ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS
Bachelor Thesis Economics & Business
Specialization: Financial Economics

Forecasting the Long-Term Equity Premium for Selected Emerging Markets: A Cross-Sectional Global Factor Model Approach

Author:	Alexandros Rizzi
Student number:	599476
Thesis supervisor:	Dr. Laurens Swinkels
Second reader:	X Ma
Finish date:	10th July 2024

Abstract

A cross-sectional global factor model (CS-GFM) models long-term country equity premiums as the expected global market equity premium plus additional factor premiums commensurate with each market's factor exposure to risks proxied by state variables. CS-GFM long-term equity premium forecasts are statistically significantly more accurate compared to the historical average method and selected benchmark models, for nine selected emerging markets. Long-term asset allocation strategies based on CS-GFM forecasts do not produce significant utility gains to investors compared to the historical average method and selected benchmark models.

Keywords: global factor model; equity premium prediction; forecasting

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

A well-specified factor model that uses theoretically sound predictor variables which show success in predicting long-term equity risk premiums out-of-sample and produces significant utility gains for investors can greatly aid both individual and institutional investors in making informed strategic asset allocation decisions to potentially boost their portfolio performance. Emerging markets, despite recent underperformance compared to developed markets (MSCI, 2024) could potentially offer opportunities due to their different risk-return profiles. Consequently, they could be viable options for both individual and institutional investors with respect to long-term strategic asset allocation.

This paper uses the cross-sectional global factor model (CS- GFM) put forward by Sakkas and Tessaromatis (2022) to forecast the long-term equity risk premiums at the country index level of nine selected emerging markets. A country's future equity premium is modelled as the expected global market equity premium plus additional global equity factor premiums commensurate with each market's factor exposure to risks proxied by state variables like the valuation ratios dividend yield (DY) and price-earnings ratio (PE), macroeconomic variables like short-term interest rate (STIR) and term spread (slope of the term structure of the interest rates) (TS), and market volatility (MVOL).

This thesis contributes to the broad equity premium literature in a few ways, making it scientifically relevant. Firstly, it provides empirical evidence on the effectiveness of a long-term cross-sectional global factor model for long-term out-of-sample equity premium prediction in nine emerging markets: Greece, South Africa, Hungary, India, Thailand, Philippines, Malaysia, Taiwan, and Pakistan[1], selected based on data availability. These are countries were chosen due to their sufficient comprehensive data over a large enough time period where all variables are simultaneously available for the application of the CS-GFM. This will enhance the theoretical understanding on cross-sectional country return predictability for the selected emerging markets, which is scientifically relevant. It will also add to the limited out-of-sample predictability evidence for state variables dividend yield (DY), price-earnings ratio (PE), short-term interest rate (STIR), term spread (TS) and market volatility (MVOL) as predictors of long-term country index returns for emerging markets.

Second, the economic utility gains of investors using the CS-GFM forecasts for long-term asset allocation decisions in the selected emerging markets is assessed. A mean-variance investor with a five-

[1] Pakistan is not technically classified as an emerging market, but as a frontier market. However, due to the available required data, and the fact that it used to be classified as an emerging market until 2021(MSCI, 2021) means it is suitable for inclusion in the analysis. The motivation for the choice of the selected emerging markets is described in Chapter 3.

year investment horizon is assumed, who decides the optimal asset mix at the start of the investment horizon and then rebalances the asset allocation monthly (to ensure the asset mix remains aligned with the five-year strategy) until the end of the five-year investment horizon. The utility gains are assessed against the utility gains from basing the strategy on the historical average equity premium forecast. This will generate practical insights that investors can implement to potentially boost portfolio performance compared to strategies based on the historical average or other benchmark models.

Accurate equity premium prediction models help investors allocate their capital more efficiently by directing resources towards the most promising opportunities, a step closer towards maximization of portfolio returns while mitigating risks. Effective risk management reduces the likelihood of significant financial losses, which is particularly beneficial for institutional investors such as pension funds, which rely on stable and predictable returns to meet their long-term obligations and are more likely to take long-term asset allocation decisions compared to short-term ones. This ensures social relevance of the research. Given the volatile nature of emerging markets, predictive models such as the CS-GFM are invaluable.

Finally, the performance of strategic asset allocation strategies based on prediction models commonly used by academics and practitioners is compared. These include panel predictive and time series models based on the individual state variables of the CS-GFM, and combinations of them.

Variables that are significant in-sample predictors of the equity premium match the rest of the literature on emerging market return prediction, apart from the surprising insignificance of the price-earnings (PE) ratio when it is the sole predictor. CS-GFM out-of-sample predictions of 5-year equity premiums statistically significantly reduce the mean-squared error (MSE) for all nine countries, compared to forecasts based on the historical average equity premium, and the benchmark models, possibly through the model's ability to combine average factor premiums and country characteristics. The persistence of the state variables is not as closely linked to long-term equity premium prediction performance as with developed markets. Interest rate variables such as the short-term interest rate (STIR) outperform valuation ratios (DY and PE) as predictors of country returns, consistent with the rest of the emerging market country return prediction literature. There is insufficient evidence to suggest the economic utility gains for investors from long-term asset allocation strategies based on the CS-GFM forecasts for the selected emerging markets are higher compared to strategies using forecasts based on the historical average equity premium, and the benchmark models. The same is true for the benchmark models compared to the historical average method. This is likely due to the small magnitude of the factor premiums, meaning that accurate forecasts from the CS-GFM do not necessarily translate into substantial utility gains for investors.

CHAPTER 2 Theoretical Framework

2.1 Global Prediction

The asset pricing literature on cross-sectional patterns for individual stock return prediction is extensive. In comparison, the empirical work on country-based equity markets is much smaller. However, as confirmed by Zaremba (2019), many of the cross-sectional patterns found for individual stocks such as value, momentum, or seasonality, have their parallels at the intermarket level and can be used for country allocation. Angelidis and Tessaromatis (2018) show that country-based and individual stock-based global investments are similar in terms of performance and risk characteristics. Below is an overview of empirical patterns found in the cross-section of country equity returns.

The tendency of assets with high past returns to continue outperforming, and vice versa, is the momentum effect. It is robust across different markets and time periods, with strong predictive power for country equity index returns (Balvers and Wu, 2006; Bhojraj and Swaminathan, 2006; Baltussen et al., 2021). The size effect is the tendency of smaller equity markets to outperform larger ones. First observed at the country level by Keppler and Traub (1993), many studies followed such as Zaremba and Umutlu (2018) demonstrating the size effect in a large international sample, and Li and Pritamani (2015) showing its link to return predictability for emerging and frontier markets. The value effect refers to stocks with low valuations ratios outperforming stocks with high valuation ratios; usually it is measured by Price-Earnings ratio (P/E), Price-book ratio (P/B) and dividend yield (DY). This was documented at the country level (Angelidis and Tessaromatis, 2018; Faber, 2012). Cross-sectional seasonality (stocks with high same-month average returns in the past tend to outperform those with low same-month returns) has been confirmed in international markets, for instance in Heston and Sadka (2010) and Keloharju et al. (2016) with country equity indices. Long-run reversal is the tendency for stocks with good long-term performance to eventually underperform, and vice versa. First documented for country equity indices by Kasa (1992) and Richards (1995, 1997), it was confirmed for a large sample by Spierdijk et al. (2012). Price risk has been explored extensively. For example, Frazzini and Pedersen (2014) show that low-beta indices outperform high-beta indices, and Bali and Cakici (2010) show a positive link between idiosyncratic volatility and returns. Jordan et al. (2014) also found return predictability for the past return variance. Non-price risks such as credit risk (Erb et al., 1995) and political risk (Dimic et al., 2015) have also been explored. Other predictors are also showing good predictive power, in particular interest rates. Hjalmarsson (2010) found decent predictive power for developed country index returns with the short-term interest rate and the term spread.

2.2 Out-of-sample Versus In-sample Prediction

While it is evident many predictor variables have in-sample success, the distinction between in-sample and out-of-sample success for return prediction is imperative. Goyal & Welch (2008) argue that most predictor variables have limited out-of-sample predictive power for individual US stock returns, and that the historical average excess stock return (often used as a benchmark for prediction models to outperform) forecasts future excess stock returns better than regressions of excess returns on predictor variables. Out-of-sample tests admittedly have worse power of statistical tests (Inoue and Kilian, 2005) and less precise estimates compared to in-sample tests, as they generally use much smaller samples. It may indeed be possible that the absence of strong out-of-sample predictability in stock returns is exclusively due to short evaluation samples (Paye and Timmerman, 2006). Also, good in-sample but bad out-of-sample performance indicates problems like overfitting or data mining bias and is not practically useful for an investor. The biggest issue for out-of-sample forecasting is structural parameter instability; inconsistent parameters across periods lead to inconsistent out-of-sample performance (Goyal and Welch, 2008). For instance, Paye and Timmerman (2006) found that for linear models of developed country equity index excess returns, there is evidence of parameter instability of the estimated regression parameters in most countries; the hypothesis of a constant regression coefficient is almost always rejected.

Another issue when relying on in-sample evidence is persistence of the predictor variables; this is the degree to which a variable's past values indicate its future values. Ang and Bekaert (2007) point out that persistent variables may lead to spurious regression results. In the general literature, most predictor variables have been highly persistent, and when corrected produce even weaker evidence of predictability. Cross-sectional predictive regressions are less susceptible to the issues caused by highly persistent predictor variables than time series; the focus on cross-sectional variability helps isolate relationships without the interference of temporal dependencies, which could be induced by the high persistence of predictors (Karolyi & Van Nieuwerburgh, 2020).

There are further advantages of a cross-sectional prediction model compared to time series predictive models. Firstly, observed characteristics from predictive regressions are better proxies than estimated beta factor exposures from time series regressions. This is shown for individual stocks by Chordia, Goyal, and Shanken (2019) and Nazaire, Pacurar, and Sy (2020). Kelly, Pruitt and Su (2019) argue that characteristics are proxies for time-varying betas on common risk factors. Further, Lewellen (2015), Green et al. (2017), and Bessembinder et al. (2019) found evidence that cross-sectional regressions provide quite reliable estimates of expected stock returns. Finally, Fama and French (2019) showed that time series models using cross-sectional factors better describe equity returns in the US market compared to models using time series factors. They argue that when stacked across t, cross-sectional regressions become a time series model which combines observable time-varying stock or county characteristics with estimated common risk factor premiums. This differs from traditional time series factor models such as Fama and French (2015) because they use prespecified factors with factor loadings assumed to be constant over time.

It follows that a global factor model using the cross-sectional approach may indeed be more successful at predicting individual stock returns. Zaremba (2019) states that many of the cross-sectional patterns found for individual stocks have their parallels at the intermarket level and could be potentially used for country allocation; therefore, a global factor model can take advantage of the cross-sectional approach to predict returns at the country index level, even using variables also proven to be good predictors for individual stock returns.

2.3 Cross-Sectional Global Factor Model

Sakkas and Tessaromatis (2022) introduce a cross-sectional global factor model (CS-GFM) designed for long-term asset allocation. It predicts country returns by using the cross-sectional approach to predict individual stock returns of Haugen and Baker (1996), Lewellen (2015), Green et al. (2017), and Bessembinder et al. (2019). The model posits that a country's equity premium equals the equally weighted global equity market premium, plus factor premiums which compensate investors for risks related to their hedging requirements (Sakkas and Tessaromatis, 2022). The factors included, labelled as state variables, are based on their extensive use and performance in the literature on individual return prediction. These are dividend yield (DY), price-earnings ratio (PE), short-term interest rate (STIR), term spread (TS) and market volatility (MVOL).

The CS-GFM uses cross-sectional regressions stacked over time, as detailed in Section 2.2. This combines observable time-varying country characteristics with estimated common risk factor premiums. Specifically, the cross-sectional specification captures heterogeneity in country-specific characteristics affecting equity premiums. Meanwhile, the time-varying effects of global common shocks are incorporated through coefficients that are expanding window averages of cross-sectional regression estimates and are therefore updated monthly. This assumes that historical averages of factor premiums provide information about future country returns. If factor premiums are compensations for risk, they should persist in the future, thus offering predictive value as in Merton's (1973) Intertemporal Capital Asset Pricing Model (ICAPM), where priced factors represent compensation for exposure to state variables.

For 12 developed markets, long-term asset allocation strategies based on CS-GFM forecasts, especially using simple valuation indicators like the DY or PE ratios, outperform significantly out-of-sample strategies using the historical mean as an input (Sakkas & Tessaromatis, 2022). Time series prediction models combined with multiple regression models, diffusion indices, or a combination of single variable-based forecasts significantly underperform the CS-GFM when used for long-term asset allocation, reconciling with the work of Karolyi and Van Nieuwerburgh (2020) mentioned earlier.

Given its effectiveness, it follows that the CS-GFM is suitable for use in the long-term asset allocation decisions of investors when allocating portfolio weights between country indices and the risk-free rate in developed markets. Emerging markets are characterized by lower market efficiency and higher business cycle variability (Hollstein et al., 2020), which should lead to higher return predictability compared to developed economies with more efficient markets. Therefore, the question arises of whether the CS-GFM will offer significant predictive power and practical utility for investors in emerging markets, as well as developed markets.

2.4 State Variables

While there are many predictors that are effective in-sample mentioned earlier, that can also be effective out-of-sample, the focus of this paper is to apply the CS-GFM of Sakkas and Tessaromatis (2022) to emerging markets. Therefore, the five state variables they used to test the model for developed markets will also be used here.

PE and DY ratios are part of the value effect, as mentioned earlier. Both Kim (2012) and Zaremba (2016) show that the value effect is stronger among the emerging markets rather than in developed countries, boding well for the CS-GFM's applicability. However, the out-of-sample evidence thus far is not promising. Hollstein et al. (2020) found generally poor out-of-sample predictability for DY, but it is better for emerging compared to developed markets. The same is true for PE. Bahrami et al. (2018) found generally poor performance of DY and PE in and out-of-sample in their study focusing only advanced emerging market countries (more developed and have a lower prevalence of market frictions than secondary emerging and frontier markets). There are some exceptions of course, like the finding of positive out-of-sample predictability for DY in Brazil and Thailand, and PE in Malaysia and Poland. Charles et al. (2017) also found poor out-of-sample performance of financial ratios, which include DY and PE. Hjalmarsson (2010) found weak out-of-sample predictability for DY and PE for emerging markets.

Interest rate variables have more convincing evidence. Charles et al. (2017) found that STIR performed well out-of-sample especially for Asian markets like South Korea and Malaysia. It performs better for developed markets, however. Hollstein et al. (2020) found the opposite; STIR has generally a larger and more often significantly positive out-of-sample R^2 than in developed markets; they reported moderate out-of-sample predictability. Bahrami et al. (2018) found that interest rates are consistent, strong predictors out-of-sample; TS also had strong out-of-sample predictability in longer horizons for specific countries. Hollstein et al. (2020) also found moderate out-of-sample performance of TS, again stronger for emerging markets compared to developed ones.

Market volatility (MVOL) is calculated as the standard deviation, each month, of the daily stock returns in that month. Hollstein et al. (2020) found a consistent, moderate to strong out-of-sample predictability; stock excess return volatility had positive out-of-sample R^2 values and significant for a substantial proportion of the emerging markets in the analysis. As for all the state variables, Sakkas and Tessaromatis (2022) draw this variable from the individual stock return prediction literature, most notably from the work of Bakshi and Kapadia (2003) who found a negative market volatility risk premium. Furthermore, Ang et al. (2006) show that market volatility is a significant factor for returns.

The state variables are both theoretically and empirically motivated by previous literature, either for country returns or individual stock returns, also for emerging markets; therefore, they are suitable to use in a CS-GFM for country equity premium prediction in emerging markets.

2.5 Emerging Markets

The state variables are both theoretically and empirically motivated by previous literature, also for emerging markets, for either country returns or individual stock returns; therefore, they are suitable to use in a CS-GFM for country equity premium prediction in emerging markets. However, the different risk-return characteristics for emerging markets may affect the applicability of the CS-GFM compared to developed markets. Emerging markets are highly volatile (Narayan et al., 2014). They are characterized by both lower market efficiency and generally higher business cycle variability (Hollstein et al., 2020). Harvey (1995) also argues that emerging markets are segmented with a high degree of return predictability. Hollstein et al. (2020) found the same in their comprehensive analysis of 81 countries over a period of up to 145 years with various predictor variables and forecast specifications; they show that out-of-sample predictability is generally better (through larger and more often significantly positive out-of-sample R^2 values) the less developed a country's capital market is.

This points towards higher out-of-sample predictability for the state variables and their combinations for emerging markets compared to developed markets, possibly leading to more significant factor premiums for the state variables compared to developed economies. Note that the analysis by Hollstein et al. (2020) included all five state variables in the CS-GFM and confirmed stronger out-of-sample predictability for emerging compared to developed markets for all of them.

Variables such as PE and DY thus far did not perform well for emerging markets, especially compared to variables like interest rates. However, they also have lackluster out-of-sample evidence for developed economies (Hjalmarsson, 2010; Jordan et al., 2014), and despite this still outperform significantly out-of-sample strategies using the historical mean as an input, when used in the CS-GFM framework. This means the cross-sectional estimation method of the CS-GFM that combines average factor premiums and country characteristics is effective for return prediction, and the same outperformance may be seen for emerging markets, despite the unconvincing out-of-sample evidence.

Donadelli and Paradiso (2014) found that the level of integration across emerging equity markets is low at the country level. This points to a higher heterogeneity in country characteristics compared to developed markets, which will be captured by the CS-GFM. Emerging markets are also more likely to undergo frequent economic, political, and regulatory changes, leading to a more volatile environment (Bekaert and Harvey, 1997). This volatility may cause unstable relationships between predictors and equity premiums over time, making parameter instability an even bigger issue. The CS-GFM is equipped to handle this due to its time-varying slopes (more details to follow in Chapter 3), meaning coefficients are not rigid as in time series models. It follows that the CS-GFM can effectively capture both persistent country-specific characteristics and the frequent temporal shocks that characterize emerging markets, which can prove robust for country equity premium prediction to guide long-term asset allocation.

It is also worth mentioning the prediction horizon. Overfitting and parameter instability reduce out-of-sample performance for short-term predictions (ranging from one month to one year) (Goyal & Welch, 2008; Paye & Timmerman, 2006), which constitute most of the return prediction literature. In contrast, long-term prediction models like the CS-GFM benefit from incorporating persistent factors and mitigate these issues by averaging out short-term market fluctuations, leading to more stable predictions. Drawing from the individual stock return prediction literature, there is evidence to suggest that long-horizon premiums might be more predictable than short-horizon ones. The equity premium is of a mean-reverting nature, meaning periods of high or low premiums eventually revert to the long-term average, enhancing predictability over extended periods (Fama & French, 2002). Long-term returns are driven by fundamental factors such as earnings growth and dividends, which are less volatile and more stable compared to short-term market noise (Arnott & Bernstein, 2002). Ibbotson and Chen (2003) also find similar drivers like real economic growth, dividends, and reinvestment returns, which are relatively stable and predictable over long horizons. Bogle and Nolan (2015) emphasize that short-term valuation anomalies are corrected in the long term, increasing the predictive power of valuation-based models. The persistence of these fundamental factors over long periods further supports higher predictability of long-horizon equity premiums over short-term ones, and points to the effectiveness of the CS-GFM as an effective long-term return prediction model.

2.6 Research Question and Hypotheses

Building on the theoretical base and empirical insights provided by previous studies, as well as the speculated reasons for the applicability of the CS-GFM to emerging markets, this thesis aims to address the following research questions:

1. How effective is the Cross-Sectional Global Factor Model (CS-GFM) in predicting long-term equity premiums for the selected emerging markets compared to the historical average equity premium, and benchmark models?

2. How effective is the Cross-Sectional Global Factor Model (CS-GFM) in producing significant utility gains for investors pursuing a long-term asset allocation strategy based on its long-term equity premium forecasts for the selected emerging markets, compared to using the historical average equity premium, and benchmark models?

Given this, the following hypotheses arise:

1. CS-GFM out-of-sample predictions of 5-year equity premiums will statistically significantly reduce the mean-squared error (MSE) compared to forecasts based on the historical average equity premium, through the model's ability to combine average factor premiums and country characteristics. This will be tested with $H_0: R^2_{OOS} \leq 0,\ H_1: R^2_{OOS} > 0$, using the 5% significance level, where the comparison model is the historical average equity premium.
2. The economic utility gains for investors from long-term asset allocation strategies based on the CS-GFM forecasts for the selected emerging markets will be significantly higher compared to strategies based on forecasts based on the historical average equity premium. This is tested with $H_0: \Delta CER_i = 0, H_1: \Delta CER_i > 0$, using the 5% significance level, where the comparison model is the historical average equity premium.

More details in the statistical tests through which the hypotheses will be tested are provided in Section 4.4. By addressing these hypotheses, this research will fill gaps in the literature on long-term country equity premium prediction for emerging markets and provide valuable insights for both academic researchers and investors in practice.

Chapter 3 Data

Emerging markets are selected based on the MSCI (2022) classification. The countries are Greece, South Africa, Hungary, India, Thailand, Philippines, Malaysia, Taiwan, and Pakistan. These nine countries were chosen because of their sufficient availability of comprehensive, reliable data on equity market return, DY, PE, 10-year government bond yields and 3-month treasury bill yields at the country index level, over a large enough time period where all variables are simultaneously available for each month, with few missing values. Only the selected countries fit this criteria, and that is the reason for their selection. Pakistan is a frontier market, but due to its availability of the mentioned variables in the right time period, it was included for a more comprehensive analysis from including an additional country. It was classified as an emerging market until 2021 (MSCI, 2021) and shares similarities with the other emerging markets, making its inclusion suitable. Descriptive statistics per country are to be found in Appendix A.

Data on country equity indices is sourced from Refinitiv Datastream. STIR is proxied by the 3-month treasury bill yields and the long-term bond yield is proxied by 10-year government bond yields; these are obtained from Finaeon Global Financial Data. TS is computed as 10-Year bond yield minus the 3-month treasury yield. MVOL is computed at a monthly frequency as the standard deviation of the available daily returns over the month. All other variables are at a montly frequency. Outlined in Appendix B are the exact variable sources, formulations and codes where applicable. All state variables (characteristics) are standardized by subtracting the characteristic's cross-sectional average and dividing by its cross-sectional standard deviation; standardized characteristics are used in all analyses.

The sample period (where all variables are available for all countries) is Feb-1999 to May-2022, leading to a time series of data with n = 280 per country; This means the length of the time series is homogeneous across countries; results are more comparable from the consistent statistical power of tests. There are few missing variables for the countries chosen, and linear interpolation is used to impute these; missing values are filled in by calculating the midpoint value between adjacent observed points, and if leading and trailing values are missing they are filled in by the values after or before, respectively. This is done for simplicity, and it supports the specification of the Ordinary Least Squares (OLS) regressions that follow, preserving the time series trends of the variable in question per country.

Chapter 4 Method

4.1 CS-GFM

Stacked cross-sectional regressions will be carried out with the q-month equity premium of country i as the dependent variable, and the 5 state variables observed at the start of the q-month period (at time $t-1$) as independent variables. Therefore, at time $t-1$ the state variables are observed, and the q-month equity premium over t to $t+q-1$ is estimated with the below Equation 1.1:

$$r_{i,t:t+q-1} = r_{z,t:t+q-1} + \sum_{j=1}^{k} r_{j,t:t+q-1} S_{i,j,t-1} + e_{i,t:t+q-1} \quad (1.1)$$

Here, $\mathrm{r}_{\mathrm{i,t:t+q-1}}$ is the equity premium of country $i = 1, \dots, n$, over the period t to $t+q-1$, k is the number of factors or state variables, and $S_{i,j,t-1}$ is a country characteristic j for country i observed at $t-1$. $r_{j,t:t+q-1}$ represent returns of zero-investment portfolios over t to $t+q-1$ with exposure equal to one to factor j and no exposure to all other factors (long on countries with high exposure to the factor j, and short on countries with low exposure to factor j), as per Fama (1976) and Fama and French (2019). Therefore, $r_{j,t:t+q-1}S_{i,j,t-1}$ is the return contribution of state variable j to the equity premium of country i based on its value observed at $t-1$. Therefore, in the case of negative factor premiums, a positive return is associated with short positions on countries with high exposure to factor j, and long on countries with low exposure to factor j. The intercept $r_{z,t:t+q-1}$ is the excess return on an equally weighted portfolio of the nine countries used in the estimation. $e_{i,t:t+q-1}$ is the error term. It folllows that country i's equity premium is equal to equally weighted global equity market premium plus factor premiums compensating investors for risk associated. q-month returns are cumulative and are compounded using a geometric mean approach: The logarithmic returns are summed over the 60-month period and exponentiated to obtain the cumulative return, following standard practice in financial literature such as in Fama and French (1988).

The multi-period forecasts of country equity premium $r^{e}_{i,t+1:t+q|t}$ are done over the $t+1$ to $t+q$ period, based on information available through month t, using the stored coefficient estimates from Equation 1.1, in Equation 1.2.

$$r^{e}_{i,t+1:t+q|t} = \bar{r}_{z,t} + \sum_{j=1}^{k} S_{i,j,t}\bar{r}_{j,t} \quad (1.2)$$

$r^{e}_{i,t+1:t+q|t}$ is the forecasted 60-month equity premium for country i over the period $t+1$ to $t+q$, based on information available up to month t. $\bar{r}_{z,t} = \frac{1}{M-q}\sum_{m=1}^{M-q}\hat{r}_{z,t-q-m}$ and $\bar{r}_{j,t} = \frac{1}{M-q}\sum_{m=1}^{M-q}\hat{r}_{j,t-q-m}$ are the monthly average intercept and slope estimates from (1), where the averages are calculated over the period M to $t-q$.

The time-varying nature of the average slopes captures global temporal common shocks. The cross-sectional nature also captures heterogeneity in structural country-specific characteristics. Therefore, both persistent structural factors and transient shocks affecting country equity premiums are captured. The out-of-sample forecasts use an expanding window after the training period (such that forecasts use all data available data at the time of the forecasting decision), with lenth M: the number of months from the start of the sample period until t. q is 60 months, given the aim to forecast the five-year equity premium. A training period of 60 months will be used in an effort to balance a large enough training period to avoid overfitting, but also have a large enough number of forecasts.

4.2 In-sample

Prior to the forecasts, an in-sample analysis will be done. 5-year factor premiums for exposure to the state variables will be estimated in-sample, over the entire sample period, using the stacked cross-sectional regressions with Equation 1.1. q is 60. Average annual post-formation factor premiums based on the state variables for years 1-5 will also be computed, along with their respective Sharpe ratios; the latter is computed by dividing the annual average post-formation factor premium by the standard deviation of these annual premiums over a specified horizon, meaning the volatility is the standard deviation of the annual factor premiums. The cross-sectional regression from Equation 1.1 will be re-run, using as dependent variable monthly country excess returns t to $t + 59$ months in the future (not cumulative), with independent variables country state variables observed at month $t - 1$. This will generate every month 60 slope coefficients, that measure the future monthly factor premiums associated with exposure to state variables observed at $t - 1$. The holding periods are non-overlapping; for example, the return displayed for year 5 is the average factor premium in the fifth year (which is the time series average of the 12-month slopes in year 5) after portfolio formation. Average monthly slopes will therefore represent estimates of 60-month monthly future factor premiums based on exposure to state variables at $t - 1$.

Aside from the persistence of the premiums associated with the state variables, it is also important to investigate the persistence of the state variable themselves. For each state variable, the monthly cross-sectional correlation between it (X) will be calculated, measured q periods apart as per Bali, Engle, and Murray (2016). For each time period t, persistence ($\rho_{\mathrm{t,t+q}}(X)$) is defined as the cross-sectional Pearson product-moment correlation between X measured at time t and X measured at time $t + q$. Therefore, $\rho_{t,t+q}(X) = \frac{\sum_{i=1}^{n}[(x_{i,t}-\bar{x}_t)(x_{i,t+q}-\bar{x}_{t+q})]}{\sqrt{\sum_{i=1}^{n}(x_{i,t}-\bar{x}_t)^2}\sqrt{\sum_{i=1}^{n}(x_{i,t+q}-\bar{x}_{t+q})^2}}$, where x is the state variable. $\rho_{t,t+q}$ is the cross-sectional correlation for each month t, and q ranging from 1 to 60 months. n is the number of countries.

4.3 Benchmarks

The first benchmark the CS-GFM will be compared with are two panel predictive models; the first is a panel country fixed effects regression (Equation 2.1). Its purpose is to control for unobserved time-invariant heterogeneity across countries, capturing time-invariant country-specific characteristics.

$$r_{i,t:t+q-1} = a_i + \sum_{j=1}^{k} \beta_j S_{i,j,t-1} + e_{i,t:t+q-1} \text{ (2.1)}$$

a_i is the time-invariant country-specific fixed effects; this represents the country-specific component of variation in equity premiums, accounting for unobserved, time-invariant heterogeneity. β_j are the common slope coefficients which capture the average effect of the predictor variable on the equity premium across all countries, assuming a homogeneous effect across countries. This approach, with common slopes but differing intercepts, is consistent with a factor model in which country equity premiums depend on common global factors plus a country-specific component. The key difference between this model and the CS-GFM is that the former assumes homogeneous predictor effects and time-invariant country characteristics, while the CS-GFM uses cross-sectional variations and time-varying factor premiums. Thus, if time-invariant country-specific factors are indeed the most important determinants of country equity premiums, the country fixed effects model may outperform the CS-GFM. The resultant forecasts are illustrated in Equation 2.2:

$$r^e_{i,t+1:t+q|t} = \hat{a}_i + \sum_{j=1}^{k} \hat{\beta}_j S_{i,j,t} \text{ (2.2)}$$

The second panel predictive model is a modified panel time fixed effects regression, given in Equation 3.1. Its purpose is to capture period-specific effects that impact all countries simultaneously, accounting for global common shocks.

$$r_{i,t:t+q-1} = a_t + \left(\sum_{p=1}^{P} \frac{\sigma_{t,p}}{\sum_{p=1}^{P} \sigma_{t,p}} \hat{b}_{t,p} \right) \sum_{j=1}^{k} S_{i,j,t-1} + e_{i,t:t+q-1} \text{ (3.1)}$$

a_t is the time fixed effects term which captures common time-specific effects each month, representing global common shocks. $\hat{b}_{t,p}$ is the estimated slope coefficient for period t in cross-sectional regression p. $\sigma_{t,p}$ is the cross-sectional volatility for period t. $w_{t,p} = \frac{\sigma_{t,p}}{\sum_{p=1}^{P} \sigma_{t,p}}$ is the weight for period t based on cross-sectional volatility. Unlike conventional time fixed effects regressions that assume constant slope b, the coefficients here are weighted based on cross-sectional volatility each period, assigning more importance to periods with higher variability. This is consistent with a global factor model where country equity premiums depend on common global shocks and exposure to state variables weighted for uses period-by-period volatility. The resultant forecasts are illustrated in Equation 3.2.

The difference between the time fixed effects specification and the CS-GFM is the time fixed effects model uses period-by-period volatility-weighted coefficients, whereas the CS-GFM uses an average of the cross-sectional slope coefficients for the forecast. The time fixed effects model can be almost likened to the CS-GFM, with differences in the slope coefficients. Indeed, Pástor, Stambaugh, and Taylor (2017) show that, for one independent variable, the slope estimator using the Fama and MacBeth (1973) regression as in the CS-GFM is the same to the estimator generated by a panel regression with time fixed effects if the panel is balanced and the variance of the independent variable is constant across all time-periods.

$$r^e_{i,t+1:t+q|t} = \hat{a}_t + \left(\sum_{p=1}^{P} \frac{\sigma_{t,p}}{\sum_{p=1}^{P} \sigma_{t,p}} \hat{b}_{t,p}\right) \sum_{j=1}^{k} S_{i,j,t} \ (3.2)$$

The other benchmark is time series predictive "kitchen sink" regression models. The purpose is to capture country-specific relationships between state variables and equity premiums over time, accounting for temporal dependencies within each country's data. Below is the model:

$$r_{i,t:t+q-1} = a + \sum_{j=1}^{k} \beta_{i,j,t} S_{i,j,t-1} + e_{i,t:t+q-1} \ (4.1)$$

Here, $\beta_{i,j,t}$ are the slope coefficients for each predictor variable $S_{i,j,t-1}$, which is observed at time $t - 1$. The intercept a represents the country's equity premium over the specified period without exposure to the included state variables. This is consistent with a factor model in which each country's equity premium is forecasted based on its own time series of predictor variables, capturing country-specific dynamics and predictor relationships over time. If country-specific trends, cycles, and shocks are the most important determinants of country equity premiums, this model may outperform the CS-GFM. The resultant forecast is displayed in Equation 4.2:

$$r^e_{i,t+1:t+q|t} = \hat{a} + \sum_{j=1}^{k} \hat{\beta}_{i,j,t} S_{i,j,t} \ (4.2)$$

In summary, for all benchmark models, forecasts are carried out in the same way. The training period is 60 months, and an expanding window is used thereafter. The estimated coefficients over the expanding window are used to predict the q-month equity premium for each country for the next month. This all ensures comparability with the CS-GFM. Every forecast (both CS-GFM and benchmarks) is compared against the historical average; this is a forecast per country based on the mean equity risk premium of that country, using an expanding window with the same sample period as the rest of the analysis.

4.4 Forecast Evaluation

The statistical evaluation of the 60-month equity premium forecasts will use the Campbell and Thompson (2008) out-of-sample R^2, given by $R_{OOS}^2 = 1 - \frac{\widehat{MSFE}_i}{\widehat{MSFE}_{i,h}}$. The one-sided Clark and West (2007) statistic will be used to test $H_0: R_{OOS}^2 \leq 0$, $H_1: R_{OOS}^2 > 0$. A positive value for the R_{OOS}^2 suggests that the model's out-of-sample predictions of 5-year equity premiums reduce the MSE compared to forecasts based on the historical average equity premium, indicating outperformance.

As important as statistical forecast evaluation is, investors want to know if statistical predictability can translate to economic gains from portfolio allocation strategies. Rapach and Zhou (2013) comment that the relationship between out-of-sample R^2 and economic utility is complicated. Kandel and Stambaugh (1996), Cenesizoglu and Timmermann (2012) and Timmermann (2018) all argue and present evidence that even statistically weak forecasting models can produce economic gains. On the other side, Cederburg et al. (2019) argue that good forecasting models applied to asset allocation may not necessarily produce economic benefits.

A mean-variance investor is assumed, with a five-year investment horizon who decides the optimal asset mix at the start of it, and rebalances monthly[2]. There is a separate portfolio for each country, where the decision is whether to allocate capital to the country index with weight equal to $w_t = \frac{1}{\gamma} \frac{\mu_{t+q}}{\sigma_{t+q}^2}$, or the risk-free rate, with weight $1 - w_t$. μ_{t+q} is the computed long-term equity premium forecast with horizon q, σ_{t+q}^2 is the expected variance of the country equity premium, and γ is the risk-aversion coefficient of the investor. $\gamma = 5$ is assumed, which is a moderate risk aversion. Therefore, allocation decisions are made for each individual country, and not for a portfolio of multiple countries, for simplicity. For example, if Country A's predicted equity premium by the model is higher, an investor may increase their allocation to Country A's equities (increase the portfolio weight of the country index) to capitalize on these anticipated higher returns. This is done within a mean-variance optimization framework, where the investor balances the expected returns against the associated volatility, while also considering their risk aversion.

The assumed investor uses all data until T_0 (the strategy's start day) to estimate the prediction model (as done in the regressions and forecasts), and then uses the forecasts to choose portfolio weights. The investor rebalances the portfolio to the strategic benchmark monthly until T (the strategy's end day). This results in a series of portfolio weights and returns, one terminal wealth at the end of the horizon $T_0 + q$ (denoted as W_{t+q}) and a realized utility of this terminal wealth $u(W_{t+q})$. An expanding window of data is used with starting date

[2] Note that the monthly rebalancing does not change the five-year investment horizon. It is an adjustment to ensure adherence to the mean-variance optimization framework, maintaining an asset allocation consistent with the five-year investment strategy.

T_0. The economic performance of asset allocation strategies using a prediction model is based on the average out-of-sample realized utility of terminal wealth of all the long-term asset allocation strategies. The Certainty-equivalent-return (CER) will be used to address economic significance as per Sakkas and Tessaromatis (2022), defined as:

$$1 + CER_i = ((1-\gamma)\bar{u}_i)^{\frac{1}{1-\gamma}}, u_i = \frac{1}{T - 60 - T_0 + 1} \sum_{t=T_0}^{T-60} \frac{W_{t+60}^{1-\gamma}}{1-\gamma}$$

Here, u_i is the average of realized utilities of 60 month investment strategies for country i using the prediction model, γ is the coefficient of risk aversion, T_0 is the starting date and T the ending date. The Certainty Equivalent return difference between the asset allocation based on model i and the historical average is: $\triangle CER_i = CER_i - CER_{ih}$. To test whether the $\triangle CER_i = 0$, the one-sided Diebold and Mariano (1995) test will be used on the time series of realized utility values as in the work of Gargano et al. (2019) to test $H_0{:}\, \Delta CER_i = 0, H_1{:}\, \Delta CER_i > 0$. A positive value for the ΔCER_i suggests that the economic utility gains for investors from long-term asset allocation strategies based on the model's forecast for the selected emerging markets are higher compared to strategies based on the historical average equity premium.

4.5 Robustness

As in Sakkas and Tessaromatis (2022), all t-statistics in this paper are corrected with the Newey and West (1987) standard errors, using the lag truncation parameter recommended by Lazarus et al. (2018), given by $S = 1.3\sqrt{T}$. S is the lag truncation parameter and T is the number of observations in the time series of the variable in question. The Lazarus et al. (2018) correction leads to more conservative estimates than regular Newey-West standard errors and is more robust against autocorrelation. The standard errors are also robust to heteroskedasticity by design of the Newey-West correction.

CHAPTER 5 Results & Discussion

5.1. Global Factor Premiums

Firstly, global 5-year factor premiums are estimated as per Equation (1.1). Preliminary Variance Inflation Factor (VIF) analysis shows a high VIF score for DY and PE (15.65 and 15.16 respectively) if Equation (1) is run with all 5 state variables, indicating high multicollinearity. As such, multivariate regressions including both DY and PE will not be estimated. DY or PE are combined with STIR, TS and MVOL in the multiple regression models as shown in Table 1; VIF scores of all variables in the multivariate models fall below 2, which is good for the purposes of this analysis to ensure standard errors are not inflated and coefficients are stable, ensuring interpretability of results. Univariate regressions for each state variable are also run. Table 1 displays the average global 5-year factor premium estimates from the in-sample regression combinations.

Table 1

Global 5-year Factor Premiums

	1	2	3	4	5	6	7
Intercept	-0.039***	-0.039***	-0.039***	-0.039***	-0.039***	-0.039***	-0.039***
(t-stat)	(-24.42)	(-24.42)	(-24.42)	(-24.42)	(-24.42)	(-24.42)	(-24.42)
DY	0.010**					0.011**	
(t-stat)	(2.12)					(2.49)	
PE		0.001					-0.006***
(t-stat)		(-0.77)					(-2.80)
STIR			-0.013***			-0.016***	-0.016***
(t-stat)			(-3.79)			(-4.56)	(-4.32)
TS				-0.008***		-0.011***	-0.016***
(t-stat)				(-4.96)		(-7.12)	(-9.41)

MVOL	-0.004***	-0.002**	-0.001
(t-stat)	(-2.62)	(-2.06)	(-0.63)

Note. Displayed are the average parameter estimates from cross-sectional regression $r_{i,t:t+59} = r_{z,t:t+59} + \sum_{j=1}^{k} r_{j,t:t+59} S_{i,j,t-1} + e_{i,t:t+59}$. Using Equation 1.1 with $q = 60$, $r_{i,t:t+59}$ is the cumulative 60-month excess market return for country i, $r_{z,t:t+59}$ is the intercept, $S_{i,j,t-1}$ is the value of the state variable j for country i observed at time t-1, and $r_{j,t:t+59}$ is the 60-month returns of zero-investment portfolios with exposure equal to one to factor j and no exposure to all other factors (Fama, 1976; Fama & French, 2019). Models 1–5 use the single state variables DY, PE, STIR, TS and MVOL for a univariate regression, while models 6 and 7 are multiple regression models that combine country market DY or PE with STIR, TS and MVOL. *, ** and *** indicate significance at the 10%, 5% and 1% levels respectively. The sample period is from February 1999 to May 2022.

As seen in Table 1, the average 60-month cumulative excess return on an equally weighted portfolio of the countries included in the regressions is -3.9%[3] over January 2004 to May 2022[4]. The average 60-month factor premiums associated with STIR, TS and MVOL are statistically significantly different from zero at the 1% level, and are -0.13%, -0.8% and -0.4% respectively. This means that the average 5-year factor premiums associated with investing long in countries with low STIR, TS, MVOL and short in countries with high STIR, TS, MVOL, respectively, amount to 0.13%, 0.8% and 0.4% on average. These factor premiums are calculated based on the ranks of the state variables across the nine countries, where countries are ranked from lowest to highest values of the state variables, and positions are taken accordingly. The average 60-month factor premium for DY is statistically significant at the 5% level and amounts to 0.1%. Therefore, the average 5-year valuation factor premiums associated with investing long in high DY countries and short in low DY countries

[3] The negative value may be surprising given the largely positive equity market return values in the descriptive statistics in Table 6 Panel A, but it should be noted that those are displayed at a monthly frequency. In contrast, the returns used in the Table 1 regressions are the cumulative 60-month returns. The accumulation method detailed in Section 4.1 explains that the cumulative 60-month returns are calculated using the returns of the following 60 months (therefore including both positive and negative values), rather than compounding a single monthly return by 60. This approach accounts for fluctuations over the period, and can result in a negative average cumulative return of the equally weighted country portfolio, despite often positive individual monthly returns.

[4] The intercept $r_{z,t:t+59}$ is the same for all regressions because it is the 60-month cumulative excess return on an equally weighted portfolio of the nine countries, calculated for the same sample period for all models, with the state variables standardised as mentioned in Chapter 3. Consequently, it does not change with inclusion or exclusion of different state variables in the different models. Exposure to the state variables is associated with average change in returns as represented by the coefficients shown in Table 1.

is 0.1% on average. Interestingly, and contrarily to the emerging market return prediction literature (Hollstein et al. (2020), etc.), exposure to the price-earnings ratio is not associated with a significant factor premium.

A multivariate regression combining DY with STIR, TS and MVOL (model 6) generates statistically significant factor premiums for all state variables included. The 60-month factor premiums associated with DY and MVOL exposure are statistically significant at the 5% level, amounting to 1.1% and -0.2% each. The premiums associated with STIR and TS as with the univariate regressions are statistically significant at the 1% level; they amount to -1.6% and -1.1%.

A multivariate regression that instead combines PE with STIR, TS and MVOL (model 7) leads to statistically significant 60-month factor premiums for PE, STIR and TS at the 1% level, which are respectively -0.6%, -1.6% and -1.6%. We see that compared to the univariate regressions, exposure to PE at t - 1 now becomes significant, whereas exposure to MVOL is now insignificant.

Overall, the information in Table 1 suggests that state variables observed today predict factor premiums in-sample over horizons up to 5 years. The factor premiums, although statistically significant, are very small in magnitude as seen from Table 1; the highest absolute value of a significant factor premium is 1.6%. Consequently, this is of little practical use to investors due to the low cumulative returns exposure to the factors would have generated. This may indicate that despite statistically significant factor premiums, they will not be economically significant in terms of producing significant utility gains to the investor assumed in Section 4.4; this will be confirmed out-of-sample in Section 5.2. The signs for the factor premiums that are significant are consistent with the signs for the in-sample factor premiums in the same variables the country return prediction literature (Hollstein et al., 2020), indicating reliability of the analysis.

Table 2 displays the average annual post-formation factor premiums in Panel A, and their corresponding Sharpe ratios in Panel B. Average monthly slopes represent estimates of 60-month monthly future factor premiums based on exposure to state variables at $t - 1$.

Table 2

Annual Post-Formation Factor Premiums

	1 year	2 years	3 years	4 years	5 years
A: Average premium per year					
DY	0.0126	0.0124	0.0121	0.0150	0.0169
PE	-0.0048	-0.0043	-0.0036	-0.0018	0.0009

STIR	-0.0349	-0.0345	-0.0334	-0.0339	-0.0325
TS	-0.0056	-0.0050	-0.0038	-0.0007	-0.0004
MVOL	0.0055	0.0036	0.0064	0.0041	0.0062
B: Sharpe ratio					
DY	0.1460	0.1479	0.1314	0.1539	0.1626
PE	-0.0916	-0.0736	-0.0627	-0.0321	0.0202
STIR	-1.0084	-0.9307	-0.9682	-1.0096	-0.9474
TS	-0.1299	-0.1065	-0.0951	-0.0195	-0.0098
MVOL	0.1467	0.0872	0.1422	0.0752	0.1149

Note. Displayed are annual post-formation factor premiums based on the state variables for years 1-5. Panel A reports annual average factor premiums, and panel B their respective Sharpe ratios; the latter is computed by dividing the annual average post-formation factor premium by the standard deviation of these annual premiums over the specified horizon, meaning the volatility is the standard deviation of the annual factor premiums. The holding periods are non-overlapping; for example, the return displayed for year 5 is the average factor premium in the fifth year (which is the time series average of the 12-month slopes in year 5) after portfolio formation. Factor premiums are estimated by running each month cross-sectional Equation 1.1 with $q = 60$, with dependent variable monthly excess returns over t to $t + 59$ months in the future on country state variables observed at $t - 1$. Therefore, average monthly slopes represent estimates of 60-month monthly future factor premiums based on exposure to state variables at $t - 1$. Factor premiums represent returns of zero-investment portfolios with exposure equal to one for factor j, and no exposure to all other factors (Fama, 1976; Fama & French, 2019), and are reported representing long positions on countries with high exposure to factor j, and short on countries with low exposure to factor j. For negative factor premiums, a positive return is associated with short positions on countries with high exposure to factor j, and long on countries with low exposure to factor j. The sample period is from February 1999 to May 2022.

Table 2 shows that the factor premiums associated with DY exposure (long high DY countries and short low DY countries) are 1.26% in the first year and then slightly increase to 1.69% in year five. The decline is gradual and small, with the same sign throughout, indicating persistence in the factor premium of this state variable. The mostly increasing risk-adjusted return represented by the Sharpe ratio (from 0.1460 in the first year to 0.1626 in year 5) also suggests this. Overall, the DY premium shows strong persistence. The factor premium for PE exposure starts at -0.48% and rises to 0.09% by year 5, becoming less negative each year and turning positive in year 5. This fluctuation and sign change indicate weaker persistence than DY; this is also indicated by the low and fluctuating Sharpe ratios (from -0.0916 in year 1 to 0.0202 in year 5, with inconsistency). Therefore, the PE premium shows weak persistence.

The STIR factor premium starts at -3.49% in year 1 and very slightly decreases to -3.25% by year 5. The negative and only slightly changing (therefore stable) premiums indicate strong persistence. The risk-adjusted return remains consistently highly negative throughout, from -1.0084 in year 1 to -0.9474 year 5, indicating a persistently poor risk-return profile; overall, this indicates strong persistence. The TS factor premium is consistently negative, starting at -0.56% in year 1 and improving to -0.04% by year 5, indicating some persistence. This is also shown by the increasing Sharpe ratios, from -0.1299 in year 1 to -0.0098 in year 5. Overall, the persistence of the TS premium is good. Finally, the MVOL premium is small and positive[5], starting at 0.55% in the first year and fluctuating slightly to 0.62% by the fifth year, but it is still relatively stable, suggesting some persistence. The Sharpe ratio starts at 0.1467 in year 1 and ends at 0.1149 in year 5. Overall, the MVOL premium shows moderate persistence.

The evidence in Tables 2 and 3 suggest that exposure to state variables at t - 1 is rewarded with factor premiums far into the future, and that some of these factor premiums associated with the state variables are more persistent than others. Figure 1 displays the time series average of the monthly cross-sectional correlations of the state variables with their lags, for up to 60-month lags; this represents the persistence of the state variables themselves.

[5] The positive premium for MVOL in years 1-5 in Table 2 contrasts with the negative 5-year cumulative factor premium in Table 1. This may be reconciled by the difference in time horizons. While the average of monthly premiums during each individual year is positive, the cumulative effect over five years is negative due to the compounding impact of negative returns in some periods, which can outweigh the positive returns in other periods.

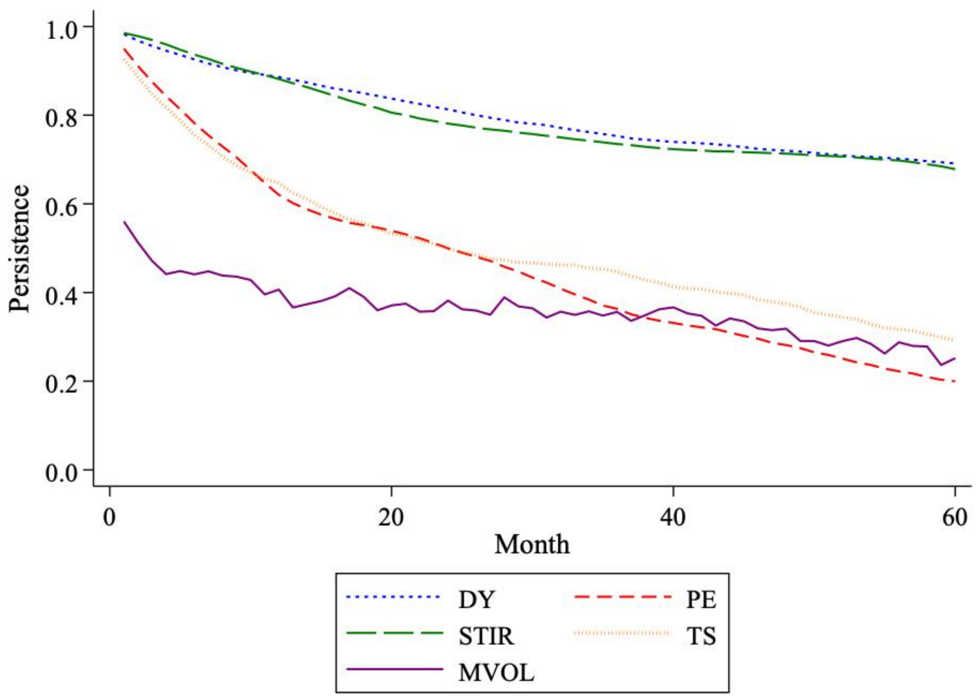

Figure 1

Persistence of the State Variables

Note. Displayed are the time series average of the monthly cross-sectional correlations between the state variable X, measured q months apart which in this figure is from 1 to 60, defined as the persistence of each state variable. This persistence measure is calculated following the Bali, Engle, and Murray (2016) procedure described in Section 4.1.

Figure 1 confirms that DY and STIR are highly persistent, with correlations close to 1 in the first year, only falling slightly below 0.7 at the end of the 60-month lag period. The correlations for both TS and PE start high in the first months but fall significantly as the lag horizon increases; by 40 months, both are below 0.4. There is less persistence for MVOL, starting at below 0.6 in the first month, and fluctuating throughout to end on less than 0.3. The high persistence of DY and STIR explains the statistically significant 60-month factor premiums reported in Table 1, whether measured based on univariate or multivariate cross-sectional regressions (1% for STIR, 5% for DY). Despite the statistical significance of a TS premium at the 1% level when measured based on both univariate or multivariate regressions, the variable itself shows less persistence than STIR or PE, which means the persistence of TS and its in-sample power are not as linked as for variables like STIR and MVOL. Contrasting with these findings, for developed markets, Sakkas and Tessaromatis (2022) find generally higher persistence of the state variables, with stronger and more consistent annual post-formation factor premiums and Sharpe ratios.

5.2 Forecast Results

This section will evaluate the performance of the CS-GFM and the benchmark models in generating 60-month country equity premium forecasts, using the methodology described in Chapter 4. The statistical evaluation of CS-GFM forecasts is displayed in Table 3 below.

Table 3

Statistical Evaluation of CS-GFM Forecasts

	DY	PE	STIR	TS	MVOL	DY, STIR, TS, MVOL	PE, STIR, TS, MVOL
Country / model based on	R^2_{oos}	R^2_{oos}	R^2_{oos}	R^2_{oos}	R^2_{oos}	R^2_{oos}	R^2_{oos}
Greece	0.093***	0.147***	0.066**	0.163**	0.166***	0.177**	0.156**
	(2.42)	(2.39)	(2.30)	(2.32)	(2.37)	(2.22)	(2.19)
South Africa	0.552**	0.533**	0.461**	0.560**	0.523**	0.409*	0.352
	(2.26)	(2.17)	(1.75)	(2.25)	(2.16)	(1.39)	(1.22)
Hungary	0.479***	0.437***	0.441***	0.465***	0.458***	0.465***	0.406***
	(3.35)	(3.29)	(2.95)	(3.36)	(3.32)	(3.23)	(2.94)
India	-1.543	-0.814	-0.040	-0.614	-0.428	-1.140	-0.436
	(-2.09)	(-4.08)	(-1.61)	(-2.87)	(-2.99)	(-2.06)	(-2.51)
Thailand	0.489***	0.506***	0.590***	0.565***	0.513***	0.635***	0.614***
	(2.53)	(2.60)	(2.82)	(2.87)	(2.57)	(2.96)	(2.80)
Philippines	0.534***	0.541***	0.550***	0.345***	0.484***	0.485***	0.476***
	(6.25)	(6.33)	(6.29)	(3.81)	(5.67)	(5.26)	(5.33)
Malaysia	0.508*	0.532*	0.598**	0.581***	0.553***	0.650***	0.610**
	(1.32)	(1.60)	(2.05)	(2.66)	(2.74)	(2.73)	(2.34)
Taiwan	0.253*	0.239*	0.510***	0.426***	0.279**	0.561***	0.494**
	(1.53)	(1.42)	(2.85)	(2.89)	(1.75)	(2.75)	(2.08)
Pakistan	0.944***	0.958***	0.969***	0.952***	0.951***	0.956***	0.963***

	(6.87)	(7.63)	(6.84)	(6.92)	(6.99)	(5.43)	(5.79)
Pooled	0.771***	0.789***	0.807***	0.791***	0.788***	0.800***	0.800***
	(4.58)	(4.71)	(4.84)	(4.74)	(4.72)	(4.47)	(4.47)

Note. Displayed are the R^2_{oos} statistics for country equity premium forecasts based on the CS-GFM cross-sectional predictive regression. Single variable models are based on the state variables DY, PE, STIR, TS and MVOL; multiple factor models are based on a combination of DY or PE with STIR, TS and MVOL. R^2_{oos} statistics are provided for all countries, and the pooled category which includes all countries. In parentheses are the one-sided Clark and West (2007) MSFE-adjusted t-statistics, corrected using the Newey and West (1987) methodology with the lag truncation parameter suggested by Lazarus et al. (2018). *, ** and *** indicate significance at the 10%, 5% and 1% levels respectively. The first forecast uses historical data from January 2004 until the end of December 2008. Subsequent forecasts use an expanding window of data.

It can be seen from Table 3 that for the pooled case, CS-GFM forecasts are associated with a reduction in MSE for the predictive regression relative to the historical average forecast for all models at the 1% significance level: for both univariate and multivariate cases we are able to reject the null hypothesis $R^2_{OOS} \leq 0$ at the 1% level for the pooled case. There is sufficient evidence to suggest the model forecasts outperform the historical average-based forecast in terms of MSFE reduction for all individual state variables, and for the two multiple variable combinations, in this sample for the pooled case. The individual predictor STIR sports the highest R^2_{oos} with 80.7%[6], followed by the multivariate regressions which both have 80%. This means that for instance with STIR, the CS-GFM univariate regression with it, reduces, on average, the MSFE by 80.7% compared to the historical average, at the 1% significance level.

At the 1% level, the most successful individual predictor for 60-month equity premiums is MVOL; it is statistically significant at this level for 6 out of 9 countries. This is consistent with the high volatility characteristic of emerging markets (Narayan et al., 2014), for which it would be expected that market volatility is a strong predictor of the equity premium. When using the 5% level, STIR, TS and MVOL are tied for the best performance; for all three, there is a significant reduction in MSFE of the model compared to the historical average for 8 out of 9 countries; for DY and PE this is 6 out of 9 countries.

Both CS-GFM models perform the exact same for the pooled case. For individual countries, the model combining DY with STIR, TS and MVOL is more successful than its counterpart with PE; there is sufficient evidence to suggest it reduces MSFE compared to the historical average at the 1% level for 6 out of 9 countries,

[6] This does not imply that while STIR shows high persistence and effectively reduces forecast errors out-of-sample, it captures 80.7% of the in-sample variance of stock returns like the traditional in-sample R^2. Rather, a forecast based on a CS-GFM using it as a predictor variable significantly improves forecast accuracy over the historical average benchmark, at the 1% level.

compared to the counterpart's 4 out of 9. This is consistent with the higher in-sample predictability of DY compared to PE as seen in Table 1. The combination with DY rather than PE also generally has higher R_{oos}^2 per country too, for all countries where it is significant apart from Pakistan.

Unlike in Sakkas and Tessaromatis (2022), here multiple variable forecasts have a higher R_{oos}^2 compared to single variable forecasts (apart from the one based on STIR). Evidently the overfitting associated with a highly parameterized model usually deemed responsible for underperformance of multiple variable predictive regression forecasts (Goyal and Welch, 2008; Rapach and Zhou, 2013)[7] was not problematic in this case. The information captured in the CS-GFM's multiple regression combinations of the individual variable forecasting model predictors must have outweighed this.

The link between the statistical performance of the state variables and their persistence is weaker, unlike in the work of Sakkas and Tessaromatis (2022). For example, DY has strong persistence yet is one of the worst individual predictors along with PE. PE shows low persistence and poor predictive performance, aligning with expectations. Interestingly, the most successful predictor in terms of effectiveness in different countries is MVOL, which is interesting given its subpar persistence. This may be due to the higher volatility and instability of emerging markets (Bekaert & Harvey, 1997), making the market volatility a key driver of country equity premiums, by capturing changes in market conditions and investor risk aversion[8]. Also, interest variables like STIR and TS are better out-of-sample predictors compared to valuation ratios, which is consistent with the emerging market country return prediction literature (Bahrami et al., 2018; Charles et al., 2017; Hollstein et al., 2020; Hjalmarsson, 2010). This is unlike what Sakkas and Tessaromatis (2022) found, which is that the best performing predictors are the valuation ratios DY and PE. This could be linked to the volatile environment in emerging markets leading to a higher sensitivity to macroeconomic conditions such as interest rates and market volatility; for instance, higher sensitivity to monetary policy changes compared to

[7] Goyal and Welch (2008) and Rapach and Zhou (2013) find that forecasts based on time series multiple regression forecasting ("kitchen sink") models underperform single variable forecasts of the same type. Rapach and Zhou (2013) attribute the poor out-of-sample performance of the multiple regression to the overfitting associated with highly parameterized models.

[8] The nature of the MVOL computation (standard deviation of daily stock returns in the month) makes it highly sensitive to changes in market conditions. The resultant fluctuation month-by-month also means the MVOL coefficients in predictive models can take longer before the model picks up structural changes in the data, compared to a predictor like PE which is a direct data point. The CS-GFM is robust against this; the stored coefficients are averaged over the expanding window for the forecasts, reducing the sensitivity to a single month's volatility spike and emphasizing prediction based on the long-term trend. It can be an issue, however, for the benchmark models, particularly the panel predictive and time series models, which do not use averaging and may suffer from parameter instability and model uncertainty due to the high volatility in MVOL.

developed economies (Frankel, 2010). Consequently, macroeconomic variables are better proxies for these than valuation ratios, which may not completely reflect these rapid changes.

Table 4 displays the economic evaluation performance of the CS-GFM combinations.

Table 4

Economic Evaluation of CS-GFM Forecasts

	DY	PE	STIR	TS	MVOL	DY, STIR, TS, MVOL	PE, STIR, TS, MVOL
Country / model based on	ΔCER	ΔCER	ΔCER	ΔCER	ΔCER	ΔCER	ΔCER
Greece	-0.282	0.792***	-0.285	1.882	1.877	-0.295	-0.337
	(-1.79)	(8.65)	(-1.72)	(1.14)	(1.19)	(-1.79)	(-1.81)
South Africa	-0.142	1.883**	-0.143	2.016	2.016	-0.142	-0.148
	(-1.17)	(2.24)	(-1.17)	(1.12)	(1.16)	(-1.17)	(-1.23)
Hungary	0.004	2.112**	0.004	2.162	2.162	0.005	-0.123
	(0.02)	(2.00)	(0.02)	(1.12)	(1.16)	(0.03)	(-0.51)
India	-0.388	1.471***	-0.388	1.770	1.770	-0.388	-0.388
	(-1.97)	(3.04)	(-1.97)	(1.13)	(1.18)	(-1.97)	(-1.97)
Thailand	-1.11	-0.580	-1.067	1.024**	0.840***	-1.097	-1.072
	(-2.18)	(-1.97)	(-2.20)	(1.79)	(2.64)	(-2.23)	(-2.25)
Philippines	-0.305	1.635***	-0.304	1.854	1.854*	-0.307	-0.303
	(-2.17)	(4.17)	(-2.16)	(1.22)	(1.33)	(-2.20)	(-2.15)
Malaysia	-0.439	-0.039	-0.448	1.677*	1.547**	-0.439	-0.448
	(-2.25)	(-0.54)	(-2.30)	(1.48)	(1.86)	(-2.25)	(-2.30)
Taiwan	-1.038	-0.622	-1.064	1.062**	0.921***	-1.070	-1.069
	(-3.99)	(-3.77)	(-4.08)	(2.00)	(2.94)	(-4.15)	(-4.13)
Pakistan	-0.266	0.643	-0.266	1.663	1.510	-0.266	-0.266
	(-2.01)	(1.19)	(-2.01)	(1.14)	(1.16)	(-2.01)	(-2.01)
Pooled	-0.440	0.812	-0.440	1.679	1.612	-0.444	-0.462

(-1.39)	(0.68)	(-1.51)	(0.37)	(0.39)	(-1.48)	(-1.56)

Note. Displayed is the economic utility of asset allocation strategies from the CS-GFM forecasts against historical average-based forecasts. Single variable models are based on the state variables DY, PE, STIR, TS and MVOL; multiple factor models are based on a combination of DY or PE with STIR, TS and MVOL. ΔCER statistics are presented for all countries, and the pooled category which includes all countries. In parentheses are t-statistics based on the one-sided Diebold and Mariano (1995) test, corrected using the Newey and West (1987) methodology with the lag truncation parameter recommended by Lazarus et al. (2018). *, ** and *** indicate significance at the 10%, 5% and 1% levels respectively. The first forecast uses historical data from January 2004 until the end of December 2008. Subsequent forecasts use an expanding window of data.

As seen in Table 4, in the pooled case, there is insufficient evidence to reject the null hypothesis of $\triangle\ CER_i = 0$ for all models. This means there is insufficient evidence to suggest that strategic asset allocation based on CS-GFM country equity premium predictions is associated with significant utility gains for the investor assumed in Section 4.4, compared to allocation based on the historical average equity premium.

For individual countries, instances of good performance can be seen. Individual predictor PE is the best performing; there is sufficient evidence to reject the null hypothesis $\triangle\ CER_i = 0$, and suggest that the difference in certainty equivalent returns is statistically significantly higher than zero for 5 out of 9 countries at the 5% level (with three of them at the 1% level), with large economic gains. DY, STIR, and both multiple factor models fail to outperform the historical average for all countries. TS and MVOL are successful at the 5% level for 2 and 3 out of 9 countries respectively.

Overall, the statistical superiority of the CS-GFM country equity premium forecasts largely does not translate into better economic performance of the asset allocation strategy based on them, for emerging markets in the sample at the required 5% level. It can be concluded that using the CS-GFM forecasts to guide asset allocation is overall not beneficial to long-term investors deciding their portfolios long-term asset mix. Following the proposed reasons for economic insignificance despite statistical significance of Cederburg et al. (2022), this is likely due to the small magnitude of the factor premiums as seen in Table 1, and the high volatility of emerging markets. While the CS-GFM is able to statistically predict the equity premium accurately, the magnitude of these predictions in terms of driving returns was not substantial enough to yield significant economic benefits when applied to asset allocation strategies. An explanation could be that for instance, if country A's predicted equity premium would lead to a gain that (although very accurate) is small, a moderately risk averse investor may not increase their allocation to country A's equities because they balance the expected returns against the associated volatility in the mean-variance optimization framework; volatility is especially high in emerging markets as mentioned earlier. This could potentially lead to conservative portfolio weights that may not fully capitalize on the model's accurate forecasts. Indeed, Sakkas and Tessaromatis (2022) reported larger in-sample 60-month factor premiums for all state variables for developed economies. The biggest differences are for the valuation ratios DY and PE which are respectively 6.33% and -7.04%. This may explain their economically significant results, compared to the smaller premiums observed for emerging markets, with higher volatility.

Moving on to the benchmark models, Table 5 presents the statistical and economie evaluation results for all benchmarks.

Table 5

Statistical and Economic Evaluation of Panel Predictive Models and Time Series Predictive Models

	DY	PE	STIR	TS	MVOL	DY, STIR, TS, MVOL	PE, STIR, TS, MVOL
Model	*A: Statistical Evaluation* R^2_{oos}						
Time Fixed Effects	0.147*	0.120*	0.136*	0.120*	0.120*	0.164*	0.142*
	(1.51)	(1.52)	(1.52)	(1.52)	(1.52)	(1.51)	(1.52)
Country Fixed Effects	-0.131*	-0.035*	-0.260	-0.120*	-0.006**	-0.431	-0.290
	(1.56)	(1.63)	(0.77)	(1.35)	(1.65)	(0.52)	(0.69)
Time Series	0.031*	-0.111	0.038	0.060	0.001*	-0.844	-0.052
	(1.49)	(1.18)	(1.19)	(1.06)	(1.50)	(-0.33)	(0.76)
	B: Economic Evaluation ΔCER						
Time Fixed Effects	2.208	2.200	2.187	2.187	2.178	2.193	2.191
	(0.41)	(0.41)	(0.41)	(0.41)	(0.41)	(0.41)	(0.41)
Country Fixed Effects	0.079	0.303	0.195	0.251	0.294	0.104	0.181
	(-0.38)	(0.46)	(-0.16)	(0.25)	(0.36)	(-0.29)	(-0.24)
Time Series	0.233	0.261	0.266	0.246	0.268	0.231	0.196
	(-0.02)	(0.15)	(0.21)	(-0.06)	(0.26)	(-0.10)	(0.03)

Note. Displayed are the results for statistical (Panel A) and economic (Panel B) evaluation of the panel predictive models and time series models outlined in Section 4.3. Only the pooled case is presented. For the statistical evaluation, the R^2_{oos} are presented, and in parentheses are the one-sided Clark and West (2007) MSFE-adjusted t-statistics. For the economic evaluation, ΔCER statistics are presented; in parentheses are t-statistics based on the one-sided Diebold and Mariano (1995) test. Both t-statistics are corrected using the Newey and West (1987) methodology with the lag truncation parameter recommended by Lazarus et al. (2018). *, ** and *** indicate significance at the 10%, 5% and 1% levels respectively. The first forecast uses historical data from January 2004 until the end of December 2008. Subsequent forecasts use an expanding window of data.

Predictions from the time fixed effects model (Table 5), while being the best-performing out of the benchmarks across all models, are dissimilar to predictions from the CS-GFM (Tables 4 and 5). For the statistical evaluation pooled case, we are only able to reject the null hypothesis $R^2_{OOS} \leq 0$ at the 10% level for all models, which is insufficient considering the 5% benchmark. This is interesting given the similarities of the model to the CS-GFM. It may be that because emerging markets are typically less integrated than developed markets (Donadelli and Paradiso, 2014) due to differences in in economic conditions, policy environments, and stages of development between each other, they may react less homogeneously to temporal shocks, reducing the predictive power of the time fixed effects model. However, being less integrated does not mean being unintegrated, and some effect from common shocks is expected; the weak result is therefore interesting.

The country fixed effects model generally performs worse than the time fixed effects, except for individual predictor MVOL, for which we are able to reject the null hypothesis of H_0: $R^2_{OOS} \leq 0$ at the 5% level. Emerging markets often experience higher idiosyncratic volatility due to unique country-specific factors (Narayan et al., 2014), perhaps explaining the better performance of MVOL. While country fixed effects capture heterogeneity of country characteristics from the fixed country-specific component, the model cannot adapt to the more frequent changes and shocks, such as policy shifts, political instability, and economic reforms characteristic of emerging markets compared to developed ones (Narayan et al., 2014); this is instead captured by the time fixed effects model.

For the time series linear regression ("kitchen sink") models, we are only able to reject $R^2_{OOS} \leq 0$ at the 10% level in the pooled case for DY and MVOL, which is insufficient. This inferiority is consistent with evidence from Rapach and Zhou (2013) regarding time series linear regression models' ability to forecast equity returns. Parameter instability and model uncertainty are regarded as responsible (Sakkas and Tessaromatis, 2022). This is especially true for emerging markets, which are more likely to undergo frequent economic, political, and regulatory changes (Bekaert & Harvey, 1997), which can lead to unstable relationships between predictors and equity premiums over time, compromising the accuracy of forecasts based on time series models. In contrast, the CS-GFM leverages persistent country-specific characteristics (such as DY and STIR) that remain relatively stable over time (unlike time series models that struggle with unstable parameters), to capture stable fundamental differences across countries for forecasting, allowing for more accurate equity premium predictions even in the volatile environment that is characteristic of emerging markets (Bekaert & Harvey, 1997).

Overall, CS-GFM out-of-sample forecasts of 5-year equity premiums statistically significantly reduce the MSE compared to forecasts based on the historical average equity premium at the 1% level in the pooled case. Therefore, is sufficient evidence to reject $H_0: R^2_{OOS} \leq 0$, and the first null hypothesis of the thesis. The CS-GFM is also more often statistically significant, at better significance levels, with higher R^2_{OOS} values compared to the panel predictive and time series models in the pooled case, meaning it outperforms all the benchmarks for the statistical evaluation. The time-varying slopes capture temporal shocks to all countries, and the cross-sectional approach captures country structural differences in characteristics. Therefore, both persistent structural factors and transient shocks are captured; this leads to strong predictive power in the relatively heterogeneous and volatile context of emerging markets.

In the pooled case, the economic utility gains for investors from long-term asset allocation strategies based on the CS-GFM, panel predictive, and time series model forecasts are not significantly higher compared to strategies based on forecasts based on the historical average equity premium. There is insufficient evidence to reject the null hypothesis $H_0: \Delta CER_i = 0$, and the second null hypothesis of the thesis at the required 5% level. Also, it is unclear which out of the CS-GFM, panel predictive, and time series models is superior due to the insignificant results for all of them in the pooled case. Again, this is likely due to the small magnitude of factor premiums associated with exposure to the state variables. Therefore, despite their significance in-sample and accurate out-of-sample forecasts of the equity premium, these may not lead to significant utility gains from portfolio performance.

5.3 Discussion

As per Stambaugh (1999), in time series regressions, overlapping returns introduce autocorrelation, leading to biased estimates which are problematic as the time horizon increases. In the cross-sectional context of the CS-GFM, each datapoint is independent of others, minimizing the impact of autocorrelation. Boudoukh, Israel, and Richardson (2021) show that in long-horizon predictive time series regressions, the bias of the coefficient estimator regressions increases with a) the horizon, b) the persistence of the predictive variable, and c) it is greater for overlapping regressions compared to nonoverlapping ones. This is commonly addressed in the return prediction literature (And & Bekaert, 2007; Karolyi & Van Nieuwerburgh, 2020).

Overlapping returns leads to inference problems; they lead to misspecification of the standard errors of the average long-horizon factor premiums generated monthly (Sakkas & Tessaromatis, 2022). Consequently, the estimates are unbiased but inefficient. In this paper, the autocorrelation from overlapping returns is corrected by calculating alternative estimates of Newey and West (1987) standard errors using the standard lag truncation parameter of Lazarus et al. (2018) $S = 1.3\sqrt{T}$ as detailed in Section 4.5. These are more conservative than the traditional Newey and West (1987) estimate, and this aids in mitigating spurious regressions from this induced autocorrelation.

Sakkas and Tessaromatis (2022) state that statistical corrections are unlikely to solve the inference problems from overlapping returns because there are few independent observations, especially when the returns in question are 60-month returns. They present arguments in favour of using overlapping returns, such as the existence of evidence of predictability of 5-year equity returns using non-overlapping returns, like in Golez and Koudjis (2018) who use dividend-to-price ratio as predictor and data over the period 1629–2015. They also state that specification errors should lead to poor out-of-sample economic performance, so including economic evaluation in the analysis can circumvent the inference problems caused by overlapping, provided this economic performance is strong.

State variables used in the CS-GFM were chosen using motivation from theory and previous research findings reported by many researchers in the stock return prediction literature as predictors of the equity market premium (Sakkas & Tessaromatis, 2022). The latter is susceptible to data mining, but as the analysis is out-of-sample this is mitigated. So while the data mining argument does not hold, uncorrected t-statistics are likely incorrect. Adjusted t-statistics provide more robust inference compared to unadjusted ones, but the results should still be interpreted with caution given the potential bias from overlapping returns. And while the strong statistical performance of the CS-GFM suggests specification errors from overlapping returns may not have impacted heavily the model, the weak out-of-sample economic performance further supports the need for caution in interpreting the results. Further robustness checks and alternative modelling approaches may be necessary to enhance the practical utility of the predictions and is a recommendation for future research in this sector.

The nature of emerging markets is such that not only are the issues present such as increased volatility due to frequent economic, political, and regulatory changes mentioned in Chapter 2, but that data quality can be compromised. This is mitigated by choosing reliable data sources that are accessible: Refinitiv Datastream and Finaeon Global Financial Data. Still, the data quality may be problematic, and could therefore have led to unreliable results. Quality aside, data availability was undoubtedly the biggest constraint, shortening the sample period considerably. Inevitably this decreased the power of statistical tests, especially for the out-of-sample analysis, and led to a less precise analysis.

Should this paper be replicated, there are many ways to improve upon it. A larger dataset for the same countries would lead to increased power of the statistical tests, therefore a more precise and perhaps conclusive analysis. As the countries to my knowledge do not have reliable data available for a longer period, this likely cannot be done without waiting a substantial period before replicating this analysis. Further tests of the CS-GFM's stacked cross-sectional approach with other theoretically motivated variables proven to be strong out-of-sample predictors for emerging markets will likely lead to valuable insights for long-term asset allocation in emerging markets. Additionally, more benchmark models used in academia and industry can be included to compare the CS-GFM's performance; examples that come to mind are using a principal component analysis (PCA) combination of individual predictors (diffusion indices), the sum-of-the-parts methodology of Ferreira and Santa-Clara (2011), and various machine learning time series models, as Sakkas and Tessaromatis (2022) used for comparison of their CS-GFM for developed markets.

Chapter 6 Conclusion

This thesis is, to my knowledge, the first paper investigating the cross-sectional global factor model (CS-GFM) of Sakkas and Tessaromatis (2020) for emerging economies. By assuming country equity premiums are driven by the CS-GFM, long-term country-level equity premium forecasts were generated for nine emerging countries. Long-term equity market premiums are modelled as the global equity market premium plus additional global equity factor premiums commensurate with each country's exposure to risks proxied by state variables that include valuation ratios price-earnings ratio and dividend yield, macroeconomic variables short-term interest rate and term spread, and market volatility.

State variables were found to be associated with significant but small factor premiums. Variables that are significant in-sample predictors of the equity premium aligned with the rest of the literature on emerging market return prediction, apart from the surprising insignificance of the price-earnings (PE) ratio in a univariate regression when it is the sole predictor. CS-GFM predictions of 5-year equity premiums for all nine emerging markets were statistically significantly better than forecasts based on the average of historical returns, and the other benchmark models. The persistence of the state variables is not as closely linked to long-term equity premium prediction performance as with developed markets. Interest rate variables such as the short-term interest rate (STIR) outperformed valuation ratios (DY and PE) as predictors of country returns, consistent with the rest of the emerging market country return prediction literature. Despite the statistical performance, there was insufficient evidence to suggest the economic utility gains for investors from long-term asset allocation strategies based on CS-GFM forecasts for the selected emerging markets are higher compared to strategies using forecasts based on the historical average equity premium, and the benchmark models. The same is true for the benchmark models compared to the historical average method. This is likely due to the small magnitude of the factor premiums, meaning that accurate forecasts from the CS-GFM do not necessarily translate into substantial utility gains for investors. This is unlike the findings of Sakkas and Tessaromatis (2022) for developed markets, where CS-GFM-based forecasts produce significant utility gains for investors, with larger in-sample factor premiums, especially for DY and PE.

This research faced limitations due to small sample sizes and potentially low data quality, which may have affected its precision and conclusiveness. Despite adjusted t-statistics, overlapping returns may have led to specification errors; this is further a cause of concern given the poor out-of-sample economic performance. Future research should aim to use larger samples, test different theoretically motivated state variables, and compare the cross-sectional global factor model with more empirically validated benchmark models. By addressing these limitations, future studies can provide better insights into the applicability of the CS-GFM for long-term strategic asset allocation in emerging markets. This will improve scientific understanding and generate practical insights for investors.

In conclusion, while this thesis highlights the statistical performance of the CS-GFM in predicting long-term equity premiums in emerging markets, as with developed economies, translating this into practical investment performance remains a challenge. This may be achieved with the further research and methodological improvements mentioned.

References

Ang, A., & Bekaert, G. (2006). Stock Return Predictability: Is it There? *Review of Financial Studies*, *20*(3), 651–707. https://doi.org/10.1093/rfs/hhl021

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The Cross-Section of Volatility and Expected Returns. *The Journal of Finance*, *61*(1), 259–299. https://doi.org/10.1111/j.1540-6261.2006.00836.x

Angelidis, T., & Tessaromatis, N. (2017). Global Equity Country Allocation: An Application of Factor Investing. *Financial Analysts Journal*, *73*(4), 55–73. https://doi.org/10.2469/faj.v73.n4.7

Arnott, R. D., & Bernstein, P. L. (2002). What Risk Premium Is "Normal"?. *Financial Analysts Journal*, *58*(2), 64–85. https://doi.org/10.2469/faj.v58.n2.2524

Bahrami, A., Shamsuddin, A., & Uylangco, K. (2016). Out-of-sample stock return predictability in emerging markets. *Accounting & Finance*, *58*(3), 727–750. https://doi.org/10.1111/acfi.12234

Bakshi, G., & Kapadia, N. (2003). Delta-Hedged Gains and the Negative Market Volatility Risk Premium. *The Review of Financial Studies*, *16*(2), 527–566. https://doi.org/10.1093/rfs/hhg002

Bali, T. G., & Cakici, N. (2010). World market risk, country-specific risk and expected returns in international stock markets. *Journal of Banking & Finance*, *34*(6), 1152–1165. https://doi.org/10.1016/j.jbankfin.2009.11.012

Bali, T. G., Engle, R. F., & Murray, S. (2016). *Empirical asset pricing : the cross section of stock returns*. John Wiley & Sons, Inc.

Baltussen, G., Swinkels, L., & Van Vliet, P. (2021). Global factor premiums. *Journal of Financial Economics*, *142*(3), 1128–1154. https://doi.org/10.1016/j.jfineco.2021.06.030

Balvers, R. J., & Wu, Y. (2006). Momentum and mean reversion across national equity markets. *Journal of Empirical Finance*, *13*(1), 24–48. https://doi.org/10.1016/j.jempfin.2005.05.001

Bekaert, G., & Harvey, C. R. (1997). Emerging equity market volatility. *Journal of Financial Economics*, *43*(1), 29–77. https://doi.org/10.1016/s0304-405x(96)00889-6

Bessembinder, H., Cooper, M. J., & Zhang, F. (2018). Characteristic-Based Benchmark Returns and Corporate Events. *The Review of Financial Studies*, *32*(1), 75–125. https://doi.org/10.1093/rfs/hhy037

Bhojraj, S., & Swaminathan, B. (2006). Macromomentum: Returns Predictability in International Equity Indices. *The Journal of Business*, *79*(1), 429–451. https://doi.org/10.1086/497416

Bogle, J. C., & Nolan, M. W. (2015). Occam's Razor Redux: Establishing Reasonable Expectations for Financial Market Returns. *The Journal of Portfolio Management*, *42*(1), 119–134. https://doi.org/10.3905/jpm.2015.42.1.119

Boudoukh, J., Israel, R., & Richardson, M. (2021). Biases in long-horizon predictive regressions. *Journal of Financial Economics*, *145*(3). https://doi.org/10.1016/j.jfineco.2021.09.013

Campbell, J. Y., & Thompson, S. B. (2007). Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? *Review of Financial Studies*, *21*(4), 1509–1531. https://doi.org/10.1093/rfs/hhm055

Cederburg, S., Johnson, T. L., & O'Doherty, M. S. (2022). On the Economic Significance of Stock Return Predictability. *Review of Finance*, *27*(2). https://doi.org/10.1093/rof/rfac035

Cenesizoglu, T., & Timmermann, A. (2012). Do return prediction models add economic value? *Journal of Banking & Finance*, *36*(11), 2974–2987. https://doi.org/10.1016/j.jbankfin.2012.06.008

Charles, A., Darné, O., & Kim, J. H. (2017). International stock return predictability: Evidence from new statistical tests. *International Review of Financial Analysis*, *54*, 97–113. https://doi.org/10.1016/j.irfa.2016.06.005

Chordia, T., Goyal, A., & Shanken, J. A. (2015). Cross-Sectional Asset Pricing with Individual Stocks: Betas versus Characteristics. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2549578

Clark, T. E., & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, *138*(1), 291–311. https://doi.org/10.1016/j.jeconom.2006.05.023

Diebold, F. X., & Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, *13*(3), 253–263. https://doi.org/10.1080/07350015.1995.10524599

Dimic, N., Orlov, V., & Piljak, V. (2015). The political risk factor in emerging, frontier, and developed stock markets. *Finance Research Letters*, *15*, 239–245. https://doi.org/10.1016/j.frl.2015.10.007

Donadelli, M., & Paradiso, A. (2014). Is there heterogeneity in financial integration dynamics? Evidence from country and industry emerging market equity indexes. *Journal of International Financial Markets, Institutions and Money*, *32*, 184–218. https://doi.org/10.1016/j.intfin.2014.06.003

Dow, C. H. (1920). Scientific stock speculation. *The Magazine of Wall Street*.

Erb, C. B., Harvey, C. R., & Viskanta, T. E. (1995). Country Risk and Global Equity Selection. *The Journal of Portfolio Management*, *21*(2), 74–83. https://doi.org/10.3905/jpm.1995.409504

Faber, M. (2012). Global value: Building trading models with the 10 year CAPE. *Cambria Quantitative Research*, No. 5. Retrieved from https://ssrn.com/abstract=2129474

Fama, E. F. (1976). *Foundations of Finance : Portfolio Decisions and Securities Prices*. Basic Books.

Fama, E. F., & French, K. R. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics*, *22*(1), 3–25. https://doi.org/10.1016/0304-405x(88)90020-7

Fama, E. F., & French, K. R. (2002). The Equity Premium. *The Journal of Finance*, *57*(2), 637–659. https://doi.org/10.1111/1540-6261.00437

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010

Fama, E. F., & French, K. R. (2019). Comparing Cross-Section and Time-Series Factor Models. *The Review of Financial Studies*, *33*(5). https://doi.org/10.1093/rfs/hhz089

Fama, E. F., & MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, *81*(3), 607–636.

Ferreira, M. A., & Santa-Clara, P. (2011). Forecasting stock market returns: The sum of the parts is more than the whole. *Journal of Financial Economics*, *100*(3), 514–537. https://doi.org/10.1016/j.jfineco.2011.02.003

Frankel, J. (2010). Monetary Policy in Emerging Markets. *Handbook of Monetary Economics*, *3*, 1439–1520. https://doi.org/10.1016/b978-0-444-53454-5.00013-x

Frazzini, A., & Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, *111*(1), 1–25. https://doi.org/10.1016/j.jfineco.2013.10.005

Gargano, A., Pettenuzzo, D., & Timmermann, A. (2019). Bond Return Predictability: Economic Value and Links to the Macroeconomy. *Management Science*, *65*(2), 508–540. https://doi.org/10.1287/mnsc.2017.2829

Golez, B., & Koudijs, P. (2018). Four centuries of return predictability. *Journal of Financial Economics*, *127*(2), 248–263. https://doi.org/10.1016/j.jfineco.2017.12.007

Green, J., Hand, J. R. M., & Zhang, X. F. (2017). The Characteristics that Provide Independent Information about Average U.S. Monthly Stock Returns. *The Review of Financial Studies*, *30*(12), 4389–4436. https://doi.org/10.1093/rfs/hhx019

Harvey, C. R. (1995). Predictable Risk and Returns in Emerging Markets. *The Review of Financial Studies*, *8*(3), 773–816. http://www.jstor.org/stable/2962239

Haugen, R. A., & Baker, N. L. (1996). Commonality in the determinants of expected stock returns. *Journal of Financial Economics*, *41*(3), 401–439. https://doi.org/10.1016/0304-405X(95)00868-F

Heston, S. L., & Sadka, R. (2010). Seasonality in the Cross Section of Stock Returns: The International Evidence. *Journal of Financial and Quantitative Analysis*, *45*(5), 1133–1160. https://doi.org/10.1017/s0022109010000451

Hjalmarsson, E. (2009). Predicting Global Stock Returns. *Journal of Financial and Quantitative Analysis*, *45*(01), 49. https://doi.org/10.1017/s0022109009990469

Hollstein, F., Prokopczuk, M., Tharann, B., & Simen, C. W. (2024). Predicting the equity premium around the globe: Comprehensive evidence from a large sample. *International Journal of Forecasting*. Advance online publication. https://doi.org/10.1016/j.ijforecast.2024.05.002

Hou, K., Xue, C., & Zhang, L. (2015). Digesting Anomalies: An Investment Approach. *Review of Financial Studies*, *28*(3), 650–705. https://doi.org/10.1093/rfs/hhu068

Ibbotson, R. G., & Chen, P. (2003). Long-Run Stock Returns: Participating in the Real Economy. *Financial Analysts Journal*, *59*(1), 88–98. https://doi.org/10.2469/faj.v59.n1.2505

Inoue, A., & Kilian, L. (2005). In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use? *Econometric Reviews*, *23*(4), 371–402. https://doi.org/10.1081/etc-200040785

Jordan, S. J., Vivian, A. J., & Wohar, M. E. (2014). Forecasting returns: New European evidence. *Journal of Empirical Finance*, *26*, 76–95. https://doi.org/10.1016/j.jempfin.2014.02.001

Jordan, S. J., Vivian, A., & Wohar, M. E. (2015). Location, location, location: currency effects and return predictability? *Applied Economics*, *47*(18), 1883–1898. https://doi.org/10.1080/00036846.2014.1000537

Kandel, S., & Stambaugh, R. F. (1996). On the Predictability of Stock Returns: An Asset-Allocation Perspective. *The Journal of Finance*, *51*(2), 385–424. https://doi.org/10.1111/j.1540-6261.1996.tb02689.x

Karolyi, G. A., & Van Nieuwerburgh, S. (2020). New Methods for the Cross-Section of Returns. *The Review of Financial Studies*, *33*(5), 1879–1890. https://doi.org/10.1093/rfs/hhaa019

Kasa, K. (1992). Common stochastic trends in international stock markets. *Journal of Monetary Economics*, *29*(1), 95–124. https://doi.org/10.1016/0304-3932(92)90025-w

Kelly, B. T., Pruitt, S., & Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*, *134*(3), 501–524. https://doi.org/10.1016/j.jfineco.2019.05.001

Keloharju, M., Linnmaimaa, J. T., & Nyberg, P. (2016). Return Seasonalities. *The Journal of Finance*, *71*(4), 1557–1590. https://doi.org/10.1111/jofi.12398

Keppler, A. M., & Traub, H. D. (1993). The Small-Country Effect. *The Journal of Investing*, *2*(3), 17–24. https://doi.org/10.3905/joi.2.3.17

Kim, D. (2012). Value Premium Across Countries. *The Journal of Portfolio Management*, *38*(4), 75–86. https://doi.org/10.3905/jpm.2012.38.4.075

Lazarus, E., Lewis, D. J., Stock, J. H., & Watson, M. W. (2018). HAR Inference: Recommendations for Practice. *Journal of Business & Economic Statistics*, *36*(4), 541–559. https://doi.org/10.1080/07350015.2018.1506926

Lewellen, J. (2015). The Cross-section of Expected Stock Returns. *Critical Finance Review*, *4*(1), 1–44. https://doi.org/10.1561/104.00000024

Li, T., & Pritamani, M. (2015). Country Size and Country Momentum Effects in Emerging and Frontier Markets. *The Journal of Investing*, *24*(1), 102–108. https://doi.org/10.3905/joi.2015.24.1.102

Merton, R. C. (1973). An Intertemporal Capital Asset Pricing Model. *Econometrica*, *41*(5), 867. https://doi.org/10.2307/1913811

MSCI. (2021, June 24). *MSCI 2021 Market Classification Review* [Press release]. https://www.msci.com/documents/10199/674fb092-7c0c-aa7d-d94d-de59e9a4996f

MSCI. (2022). *Market classification*. Www.msci.com. https://www.msci.com/our-solutions/indexes/market-classification

MSCI. (2024). *MSCI Emerging Markets Index* [Fact sheet]. MSCI Inc. Retrieved from https://www.msci.com/documents/10199/b33f4f52-6d5c-4c8a-8ad8-be161c60cb86

Narayan, P. K., Narayan, S., & Thuraisamy, K. S. (2014). Can institutions and macroeconomic factors predict stock returns in emerging markets? *Emerging Markets Review*, *19*, 77–95. https://doi.org/10.1016/j.ememar.2014.04.005

Nazaire, G., Pacurar, M., & Sy, O. (2020). Betas versus characteristics: A practical perspective. *European Financial Management*, *26*(5), 1385–1413. https://doi.org/10.1111/eufm.12263

Pástor, Ľ., Stambaugh, R. F., & Taylor, L. A. (2017). Do Funds Make More When They Trade More? *The Journal of Finance*, *72*(4), 1483–1528. https://doi.org/10.1111/jofi.12509

Paye, B. S., & Timmermann, A. (2006). Instability of return prediction models. *Journal of Empirical Finance*, *13*(3), 274–315. https://doi.org/10.1016/j.jempfin.2005.11.001

Rapach, D., & Zhou, G. (2013). Forecasting Stock Returns. *Handbook of Economic Forecasting*, 328–383. https://doi.org/10.1016/b978-0-444-53683-9.00006-2

Richards, A. J. (1997). Winner-Loser Reversals in National Stock Market Indices: Can They be Explained? *The Journal of Finance*, *52*(5), 2129–2144. https://doi.org/10.1111/j.1540-6261.1997.tb02755.x

Richards, A. J. (1995). Comovements in national stock market returns: Evidence of predictability, but not cointegration. *Journal of Monetary Economics*, *36*(3), 631–654. https://doi.org/10.1016/0304-3932(95)01225-7

Sakkas, A., & Tessaromatis, N. (2022). Forecasting the Long-Term Equity Premium for Asset Allocation. *Financial Analysts Journal*, *78*(3), 1–21. https://doi.org/10.1080/0015198x.2022.2073782

Spierdijk, L., Bikker, J. A., & van den Hoek, P. (2012). Mean reversion in international stock markets: An empirical analysis of the 20th century. *Journal of International Money and Finance*, *31*(2), 228–249. https://doi.org/10.1016/j.jimonfin.2011.11.008

Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics*, *54*(3), 375–421. https://doi.org/10.1016/s0304-405x(99)00041-0

Timmermann, A. (2018). Forecasting Methods in Finance. *Annual Review of Financial Economics*, *10*(1), 449–479. https://doi.org/10.1146/annurev-financial-110217-022713

Welch, I., & Goyal, A. (2008). A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. *Review of Financial Studies*, *21*(4), 1455–1508. https://doi.org/10.1093/rfs/hhm014

Zaremba, A. (2016). Risk-based explanation for the country-level size and value effects. *Finance Research Letters*, *18*, 226–233. https://doi.org/10.1016/j.frl.2016.04.020

Zaremba, A. (2019). The Cross Section of Country Equity Returns: A Review of Empirical Literature. *Journal of Risk and Financial Management*, *12*(4), 165. https://doi.org/10.3390/jrfm12040165

Zaremba, A., & Mehmet Umutlu. (2018). Size matters everywhere: Decomposing the small country and small industry premia. *The North American Journal of Economics and Finance*, *43*(C), 1–18. https://doi.org/10.1016/j.najef.2017.09.002

Appendix A Descriptive Statistics

Table 6

Descriptive Statistics

Panel A: Mean

Countries	Equity Market Return	DY	PE	STIR	TS	MVOL
Greece	-0.001	2.577	15.506	0.043	0.024	0.016
South Africa	0.013	3.189	14.927	0.076	0.017	0.011
Hungary	0.009	2.502	15.392	0.057	0.004	0.013
India	0.01	1.489	18.459	0.067	0.012	0.012
Thailand	0.01	2.924	14.085	0.021	0.017	0.012
Philippines	0.008	1.865	17.658	0.042	0.037	0.01
Malaysia	0.007	3.02	16.265	0.028	0.013	0.007
Taiwan	0.008	3.192	17.311	0.012	0.008	0.012
Pakistan	0.011	5.634	10.05	0.087	0.015	0.012

Panel B: Standard Deviation

Countries	Equity Market Return	DY	PE	STIR	TS	MVOL
Greece	0.086	1.091	6.582	0.054	0.056	0.009
South Africa	0.049	0.58	3.082	0.024	0.018	0.005
Hungary	0.07	0.967	7.629	0.043	0.023	0.007
India	0.001	0.367	4.759	0.019	0.011	0.007
Thailand	0.07	1.061	4.213	0.011	0.012	0.007

Philippines	0.058	0.614	3.493	0.029	0.019	0.005
Malaysia	0.045	0.646	2.997	0.005	0.009	0.004
Taiwan	0.064	1.279	5.69	0.012	0.005	0.006
Pakistan	0.185	1.317	2.507	0.031	0.025	0.007

Note. The table presents means (Panel A) and standard deviations (Panel B) for country state variables and returns. The state variables are the dividend yield (DY), the price-earnings ratio (PE), the short-term interest Rate (STIR), the term spread (TS), and the market volatility (MVOL). The equity market return is the one-month return. The sample period is from February 1999 to May 2022 for all variables. All variables are presented at a monthly frequency. Note that MVOL is presented monthly, but calculated using the daily returns within a month, using the daily country equity market return.

Appendix B Variable Sources

Table 7

Datastream Variables

Country	Total Return Index of the equity market index (RI)	Dividend yield of the equity market index (DY)	Price to Earning (PE) ratio of the equity market index
Greece	TOTMKGR(RI)	TOTMKGR(DY)	TOTMKGR(PE)
South Africa	TOTMKSA(RI)	TOTMKSA(DY)	TOTMKSA(PE)
Hungary	TOTMKHN(RI)	TOTMKHN(DY)	TOTMKHN(PE)
India	TOTMKIN(RI)	TOTMKIN(DY)	TOTMKIN(PE)
Thailand	TOTMKTH(RI)	TOTMKTH(DY)	TOTMKTH(PE)
Philippines	TOTMKPH(RI)	TOTMKPH(DY)	TOTMKPH(PE)
Malaysia	TOTMKMY(RI)	TOTMKMY(DY)	TOTMKMY(PE)
Taiwan	TOTMKMY(TA)	TOTMKMY(TA)	TOTMKMY(TA)
Pakistan	TOTMKMY(PK)	TOTMKMY(PK)	TOTMKMY(PK)

Note. Displayed are all variables obtained from Refinitiv Datastream for all countries in the analysis, including the respective codes. These are for the state variables dividend yield (DY) and price-to-earnings (PE) ratio. Monthly market volatility for month t is computed as the standard deviation of the daily equity market total return index (RI) returns over month t. The returns used for equity premiums are in column 2 as the total return index of the equity market index (RI). The equity risk premium is calculated as the difference between the country returns computed using this index, and the short-term interest rate proxied by the 3-month treasury bill yield (more details in Table 8 below).

Table 8

Finaeon Global Financial Data Variables

Country	10-Year Government Bond Yield	3-Month Treasury Bill Yield

Greece	IGGRC10D	ITGRC3D
South Africa	IGZAF10D	ITZAF3D
Hungary	IGHUN10D	ITHUN3D
India	IGIND10D	ITIND3D
Thailand	IGTHA10D	ITTHA3D
Philippines	IGPHL10D	ITPHL3D
Malaysia	IGMYS10D	ITMYS3D
Taiwan	IGTWN10D	ITTWN3D
Pakistan	IGPAK10D	ITPAK3D

Note. Displayed are all variables obtained from Finaeon Global Financial Data for all countries in the analysis, including the symbols. These are for the state variables short-term interest rate (STIR) and term spread (TS). The term spread is defined as the difference between the long-term and short-term interest rate; in this case it is proxied by the difference between the 10-Year government bond yield and the 3-month treasury bill yield.