., B. Höppner, A. Neuhierl, and M. Weber (2024). Missing data in asset pricing panels. *The Review of Financial Studies*, hhae003.

Freyberger, J., A. Neuhierl, and M. Weber (2020). Dissecting characteristics nonparametrically. *The Review of Financial Studies 33*(5), 2326–2377.

Ghysels, E. (1998). On stable factor structures in the pricing of risk: do time-varying betas help or hurt? *Journal of Finance 53*(2), 549–573.

Gonçalves, S., J. Koh, and B. Perron (2025). Bootstrap inference for group factor models. *Journal of Financial Econometrics 23*(2), nbae020.

Grubel, H. G. (1968). Internationally diversified portfolios: welfare gains and capital flows. *The American economic review 58*(5), 1299–1314.

Harvey, C. R. (1991). The world price of covariance risk. *The Journal of Finance 46*(1), 111–157.

Heston, S. L. and K. Rouwenhorst (1994). Does industrial structure explain the benefits of international diversification? *Journal of Financial Economics 36*(1), 3–27.

Ince, O. S. and R. B. Porter (2006). Individual equity return data from thomson datastream: Handle with care! *Journal of Financial Research 29*(4), 463–479.

Jensen, T. I., B. Kelly, and L. H. Pedersen (2023). Is there a replication crisis in finance? *The Journal of Finance 78*(5), 2465–2518.

Karolyi, G. A. and R. M. Stulz (2003). Are financial assets priced locally or globally? *Handbook of the Economics of Finance 1*, 975–1020.

Kelly, B., S. Pruitt, and Y. Su (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics 134*, 501–534.

Kim, S. and R. A. Korajczyk (2024). Large sample estimators of the stochastic discount factor. *Journal of Financial Econometrics 22*(5), 1672–1713.

Kim, S., R. A. Korajczyk, and A. Neuhierl (2021). Arbitrage Portfolios. *Review of Financial Studies 34*(6), 2813–2856.

Kim, S. and G. Skoulakis (2018). Ex-post risk premia estimation and asset pricing tests using large cross sections: The regression-calibration approach. *Journal of Econometrics 204*(2), 159–188.

Korajczyk, R. A. and C. J. Viallet (1989). An empirical investigation of international asset pricing. *The Review of Financial Studies 2*(4), 553–585.

Korajczyk, R. A. and C. J. Viallet (1992). Equity risk premia and the pricing of foreign exchange risk. *Journal of International Economics 33*(3–4), 199–219.

Linton, O. B., H. Tang, and J. Wu (2025). A large confirmatory dynamic factor model for stock market returns in different time zones. *Journal of Econometrics 249*, 105971.

Markowitz, H. (1952). Modern portfolio theory. *Journal of Finance 7*(11), 77–91.

Patton, A. J. and B. M. Weller (2022). Risk price variation: The missing half of empirical asset pricing. *The Review of Financial Studies 35*(11), 5127–5184.

Pukthuanthong, K. and R. Roll (2009). Global market integration: An alternative measure and its application. *Journal of Financial Economics 94*(2), 214–232.

Quinn, D. P. and H.-J. Voth (2008). A century of global equity market correlations. *American Economic Review 98*(2), 535–540.

Raponi, V., C. Robotti, and P. Zaffaroni (2020). Testing beta-pricing models using large cross-sections. *The Review of Financial Studies 33*(6), 2796–2842.

Roll, R. and S. A. Ross (1980). An empirical investigation of the arbitrage pricing theory. *Journal of Finance 35*(5), 1073–1103.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory 13*(3), 341–360.

Sandulescu, M. and P. Schneider (2021). International arbitrage premia. Research Paper Series 21-14, Swiss Finance Institute.

Sandulescu, M., F. Trojani, and A. Vedolin (2021). Model-free international stochastic discount factors. *Journal of Finance 76*(2), 935–976.

Shanken, J. (1992). On the estimation of beta-pricing models. *The review of financial studies 5*(1), 1–33.

Sinquefield, R. A. (1996). Where are the gains from international diversification? *Financial analysts journal 52*(1), 8–14.

Solnik, B. H. (1974). Why not diversify internationally rather than domestically? *Financial analysts journal 30*(4), 48–54.

Verdelhan, A. (2018). The share of systematic variation in bilateral exchange rates. *The Journal of Finance 73*(1), 375–418.

Zaffaroni, P. (2025). Factor models for conditional asset pricing. *Journal of Political Economy*. Forthcoming.