US Risk Premia under Emerging Markets Constraints

Elias Cavalcante-Filho, Fernando Chague, Rodrigo De-Losso, Bruno Giovannetti*

Área ANPEC: 08 Microeconmia, Métodos Quantitativos e Finanças

Resumo

O mercado dos EUA é a referência para as finanças empíricas e é considerado o exemplo mais próximo de um mercado eficiente. Por outro lado, resultados divergentes dos observados nos EUA estão frequentemente considerados não confiáveis devido aos possíveis desvios das hipóteses de eficiência de mercado. No entanto, como os resultados do mercado norte-americano se comportariam se os dados fossem submetidos as mesmas restrições de uma economia emergente? Para responder a essa pergunta, analisamos a estimativa de prêmios de risco sob as restrições típicas dos mercados de ações emergentes: o baixo número de ativos e curto histórico de dados. Utilizamos como referência os parâmetros de tamanho da amostra, número de ativos e distribuição de variáveis contábeis do mercado acionário brasileiro. Surpreendentemente, concluímos que os prêmios de risco dos EUA apresentam as mesmas características que as obtidas com dados brasileiros quando sob as mesmas restrições de tempo. Em seguida, avaliamos as duas possíveis causas decorrentes do T pequeno: i) viés de pequenas amostras nos betas e ii) divergência entre os prêmios de risco *ex-post* e *ex-ante*. Através de simulações de Monte Carlo, concluímos que com T de aproximadamente 5 anos as estimativas dos beta não são mais um problema. No entanto, é necessário analisar uma série temporal superior a 40 anos para se obter um prêmio de risco *ex-ante* robusto.

Palavras-chaves: Prêmios de risco, Precificação de ativos, Modelos multifatorias.

Códigos JEL: G12, G17.

Abstract

USA market is the benchmark for empirical finance and considered the closest example of how an efficient market should behave. On the other hand, divergent results from the observed in the USA are often associated with unreliable and due deviations from efficient hypothesis. However, how would the US market results behave had the data the same constraints as an emerging market economy? To answer that question we analyze the risk premia market estimation under the typical constraints from emerging equity markets: the small number of assets and the short time-series sample available for estimation. We use parameters of time-series length, number of assets and accounting variables distribution from the Brazilian equity market. Surprisingly, we conclude that the US market risk premia convey the same data features as the Brazilian risk premia if under the same time constraints. Then, we evaluate two potential causes of problems in risk premia estimations with small T: i) small sample bias on betas, and ii) divergence between *ex-post* and *ex-ante* risk premia. Through Monte Carlo simulations, we conclude that for the T around 5 years the beta estimates are no longer a problem. However, it is necessary to analyze a time-series sample exceeding 40 years to obtain robust *ex-ante* risk premia.

Keywords: Equity Risk premia, Asset pricing, Multi-factor model.

JEL Codes: G12, G17.

*E-mail addresses: e.cavalcante@usp.br (E. Cavalcante-Filho), fernando.chague@fgv.br (F. Chague), delosso@usp.br (R. De-Losso), bruno.giovannetti@fgv.br (B. Giovannetti).

1 Introduction

The estimation results of emerging markets risk premia show no consensus in the literature.[2] The estimation results of Brazilian risk premia, for instance, not only disagree with the US and developed markets results but also vary among themselves. The Brazilian risk premia estimates collected from several studies[3] and presented in Figure 1 illustrate this lack of consensus.

The estimates in Figure 1 are obtained by exploring many sub-periods between 1976 and 2015 and by applying a variety of techniques[4]. The estimates of the market (*Mkt*) risk premium are positive in 41 cases, negative in 18 cases, and not significant in 74 cases. The size (*SMB*) risk premium does not present any positive and statistically significant estimates. The value (*HML*) risk premium presents the most robust estimates. Of the 88 estimates, 46 are positive and significant, and only 2 are significant and negative. Finally, there are 26 estimates for the momentum (*WML*) risk premium, but only a few cases are significant, with three positives and one negative.

This lack of consensus is normally attributed to deviations from efficient hypothesis or imperfections on the emerging markets. In light of these circumstances, we turn to the US data and estimate risk premia under the restrictions faced by a typical emerging stock market. We chose the US since the explanatory power of risk factors in this market is widely validated and accepted by the literature and we apply the restrictions observed for the Brazilian market as reference for a typical emerging market.

First, we analyze the sensitivity of the US risk premia estimations on two relevant constraints present in emerging stock market: the small number of assets (small N) and the short time-series samples (small T).[5,6]

[2] Grandes et al. (2010) find that the size and value premiums are not statistically significant in the stock markets of Latin America. Chui and Wei (1998) find in Pacific-Basin emerging markets (Hong Kong, Korea, Malaysia, Taiwan, and Thailand) a weak relationship between average returns and beta, a value premium in three of the markets, and size premium in four of the markets. Barry et al. (2002) investigate the size and value effects in 35 emerging markets and find evidence of a value premium, but not robust evidence for the size premium. Cakici et al. (2013) examine 18 emerging stock markets from Asia, Latin America, and Eastern Europe and find strong evidence for the value effect in the whole sample, momentum effect for all markets but Eastern Europe, and very weak evidence for size effect. Bauman et al. (1998) find evidence of both the size and value premiums for a range of Asian stock markets. Foye et al. (2013) test the three-factor model in the European Union markets and find a value premium, but no size premium. On the other hand, Lischewski and Voronkova (2012) examine only Polish stocks and report that the market, size, and value factors have explanatory power on average returns, but do not explain portfolio returns fully.

[3] Fama and French (1998); Rouwenhorst (1999); Bonomo and Garcia (2001); Bonomo et al. (2002); Sampaio (2002); Malaga and Securato (2004); Matos (2006); Chague and De-losso (2007); Bellizia (2009); Mussa et al. (2009); Brito and Murakoshi (2009); Grandes et al. (2010); Mussa et al. (2011); Bodur (2011); Rizzi (2012); Mussa et al. (2012); Varga and Brito (2015); Eid Jr and Martins (2015), and Piccoli et al. (2015).

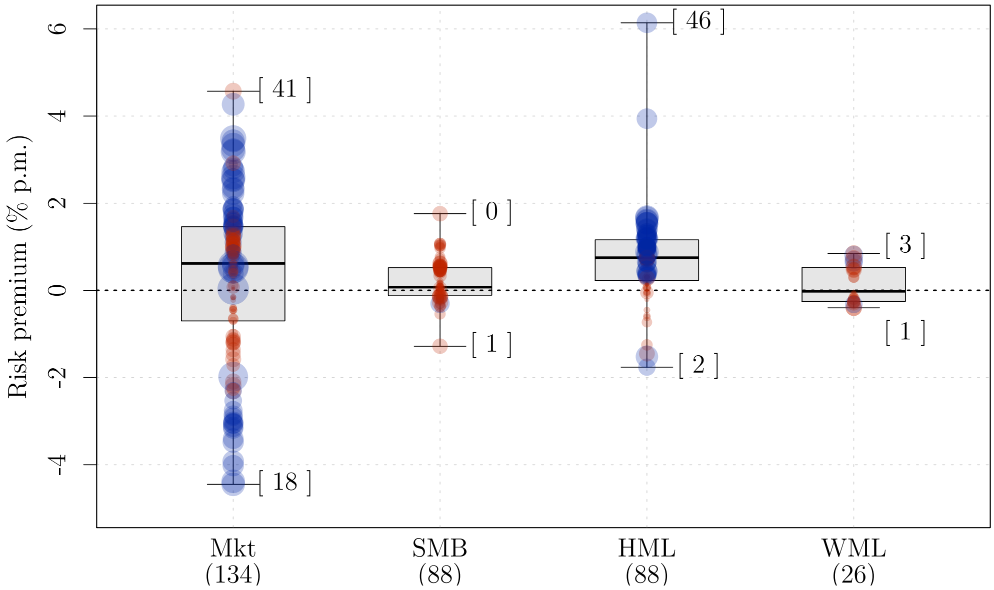
[4] Sample means, (Fama and MacBeth, 1973), Generalized Method of Moments (GMM), and Iterative Nonlinear Seemingly Unrelated Regression Estimation (ITNLSUR).

[5] There are few liquid stocks in Brazil. In 2000, only 37 stocks could be considered liquid. In 2014, this number increased to 137. The details on the liquidity criteria can be found at the Brazilian Center for Research in Financial Economics of the University of São Paulo (NEFIN) website – Núcleo de Pesquisas em Economia Financeira da Universidade de São Paulo <http://nefin.com.br>.

[6] Until 1999, the risk-free rate was used as an instrument of the pegged exchange rate regime, and it

Figure 1: Dispersion in the risk premia estimations for Brazil

The figure shows the dispersion in the estimations of market (*Mkt*), size (*SMB*), value (*HML*), and momentum (*WML*) risk premia in the Brazilian stock market. The figure shows a box-plot for each risk factor, and each circle represents one reported estimation (% p.m.). The box-plot reports the 0^{th}, 25^{th}, 50^{th}, 75^{th}, and 100^{th} percentiles. The numbers within the brackets to the right of the 0^{th} and 100^{th} percentiles are the number of negative and positive estimates of each factor respectively. The higher the t-value reported, the larger the circle. A significant estimate ($|\text{t-value}| \geq 1.64$) is indicated by a blue circle. The coordinate axis reports the factor's name followed by the number of estimates reported.



We conclude that the restriction imposed by the small T is more relevant than that imposed by the small N. While Brazilian data offer values of T over 14 years, our analysis indicates that it is necessary to analyze a time-series sample using data exceeding 40 years to obtain robust risk premia estimates. On the other hand, values of N as observed for the Brazilian market do not pose an issue.

Given these results, we then investigate the problems caused by the small T. One problem could be the use of poorly estimated betas in the second stage of the estimation. Another problem could be the use of poorly estimated expected returns of stocks. Both would induce errors in the estimation of the risk premia. Poorly estimated betas generate biased risk premia estimates. Poorly estimated expected returns lead to the estimation of *ex-post* instead of *ex-ante* risk premia.[7]

To assess the relative importance of these two issues, we perform Monte Carlo simulations. We conclude that the most important issue is the use of poorly estimated expected

was often set to very high levels, and before 1994 the Brazilian economy has faced a strong inflationary period. The results obtained after including data from the period between 1994 and 1999 remain the same and are presented in Appendix A, which can be provided under requested. Therefore, to estimate the risk premia in Brazil, one commonly uses data beginning in the year 2000.

[7]The *ex-ante* risk premium is the excess return that the investor expects to receive when investing in the asset. This is the value to be estimated. The *ex-post* risk premium is the realized excess return after shocks.

returns. Indeed, the difference between *ex-post* and *ex-ante* risk premia proves to be significant when the estimation is performed with small T. For instance, when data are simulated using a market risk premia of 0.65% p.m., we estimate a positive and significant risk premium in only 17% of the samples if T is set at 14.

To the best of our knowledge nobody has restricted the US data to understand how they behave. By adopting such a procedure, we show that data behavior of emerging market economies are not that awkward. By contrast, we are able to show that the US data behavior is not robust by shortening the sample. We also conclude that it takes many years (roughly 40 years) to safely estimate the *ex-ante* risk premia. Therefore, anyone interested in computing the cost of equity for firms in emerging economies should use the local data only to estimate the betas. In turn, the price of risk should be taken from data with longer time-series sample, US data, for instance.

The remainder of the paper is organized as follows. Section 2 describes the dataset, followed by the first analysis of portfolios and risk factors. Section 3 presents the factor model we use in the estimation as well as the methodology supporting it. Section 4 compares the premia results for the US with those found for Brazilian markets (4.1), followed by the analysis of the impact of estimating risk premia with a small number of assets (4.2) and with small time-series samples (4.3). This section ends with the assessment of the consequences of estimating risk premia with small time-series samples (4.4). Section 5 concludes the paper by summarizing the findings and its implications.

2 Data

The main paper's datasets consist of monthly portfolio returns from and risk factors of the Brazilian and the US stock markets. The US information is obtained from French's website[8] and covers the period between January 1927 and December 2014. The Brazilian information is taken from the Brazilian Center for Research in Financial Economics of the University of São Paulo (NEFIN) website and covers the period from January 2001 to December 2014. The Brazilian stock market was already operational prior to this period. However, until 1999, the risk-free rate was used as an instrument of the pegged exchange rate regime, and it was often set to very high levels. Therefore, to estimate the risk premia in Brazil, one commonly uses data beginning in the year 2000. The US portfolio and risk factor returns are value-weighted, while the Brazilian ones are equally weighted[9].

The 22 portfolios used in this work are presented in Table 1. This table is organized as follows. The first column shows the variables used to arrange the assets into portfolios. The second column names each portfolio, the third and fourth columns contain the percentiles used as breakpoints for the construction of the portfolios, the fifth and sixth columns report the means of the portfolios' returns, and, finally, the last two columns list the autocorrelation consistent standard deviations. The data available at NEFIN's and French's websites have different numbers of portfolios. Therefore, in order to conduct a fair comparison between the Brazilian and the US markets, some portfolios are combined by calculating their value-weighted returns. The first 17 portfolios are defined based on information about Size, Book-to-market, and Momentum, and the other portfolios are

[8] `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

[9] We test both measures for the Brazilian market, and they remain mostly unchanged.

organized by grouping assets of the same industry. Note that both countries have the same number of industry portfolios; however, some industries appear only in one of the markets.

Table 1: US and Brazilian portfolios

Portfolios of the the US and Brazilian asset market. The portfolios differ between countries by their breakpoints and the period contemplated. Their means were calculated on the excess of return and the standard deviation are autocorrelation consistent.

Variables	Labels	Breakpoints US Jan/1927 - Dec/2014 (1056 months)	Breakpoints BRA Jan/2001 - Dec/2014 (168 months)	Mean (% p.m.) US	Mean (% p.m.) BR	Sd (% p.m.) US	Sd (% p.m.) BR
Size	Small	[0,30]	[0,33.3]	1.00***	0.25	8.44	8.01
	Medium size	[30,70]	[33.3,66.6]	0.88***	0.25	6.78	6.99
	Big	[70,100]	[66.6,100]	0.63***	0.12	5.25	6.20
Book-to-market	Low	[0,10]	[0,33.3]	0.58***	0.10	5.77	6.74
	Medium bm	[30,70]	[33.3,66.6]	0.72***	0.13	5.78	6.99
	High	[90,100]	[66.6,100]	1.09***	0.49	9.22	7.17
Momentum	Loser	[0,30]	[0,33.3]	0.37	-0.42	7.71	8.61
	Normal	[30,70]	[33.3,66.6]	0.63***	0.44	5.62	6.29
	Winner	[70,100]	[66.6,100]	0.98***	0.71	5.61	6.36
Size x Book-to-market	Small Low	[0,50 ; 0,20]	[0,50 ; 0,50]	0.65***	0.24	8.09	7.70
	Small High	[0,50 ; 80,100]	[0,50 ; 50,100]	1.27***	0.47	8.77	8.06
	Big Low	[50,100 ; 0,20]	[50,100 ; 0,50]	0.62***	0.15	5.47	6.21
	Big High	[50,100 ; 80,100]	[50,100 ; 50,100]	0.99***	0.08	8.12	6.77
Size x Momentum	Small Loser	[0,50 ; 0,30]	[0,50 ; 0,50]	0.55*	-0.14	9.21	8.88
	Small Winner	[0,50 ; 70,100]	[0,50 ; 50,100]	1.35***	0.84	7.30	7.08
	Big Loser	[50,100 ; 0,30]	[50,100 ; 0,50]	0.38	-0.08	7.63	7.25
	Big Winner	[50,100 ; 70,100]	[50,100 ; 50,100]	0.94***	0.42	5.56	5.79
Industry	Basic Products	-	Basic Products		0.48		7.83
	Consumer	Consumer	Consumer	0.72***	-0.09	5.35	6.81
	Energy	-	Energy		0.26		7.13
	HiTec	HiTec	-	0.67***	-	5.64	-
	Healthcare	Healthcare	-	0.81***	-	5.63	-
	Manufacturing	Manufacturing	Manufacturing	0.69***	0.88	5.59	8.70
	Other	Other	Other + Finance	0.63***	0.40	6.50	7.39

Significance: * 10%; ** 5%; ***1%.

Table 1 shows that the US portfolios' returns are negatively correlated with Size and positively correlated with Book-to-market and Momentum, as described in the literature (Fama and French, 1992, 1993; Carhart, 1997). The Brazilian portfolios also generally exhibit the same behavior as those of the US market. However, three points require attention: i) the Book-to-market effect is not observable among the Big High and Big Low portfolios, since Big High presents an average return of 0.08, which is smaller than the average return of 0.15 for Big Low, ii) the Size effect is also not observed between the Big Loser and Small Loser portfolios, since the Small Loser has lower average return, and iii) none of the Brazilian portfolios' returns have a significant average.

Besides the mentioned portfolios, we use four risk factors and a risk-free rate for each economy. Table 2 presents the mean and standard deviation of those variables. The names used for each factor are provided in the first column. They are *Mkt* for the Market factor, *SMB* for the Size factor, *HML* for Book-to-market, and *WML* for Momentum. The risk-free rate used for the US is the Treasury bill month rate, and for the Brazilian market, we use the 30-day Deposito Interbancário (DI) swap rate.

Again the variables for the US follow the pattern documented in the literature, that is, all factors have positive and statistically significant return averages. On the other hand,

Table 2: Factors and risk-free rate

Factors and risk-free rates of the US and Brazilian markets. The information refers to the period between January 1927 and December 2014 for the American market and between January 2001 and December 2014 for the Brazilian market. The table present the means for each factor and the respective autocorrelation consistent standard deviation.

Factor	**Mean (% p.m.)**		**Sd (% p.m.)**	
	US	**BR**	**US**	**BR**
Mkt	0.6502***	0.2411	5.41	6.18
SMB	0.2357**	0.0147	3.23	4.81
HML	0.3973***	0.4460	3.54	4.52
WML	0.6755***	1.2493***	4.74	5.50
risk-free	0.2840***	1.0655***	0.25	0.35

Significance: * 10%; ** 5%; ***1%.

the Brazilian data show significance only for *WML*, despite the fact that all factors also have positive averages.

3 Methodology of risk premium estimation

This section explains how the risk premium is estimated. In order to do so, we first present the multi-factor model in subsection 3.1, and then explain the methodology for the model estimation.

3.1 Multi-factor model

Define K as the number of risk factors, N as the number of assets, and T as the number of observed periods. The multi-factor model assumes that excess asset returns are governed by the following linear relation:

$$E(R_i^e) = \alpha + \beta_i' \lambda \tag{1}$$

where R_i^e is the excess return of an asset $i \in \{1, 2, ..., N\}$, α is the model pricing error, λ is a Kx1 vector with the risk premia for the K factors, and β_i is a Kx1 vector with the risk measures of asset i for each factor.

The model also proposes that the β_i vector respect the following relation in the time series:

$$R_{it}^e = a_i + \beta_i' f_t + \epsilon_{it} \tag{2}$$

where R_{it}^e is the excess return of asset i in period $t \in \{1, 2, ..., T\}$, a_i is the expected pricing error of asset i, f_t is a $Kx1$ vector with the realizations of the factors in period t, and ϵ_{it} is the random error of asset i in period t.

3.2 Risk premium estimation

The model estimation is conducted using the GMM methodology introduced by Hansen (1982), with a similar framework proposed by Cochrane (2001). This methodology provides a joint estimation of all parameters of the model and easily handles the problems of serial correlation and conditional heteroscedasticity.

The GMM estimation is based on the hypotheses derived from the model's introduction in Section 3.1. We can build the following matrix using the model's assumptions:

$$g_t(a, \beta, \alpha, \lambda) = \begin{bmatrix} (R_t^e - a - \beta f_t) \\ [(R_t^e - a - \beta f_t) \otimes f_t] \\ [\mathbf{1}, \beta'] \left[(R_t^e - \alpha - \beta \lambda) \right] \end{bmatrix}_{([N+NK+1+K]\text{x}1)} \tag{3}$$

where R_t^e is the Nx1 vector of excess returns in period t, f_t is the Kx1 vector of risk factors in period t, such that $t \in \{1, 2, ..., T\}$, $\otimes$ is the Kronecker operator, a is the Nx1 vector of expected pricing errors for each asset, β is the NxK matrix of risk related to each asset and factors of the model, α is the expected pricing error of the model, λ is the Kx1 vector of risk premia for each risk factor, and $\mathbf{1}$ is a 1xN vector of ones.

From the model's assumptions, we know that the expected value of each line of the matrix equals zero. The GMM method estimates the parameters $(\hat{a}, \hat{\beta}, \hat{\alpha}, \hat{\lambda})$ by solving the following optimization:

$$\left\{ \hat{a}, \hat{\beta}, \hat{\alpha}, \hat{\lambda} \right\} = \underset{\{a,\beta,\alpha,\lambda\}}{\operatorname{argmin}} \sum_{t=1}^{T} \left[g_t(a, \beta, \alpha, \lambda) \right]' W^{-1} \sum_{t=1}^{T} \left[g_t(a, \beta, \alpha, \lambda) \right] \tag{4}$$

where W is a weighting matrix of moments, which is typically set to generate as effective estimates as possible. However, since the number of parameters and equations are the same, the weighting matrix does not affect the estimation. Thus, we use the identity matrix to solve the problem.

Finally, we have the following matrix of variance and covariance parameters:

$$Var(\hat{a}, \hat{\beta}, \hat{\alpha}, \hat{\lambda}) = (d'Wd)^{-1}(d'WSWd)(d'Wd)^{-1} \tag{5}$$

where d is the derivate of g_t with respect to the parameter vector ($d = E\left[\partial g_t / \partial(a, \beta, \alpha, \lambda)\right]$), W is the identity matrix ($[N + NK + 1 + K]$), and $S = E(g_t g_t')$ is estimated using the Parzen kernel with the same band as the entire value of $0.75T^{1/3}$.

4 Risk premium analysis

4.1 Full samples analysis

The risk premia diagnostic starts by comparing the US and the Brazilian stock market results. The *Mkt*, *SMB*, *HML*, and *WML* risk premia are estimated using data from January 1927 to December 2014 for the US stock market, and from January 2001 to December 2014 for the Brazilian one.

Table 3 presents the results, which are organized in two panels. Panel A presents the estimated parameters for the time series regression, with the risk measures of each portfolio. Panel B arranges the cross-section regression results, with the risk premia estimates followed by their p-values and standard error deviations. In both panels, the values on the left refer to the US, and on the right, to Brazil.

The results for the time series in Panel A show some similarities between both markets. As expected, the US market data confirm the patterns widely documented in the literature. Most values of the intercept a are not significant and, despite some significant cases, they do

Table 3: **Full sample regression**

The table presents the estimated values for the United States (US) and Brazil (BR) of the parameters of the time series regression (equation 2) in Panel A and cross-section regression (equation 1) in Panel B. The periods used for the estimates cover January 1927 to December 2014 to the United States, and January 2001 to December 2014 for Brazil.

Panel A: Time series regression										
$R^e_{it} = a_i + b_i Mkt_t + s_i SMB_t + h_i HML_t + m_i WML_t$										
	US					**BR**				
Portfolio	**a**	**b**	**s**	**h**	**m**	**a**	**b**	**s**	**h**	**m**
Small	-0.08**	1.04***	1.19***	0.43***	-0.05**	0.17	0.92***	0.86***	0.01	-0.14***
Medium size	-0.01	1.06***	0.56***	0.19***	0.00	0.13	0.93***	0.25***	0.15**	-0.14***
Big	0.02***	0.99***	-0.13***	-0.01	-0.01**	0.05	0.95***	-0.12***	0.00	-0.12***
Low	0.04	1.07***	-0.07***	-0.34***	-0.02	0.18	0.9***	0.33***	-0.4***	-0.09***
Medium b/m	-0.02	0.99***	-0.05***	0.31***	-0.02	0.11	0.95***	0.28***	0.04	-0.18***
High	-0.14	1.15***	0.51***	1.07***	-0.1***	0.12	0.91***	0.34***	0.59***	-0.09**
Loser	0.10***	1.07***	0.08***	0.04**	-0.67***	0.11	0.95***	0.38***	0.08*	-0.64***
Normal	0.06	0.98***	-0.09***	0.10***	-0.12***	0.37*	0.87***	0.19***	0.08*	-0.14***
Winner	-0.01	1.07***	0.06***	0.05***	0.39***	-0.05	0.97***	0.35***	0.09**	0.39***
Small low	-0.18***	1.12***	1.11***	-0.25***	-0.07***	0.37*	0.90***	0.65***	-0.35***	-0.17***
Small high	0.06*	1.05***	0.98***	0.86***	-0.06***	0.17	0.96***	0.71***	0.39***	-0.10***
Big low	0.09***	1.03***	-0.09***	-0.28***	-0.01	0.08	0.92***	0.00	-0.20***	-0.05
Big high	-0.10	1.14***	0.06**	0.94***	-0.06***	-0.08	0.90***	-0.08	0.42***	-0.20***
Small loser	-0.08	1.07***	1.00***	0.25***	-0.58***	0.20	0.95***	0.62***	0.14**	-0.51***
Small winner	0.10**	1.08***	0.94***	0.23***	0.36***	0.25	0.97***	0.57***	0.13**	0.23***
Big loser	0.14***	1.08***	-0.06***	0.03	-0.68***	0.29	0.89***	-0.15***	0.07	-0.49***
Big winner	-0.03	1.07***	0.00	0.05***	0.38***	-0.11	0.91***	0.13***	0.01	0.25***
Basic Products	-	-	-	-	-	0.22	0.76***	0.51***	-0.08	0.08
Consumer	0.13**	0.92***	0.01	-0.01	-0.01	0.09	0.83***	0.22***	-0.09*	-0.27***
Energy	-	-	-	-	-	-0.09	0.84***	0.17**	0.54***	-0.07
HiTec	0.21***	0.97***	0.04	-0.35***	-0.07***	-	-	-	-	-
Healthcare	0.31***	0.89***	-0.10*	-0.17***	0.02	-	-	-	-	-
Manufacturing	-0.03	0.99***	-0.10***	0.19***	0.04**	0.40	1.12***	0.4***	0.27***	0.06
Other	-0.14**	1.04***	0.07**	0.33***	-0.08***	0.27	1.02***	0.25***	-0.01	-0.1**
Panel B: Cross-section regression										
$E(R^e_i) = \alpha + b_i\lambda_{mkt} + s_i\lambda_{smb} + h_i\lambda_{hml} + m_i\lambda_{wml}$										
factor	α	λ_{mkt}	λ_{smb}	λ_{hml}	λ_{wml}	α	λ_{mkt}	λ_{smb}	λ_{hml}	λ_{wml}
estimate (% p.m.)	1.06***	-0.34	0.20**	0.33***	0.62***	-0.47	0.85	0.20	0.26	1.09**
p-value	0.001	0.334	0.046	0.003	0.000	0.636	0.481	0.611	0.495	0.019
$se(\lambda)$	(0.31)	(0.35)	(0.10)	(0.11)	(0.15)	(1.00)	(1.20)	(0.40)	(0.39)	(0.46)

Significance: * 10%; ** 5%; ***1%.

not show correlation with the variables Size, Book-to-market, and Momentum. The values of b are mostly around 1.00, with low variation between portfolios, and the parameters s, h, and m behave according to the ordering patterns reported for the portfolios returns. In other words, the lower the value of the assets that integrate the portfolios, the higher the estimated values of s; on the other hand, h and m grow positively correlated with portfolios ordered by the variables Book-to-market and Momentum, respectively.

In the Brazilian case, most of the patterns observed in the US time-series regression repeat themselves with a few caveats. First, the estimates of a show a lower frequency of significant cases, which goes in favor to the model's adjustment to the data . In addition, the parameters' estimates follow the same order highlighted in the results of the US data: s is negatively correlated with Size, and h and m are positively correlated with Book-to-market and Momentum respectively, despite the lower frequency of significant estimates than those observed for the US market.

In contrast, the cross-section results in Panel B indicate some divergence between Brazil's and the US' risk premia estimations. The results for the US show significant risk premia for all factors except the market factor, which supports the capacity of the model to fit the US data. On the other hand, the Brazilian results reject the model's ability to replicate the returns of assets, since only the factor *WML* shows a positive and significant risk premium.

The analysis of this result by itself could lead to the precipitated conclusion that the multi-factor model does not fit Brazilian data. However, this result should be interpreted with caution. Therefore, in order to identify the source of the problem, we conduct a more careful analysis of two points that distinguish both markets: i) the size of N, that is, the number of assets used to build the portfolios in the estimation process, and ii) the size of T, that is, the length of the time-series sample of the Brazilian data. The next step is to investigate the impact of these divergences.

4.2 Does the size of N restrict the estimation?

There is a huge difference between the number of assets available in the US and Brazilian markets. Table 4 reports the number of eligible assets[10] for both markets and organizes this information in sub-periods[11].

Table 4: Number of eligible assets

The table indicates the maximum and minimum number of eligible assets observed for each sub-period and stock market.

Period	US		BR	
	Minimum	Maximum	Minimum	Maximum
[1927 ; 1962]	478	1,097	-	-
[1963 ; 1972]	1,924	2,365	-	-
[1973 ; 2000]	4,353	7,123	-	-
[2001 ; 2014]	3,546	5,874	37	137

While the number of US assets varies, ranging from 478 to 7,123 in a history of 84 years, we observe a maximum of 137 assets over the 14-year period analyzed for Brazil. This feature may impact on the portfolios' behavior, since the fewer the assets used to build a portfolio, the greater the idiosyncratic risk impact on the portfolio's returns.

The small number of assets could affect the risk premium estimation in two ways: i) by generating distortion in the portfolios' returns, which are used as dependent variables in

[10]In both markets, some eligibility rules are applied to select the assets that compose the portfolios. The US' and Brazil's rules are listed at French's website and NEFIN's website respectively.

[11]The sub-periods are as follows. The first sub-period is the first period built by Center for Research in Security Prices (CRSP) and spans from 1927 to 1962. The second sub-period begins from 1963 and continues until the foundation of the NASDAQ in 1972. The third sub-period begins in 1972 and continues until the starting point of the data period for Brazil, and it ends immediately after the end of the dot-com bubble in 2000. Finally, the last sub-period ranges from 2001 to 2014 and covers the same period as the Brazilian data.

the regressions, and ii) by impacting the estimations of the risk factors. The first concern is discarded by analyzing the standard deviations of the returns on the 22 portfolios for both markets, reported in Table 1. One can see that the standard deviations are quite similar between the markets, and there is no pattern. This means that the standard deviations of the Brazilian data are not always greater than those of the US data. This indicates that no relevant distortions are generated in the Brazilian portfolios, which are used as dependent variables. However, the second highlighted concern demands more attention, considering that, as reported in Table 2, all the risk factors in the Brazilian case show higher standard deviations than those of the US.

The procedure to calculate the risk factors is based on building portfolios using the correlated assets' characteristics with the returns and then defining the returns on those portfolios as realizations of the risk factors (Fama and French, 1993; Carhart, 1997). For example, the factor realization *SMB* is obtained by calculating the return from a portfolio long in small assets and short in big assets. However, considering the low number of assets available in Brazil, risk factor estimation by this procedure can be affected by the asset's idiosyncratic risk. This could explain the pattern observed with regard to the standard deviation of Brazilian factors in Table 2, or the results observed for the risk premia estimations in Panel B of Table 3, in which most parameters have non-significant estimates.

In order to measure the impact of this feature on risk premia estimation, we decrease the number of assets used in the estimation of the US risk factors to a similar number of assets available for the Brazilian market and verify if the risk premia significance frequency is affected. Therefore, we run the following procedure 1,000 times: i) for each year of the 84 years of historical data, we select a sample of 37 and 137 assets, and use them to estimate the risk factor realization for the months of the respective year, and ii) we process the risk premia estimation and verify if the estimated λ_k, such that $k \in \{mkt, smb, hml, wml\}$, are positive and significant, that is, have a t-value ≥ 1.64.

Two methods of sample selection are applied. The first one, denominated "Random," is a random selection from the eligible assets of each year of the 84 years' samples. The second, denominated "Size x Book-to-market", intends to preserve the distribution on the variables Size and Book-to-market among the selected assets[12].

The results obtained from this procedure are presented in Table 5. The first column reports the selection method applied, the second indicates the number of assets selected each year, and the remaining columns show the percentages from the 1,000 estimations that return positive and significant estimates of the parameters α, λ_{mkt}, λ_{smb}, λ_{hml}, and λ_{wml}. We run both selection methods with 37 and 137 assets, the two extreme cases observed in the Brazilian data. The results using 37 assets are presented in Figure 2, where we plot the observed density of the t-values estimated for each risk factor.

The results show that most of the parameters are barely affected. The most affected parameter is the *SMB* risk premium, which has around 85% to 92% of significant estimates when 37 assets are used to build the risk factors, but this result changes to around 99% when 137 assets are used. The second most affected risk premium is *HML*, and its biggest impact is observed with the random selection with 37 assets; 97.8% of the estimates are significant. *WML* has a very small impact and shows more than 98% significant estimates

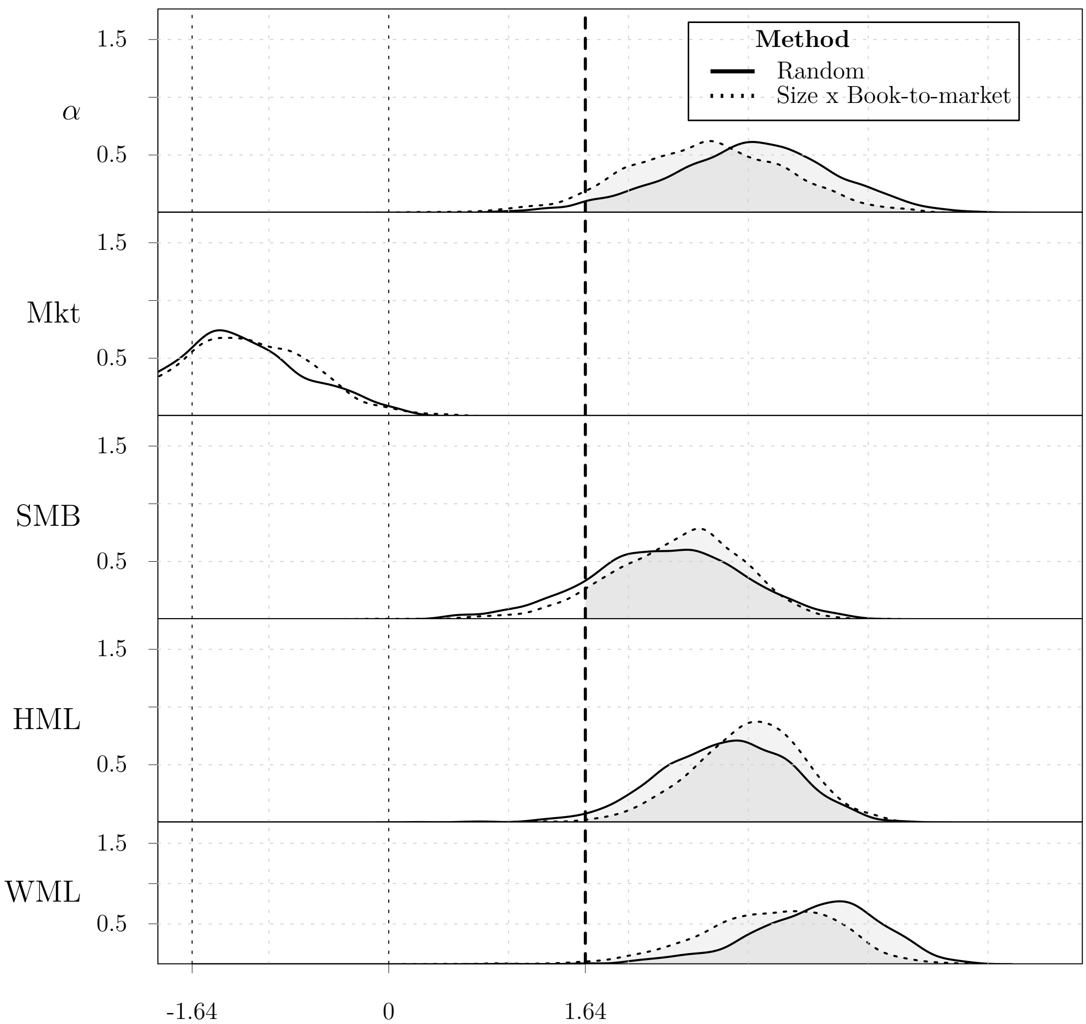
[12] The procedure applied to build the new risk factors' time-series is detailed in Appendix B, which can be provided under requested.

in all cases. *Mkt* shows non-significant results in all cases. Finally, the mean pricing error (α) has more than 93% of significant and positive estimates.

Table 5: Percentage of significant cases by the number of assets

Selection Method	**Number of assets**	**percentage** ($t \geq 1.64$)				
		α	λ_{mkt}	λ_{smb}	λ_{hml}	λ_{wml}
Random	37	97.4	0.0	85.0	97.8	99.8
	137	99.9	0.0	99.7	100.0	100.0
Size x Bookt-to-market	37	93.5	0.0	91.6	99.4	98.8
	137	100.0	0.0	100.0	100.0	100.0

Figure 2: Density of $t_{\hat{\lambda}}$ estimated with factors for 37 assets



The results show that most of the risk factors calculated using 37 or more assets reach the same outcome as when the whole set of assets is used. In other words, there is no indication that the number of assets available for the risk premium estimation generates a big impact in the Brazilian case.

4.3 Does the size of T restrict the estimation?

While the US risk premia results are based on 84 years of historical data, the Brazilian market, for which we conduct the regression, has only 14 years of data. In order to verify whether this divergence is a constraint for Brazilian risk premia estimations, we use the US data again as a benchmark and verify the impact of time-series length on the risk premia estimations.

The analysis uses the US risk factors and portfolios' returns presented in Section 2, and the procedure is described as follows: i) to define a time window length that has the same number of months as the Brazilian data (168 months), ii) to estimate as many regressions as possible on the US data using only the number of months defined in the previous step, that is, starting with the oldest 168-month window allowed until the most recent one, and always dropping the oldest month in exchange for a more recent one. The results are the risk premia estimates for December 1940 to December 2014, using the data of only the last 168 months available. At the end of this procedure, we have a total of 889 sets of estimates.

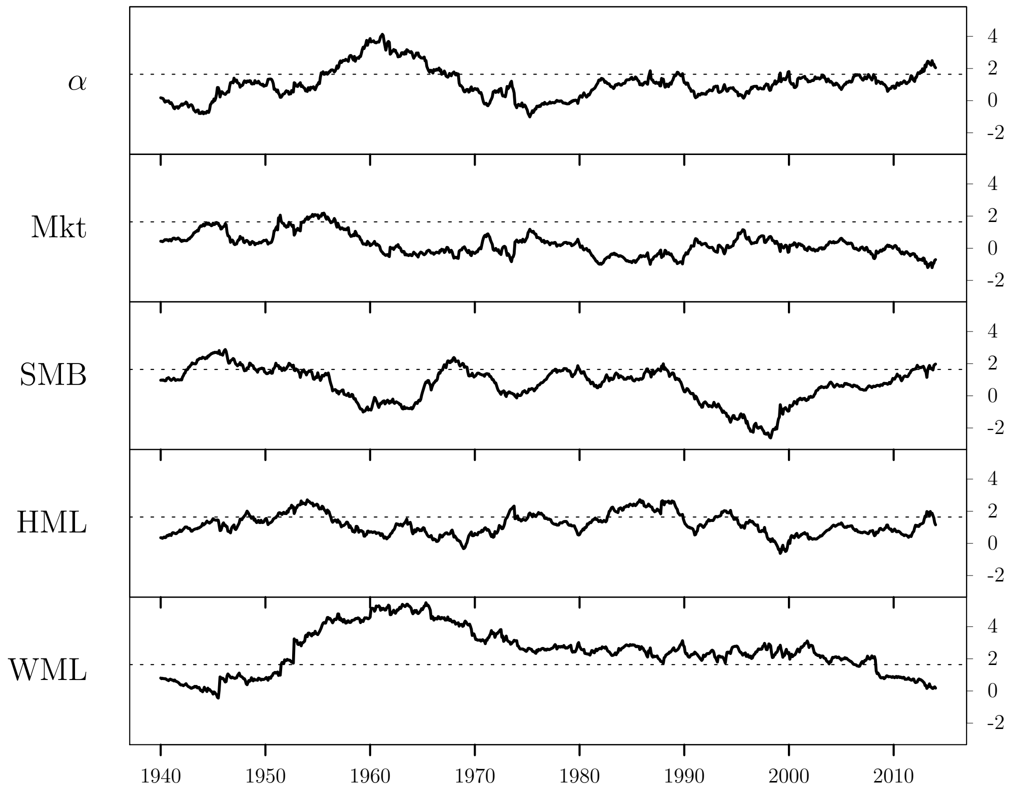
Based on the results obtained from the described procedure, we draw Figure 3. Each graphic in the figure is related to one parameter, α, λ_{mkt}, λ_{smb}, λ_{hml}, or λ_{wml}, and each point of each graphic indicates the t-value obtained for a particular parameter from an estimated model with the last 168 months observed at the reference date. Therefore, Figure 3 is nothing but a history of 889 t-values for each risk premium estimated with the US data, using fixed windows of 168 months. In addition, each graph is accompanied by a dotted line at value 1.64, which is the critical value adopted for rejecting the hypothesis that the parameter is less than or equal to zero.

The results indicate the importance of time-series sample size for the factor model estimation and show that, even with the US data, it is not uncommon to find that risk premia are not significant when short time-series samples are used. It is also notable that *WML* is the most robust factor, just as in the Brazilian case. At the start of the analysis of the market risk premium, we observe that, in most estimates, this parameter is not significant, except for a few points at the beginning of the historical data. *SMB* risk premium shows significance only in a few periods, mostly at the beginning of the historical data, but we can find some significant points in the middle and in very recent periods. The parameters relate to *HML* and *WML*, which, on the other hand, are more robust, since their t-values exceed the critical value on many more occasions. This is especially true for the second factor, which is significant for almost the entire history.

After establishing the importance of time-series sample size to risk premia estimation, we verify how sensitive the estimation is to the time-series sample size. To do so, we apply the following procedure: i) we select several window lengths (48, 72,..., 1056 months), ii) for each option selected in the previous step, we repeat the procedure applied to build Figure 3 and compute the percentage that each risk premium is significant. This analysis is presented in Figure 4. The five graphs in the figure indicate the significance of each parameter according to the window used for the estimations. The dotted line crossing the graphics vertically highlights the percentage obtained with windows of 14 years, the same number available for the Brazilian data.

Figure 3: History of t-values for an estimation window of 14 years

Each point in the figure is the t-value of the parameter α, λ_{mkt}, λ_{smb}, λ_{hml}, or λ_{wml} resulting from the estimation of a model with all the factors and 168 months (14 years) of data. The horizontal dotted line crosses the ordinate axis at 1.64, the critical value to reject the hypothesis that the estimated values are equal to or smaller than zero, with a significance level of 5%.

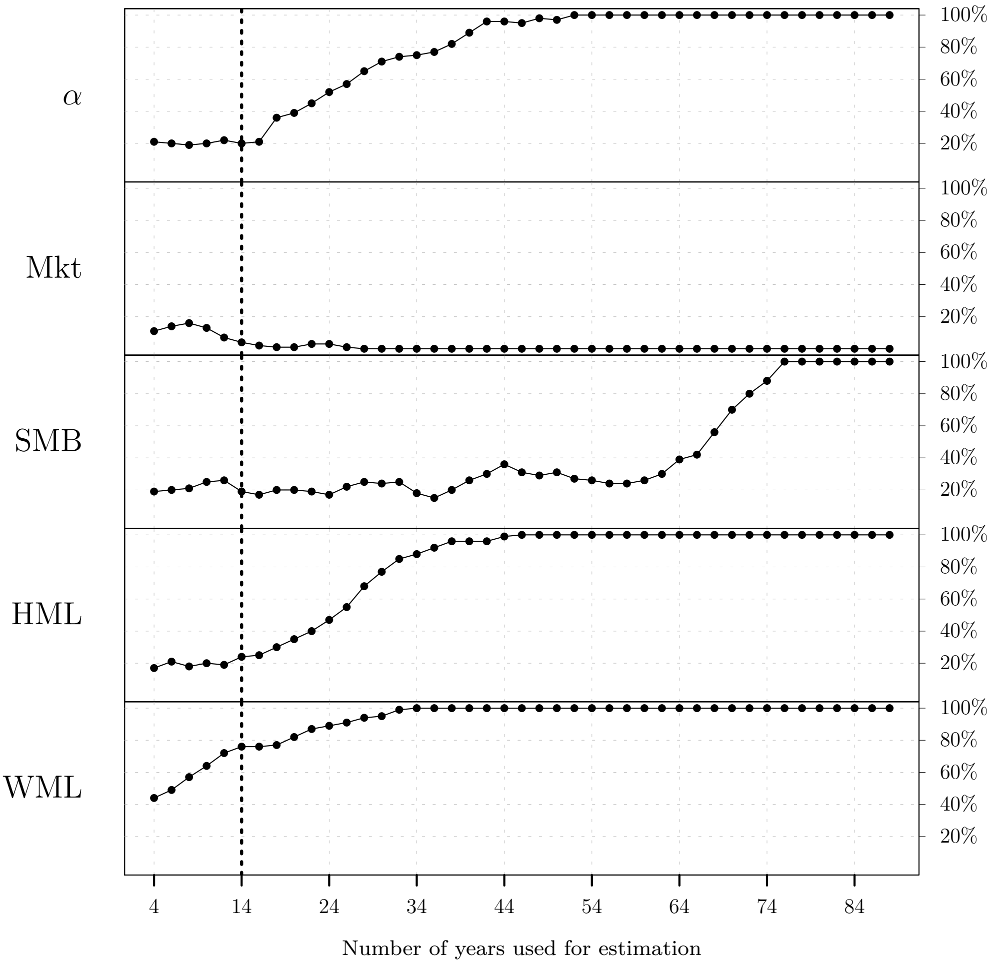


The results indicate that a large time-series sample is required in order to obtain robust estimates. We start with the parameter α, which shows positive and significant results around 20% of the time when 14 years are used for the estimation. However, as the time-series sample becomes larger, it becomes increasingly evident that the data do not support the zero mean error hypothesis. The market risk premium, on the other hand, shows robust results independent of the time-series sample length. Most results for this parameter do not show positive and significant estimates. The results of the *SMB* risk premium are very sensitive to time-series sample size. For this parameter, less than 40% of the estimates are positive and significant even with 64 years. The other two parameters again show themselves as more robust. Their estimations seem to be less sensitive to window size. *HML* risk premium has about 20% significant estimations when 14 years are used, and this percentage reaches 80% with 30 years. The *WML* factor seems to be the most robust of all. Its results are positive and significant about 80% of the time when 14 years or more are used for the estimation.

This section shows that the time-series sample size is a relevant restriction on Brazilian risk premia estimation. Based on the US data, we show how common it is to estimate non-significant risk premia when the number of observed periods is too short. Furthermore, it seems that one should not expect robust results on factor models to which time-series

Figure 4: Percentage of significant cases by the number of years

The graph below shows the percentages for which the t-values of the parameters α, λ_{mkt}, λ_{smb}, λ_{hml}, and λ_{wml} are greater than 1.64, according to the time-series sample size used for the estimations. The dotted line crossing the graphics vertically highlights the percentage obtained with windows of 14 years.



samples shorter than 40 years are applied.

4.4 Why is the impact of T so high?

The analysis of Brazilian risk premia shows that most of the parameters have non-significant estimates. In addition, we show that the source of the problem is not the small number of assets available or their characteristics, but the short historical data of the Brazilian market. The next step is to understand why the "size of T" has such a considerable impact and the consequences of applying a factor model to such a short time-series sample.

The literature on risk premiums estimated with a small T began with Shanken (1992), followed by Jegadeesh and Noh (2013); Kim and Skoulakis (2014); Raponi et al. (2015), and Bai and Zhou (2015). Two points emerge from an analysis of risk premium estimations with a small time-series sample: i) small sample bias on betas, and ii) divergence between *ex-post* and *ex-ante* risk premia.

To understand these consequences, note that as presented in Equation 1, the expected excess return from asset i equals a linear relation between the pricing error (α) and risk compensation ($\beta_i'\lambda$). However, the relation used in the risk premia estimation is

$$\bar{R}_i^e = \alpha^* + \hat{\beta}_i'\lambda^* \tag{6}$$

where $\bar{R}_i^e$ is the average excess return from asset i, $\hat{\beta}_i$ is the estimated risk vector from asset i, and α^* and λ^* are the parameters resulting from this relation. Thus, while Equation 1 has only true parameters, Equation 6 consists only of estimated values.

The first divergence arises from using the estimated beta ($\hat{\beta}_i$) instead of the true beta (β_i). According to Shanken (1992), since the independent variable in Equation 6 is measured with error, the estimator is subject to an errors-in-variables problem, making it biased in small samples. However, the measurement error declines as T increases. Hence, Shanken (1992) shows how the asymptotic standard errors are influenced by the estimation error in the betas and proposes an adjustment for the standard errors and a bias-adjusted estimator. Simulating studies with the US data show a bias of about -16% and -20% when less than 172 months are used in the risk premium estimation (Raponi et al., 2015; Bai and Zhou, 2015; Jegadeesh and Noh, 2013).

The second divergence is caused by the use of the average excess return instead of its true expected value. Averaging (2) over time, imposing (1), and noting that $E(R_i^e) = a_i + \beta_i' E(f)$ yields

$$\bar{R} = \alpha + \beta\left[\lambda - E(f) + \bar{f}\right]. \tag{7}$$

Equation 7 demonstrates that the relation between the true beta and the average excess return results in the so-called *ex-post* risk premium, $\lambda^P = \lambda - E(f) + \bar{f}$, which is equal to the sum of the *ex-ante* risk premium and the unexpected factor outcomes. Since one cannot hope for $\bar{f}$ to be a good estimation of $E(f)$ unless T is large, as Shanken (1992) points out, it is not possible to obtain a consistent estimate of λ when T is fixed.

In order to analyze the distortion that these two divergences may cause in the risk premia estimation, we perform a Monte Carlo simulation based on the following set-up:

$$f_t = \lambda + \epsilon_t \tag{8}$$

$$R_t = \beta f_t + e_t \tag{9}$$

such that $t \in \{1, 2, ..., T\}$, $\epsilon_t \sim N(0, \sigma^2)$, $e_t \sim N(0, \Sigma)$, and $\epsilon_t \perp e_t$. $\lambda = 0.6502$ and $\sigma = 5.41$, which are the mean and standard deviation respectively with regard to the US market risk factors in Table 2. β is the 1×22 vector of the market risk measure from the US market, whose values are presented in Panel A (see the first column) of Table 3. Σ is the 22×22 residual covariance matrix resulting from the US time-series regression in Section 4.1.

Based on this set-up, we select several values of T, and for each of them, we simulate 10,000 draws. We then estimate the risk premium of each draw by two methods: i) the same method presented in Section 3.2 and applied to our paper so far, and ii) a modified estimation in which the risk premium is estimated using the true betas instead of the estimated values. The simulation allows the analysis of several time-series sample sizes, and we can isolate the *ex-post* impact using the true betas in the estimation procedure.

Table 6 presents the results obtained for the risk premium estimation using the estimated beta or the true beta, both for several values of T, varying from 72 to 1056 months. The first column indicates the risk measure used as the independent variable, the estimated beta ($\hat{\beta}$) or true the estimated beta (β). The second column shows the number of months used (T), the third column reports the means of the estimated risk premiums, the fourth column reports the percentage of the 10,000 simulations with a positive estimated value, and the fifth column shows the percentage of positive and significant estimates.

Table 6: Simulation results

The table presents the results obtained for the risk premium estimation using the estimated beta or the true beta for time-series samples varying from 72 to 1056 months. The first column shows which risk measure was used as the dependent variable, the estimated beta ($\hat{\beta}$) or the true beta (β). The second column shows the number of months used. The third column reports the means of the estimated risk premiums. The fourth column indicates the percentage the 10,000 simulations with a positive estimated value, and the fifth column reports the percentage of positive and significant estimates.

Cross-section independent variable	n. of periods (months | years)	$\bar{\hat{\lambda}}$ (% p.m.)	$\hat{\lambda} \geq 0$ (%)	$t_{\hat{\lambda}} \geq 1.64$ (%)
$\hat{\beta}$	72 | 6	0.5224	69.23	9.00
	168 | 14	0.5953	79.12	17.08
	312 | 26	0.6192	87.01	27.46
	456 | 38	0.6290	90.42	36.15
	600 | 50	0.6343	93.58	44.14
	744 | 62	0.6350	95.14	50.89
	888 | 74	0.6353	96.41	57.37
	1056 | 88	0.6422	97.71	64.44
β	72 | 6	0.6404	69.64	13.84
	168 | 14	0.6475	79.27	20.37
	312 | 26	0.6480	87.10	29.60
	456 | 38	0.6472	90.58	37.73
	600 | 50	0.6498	93.64	45.43
	744 | 62	0.6459	95.22	52.16
	888 | 74	0.6453	96.53	58.29
	1056 | 88	0.6502	97.79	65.25

The results confirm that, as expected, the risk premium obtained with the estimated beta is in fact biased, as demonstrated by Raponi et al. (2015); Bai and Zhou (2015); Jegadeesh and Noh (2013). When the time-series sample has only 72 or 168 months of data, estimates are biased by about -20% and -8% respectively , and for the estimations with 1056 months, the bias is around 1%. In contrast, the estimation provided by the true beta shows almost no bias regardless of the time-series sample length.

However, the percentage of positive and significant estimates is almost the same irrespective of the beta used. The percentage of positive and significant estimates is between 9% and 14% for the 72-month time-series sample, about 17% to 20% for the time-series samples of 168 months, and about 64% to 65% for 1056 months.

These results indicate that even though the beta bias is actually a problem in small samples, it does not appear to be a major problem, considering that the magnitude of the bias is small in relation to the *ex-post* distortion on risk premium. The difference between the *ex-post* and the *ex-ante* risk premium lies in the unexpected factor outcomes that have zero mean but high volatility, as the data indicate. Consequently, the *ex-post* risk premium, under an estimation scenario of short time-series samples, may have a wide range of potential values and shows large divergence from the *ex-ante* risk premium. Thus, the results obtained from small time-series samples should not be used to draw conclusions about the actual behavior of the stock market.

5 Conclusion

Although US results are the benchmark for empirical finance literature, we show in this paper that its risk premia estimation is not robust if it is subject the same restrictions observed in emerging markets. In addition, the pattern of US results is the same as the one observed in emerging markets if both are subject to same time-series sample length constraints.

The equity risk premia estimation on emerging market is not robust. Brazilian equity market, for instance, among the 133 market risk premium estimates reported in the literature, 41 are positive, 18 are negative, and the remainder are not significant. This lack of robustness is frequently attributed to deviations from an efficient market. However, conclusions should not be made without inspect what are the observed results with US data if they are subject to a typical emerging market constraints: small number of assets (small N) and short time-series samples (small T).

To the best of our knowledge, nobody has restricted the data to understand how they behave. By adopting such a procedure, we show that data behavior of emerging market economies are not that awkward. By contrast, we are able to show that the US data behavior is not robust by shortening the sample. Therefore, we analyze the sensitivity of US risk premia estimation if restrictions are imposed on the data used.

We conclude that the source of the problem on the emerging market estimations is not the number of assets or their characteristics. The real problem in the estimations lies in the fact that the available time-series sample is short. Accordingly, we investigate which troubles the small T causes. We show that the real problem is due to the high dispersion observed in the risk factors' outcomes, which induces high divergence between *ex-post* and *ex-ante* risk premia. On the other hand, we point out that the betas estimated with time-series samples around 5 years do not generate relevant distortions in the results.

Based on these results, it should not be stated that the problem in the robustness of the results documented for emerging market is data deficiency per se, but in fact is the short time interval of available data. Our analysis indicates that it is necessary to have a time-series sample greater than 40 years in order to obtain robust results.

Thus, the application of the factors model for risk premia estimation poses a significant concern. Practitioners often face this problem when estimating cost of capital and usually apply an alternative solution. While they estimate the risk measure (betas) using short time-series sample, they use a risk premium calculated with longer time-series from other economies, such as that of the US Damodaran (1999). The results of this paper corroborate this alternative solution, since we show there is no relevant distortion from the beta estimation made using short time-series samples, however, the risk premia for those samples are not valid.

References

Bai, J. and G. Zhou (2015). Fama-Macbeth Two-Pass Regressions: Improving Risk Premia Estimates. *Available at SSRN*.

Barry, C. B., E. Goldreyer, L. Lockwood, and M. Rodriguez (2002). Robustness of size and value effects in emerging equity markets, 1985–2000. *Emerging Markets Review 3*(1), 1–30.

Bauman, W. S., C. M. Conover, and R. E. Miller (1998). Growth versus value and large-cap versus small-cap stocks in international markets. *Financial Analysts Journal 54*(2), 75–89.

Bellizia, N. W. (2009). *Aplicação do CAPM para a determinação do custo de capital próprio no Brasil.* Ph. D. thesis, Universidade de São Paulo.

Bodur, F. J. (2011). Uma comparação entre os modelos capm, fama-french e fama-french-carhart.

Bonomo, M. and R. Garcia (2001). Tests of conditional asset pricing models in the Brazilian stock market. *Journal of International Money and Finance 20*(1), 71–90.

Bonomo, M. A., P. L. V. Pereira, and A. Schor (2002). Arbitrage Pricing Theory (APT) e variáveis macroeconômicas: um estudo empírico sobre o mercado acionário brasileiro. *Revista de Economia e Administração 1*(1).

Brito, R. D. and V. Y. Murakoshi (2009). Fatores comuns de risco de mercado, tamanho, valor e diferenciais de juros nos retornos esperados das ações brasileiras. *Revista de Economia e Administração 8*(2).

Cakici, N., F. J. Fabozzi, and S. Tan (2013). Size, value, and momentum in emerging market stock returns. *Emerging Markets Review 16*, 46–65.

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance*, 57–82.

Chague, F. D. and R. De-losso (2007). The CAPM and Fama-French models in Brazil: a comparative study. *In: Encontro brasileiro de finanças*, 1–44.

Chui, A. C. W. and K. C. J. Wei (1998). Book-to-market, firm size, and the turn-of-the-year effect: Evidence from Pacific-Basin emerging markets. *Pacific-Basin finance journal 6*(3-4), 275–293.

Cochrane, J. H. (2001). Asset Pricing. *Experimental Economics 1*(07), 1–553.

Damodaran, A. (1999). Avaliação de investimentos: ferramentas e técnicas para a determinação do valor de qualquer ativo. rio de janeiro: Qualitymark, 2009. _. *Estimating Equity Risk Premiums. Stern School of Business. 24p. New York.*

Eid Jr, W. and C. C. Martins (2015). Pricing Assets with Fama and French 5-Factor Model: a Brazilian market novelty.

Fama, E. F. and K. R. French (1992). The Cross-Section of Expected Stock Returns. *Journal of Finance 47*(2), pp. 427–465.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics 33*(1), 3–56.

Fama, E. F. and K. R. French (1998). Value versus Growth: The International Evidence. *Journal of Finance 53*(6), pp. 1975–1999.

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

Foye, J., D. Mramor, and M. Pahor (2013). A Respecified Fama French Three-Factor Model for the New European Union Member States. *Journal of International Financial Management & Accounting 24*(1), 3–25.

Grandes, M., D. T. Panigo, and R. A. Pasquini (2010). On the estimation of the cost of equity in Latin America. *Emerging Markets Review 11*(4), 373–389.

Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica 50*(4), pp. 1029–1054.

Jegadeesh, N. and J. Noh (2013). Empirical tests of asset pricing models with individual stocks. *Available at SSRN 2382677*.

Kim, S. and G. Skoulakis (2014). Estimating and testing linear factor models using large cross sections: the regression-calibration approach. *Unpublished working paper. University of Maryland, College Park*.

Lischewski, J. and S. Voronkova (2012). Size, value and liquidity. do they really matter on an emerging stock market? *Emerging Markets Review 13*(1), 8–25.

Malaga, F. K. and J. R. Securato (2004). Aplicação do modelo de três fatores de Fama e French no mercado acionário brasileiro: um estudo empírico do período 1995-2003. *XVIII Encontro da ANPAD. Curitiba: Enanpad*.

Matos, F. O. (2006). Evidêcia empírica do modelo capm para o brasil.

Mussa, A., R. Fama, and J. O. dos Santos (2012). A adição do fator de risco momento ao modelo de precificação de ativos dos três fatores de Fama e French aplicado ao mercado acionário brasileiro. *REGE Revista de Gestão 19*(3).

Mussa, A., P. Rogers, and J. R. Securato (2009). Modelos de retornos esperados no mercado brasileiro: testes empíricos utilizando metodologia preditiva. *Revista de Ciências da Administração 11*(23), 192–216.

Mussa, A., J. R. Securato, J. O. dos Santos, and R. Famá (2011). A influência das condições do mercado acionário e da política monetária no comportamento dos indicadores de risco tamanho, índice book-to-market e momento, no mercado acionário brasileiro. *Revista de Ciências da Administração 13*(29), 152.

Piccoli, P. G. R., A. Souza, W. V. da Silva, and J. A. W. Cruz (2015). Revisitando as estratégias de momento: o mercado brasileiro é realmente uma exceção? *Revista de Administração 50*(2), 183.

Raponi, V., C. Robotti, and P. Zaffaroni (2015). Ex-Post Risk Premia and Tests of Multi-Beta Models in Large Cross-Sections.

Rizzi, L. J. (2012). *Análise comparativa de modelos para determinação do custo de capital próprio: CAPM, três fatores de Fama e French (1993) e quatro fatores de Carhart (1997).* Ph. D. thesis, Universidade de São Paulo.

Rouwenhorst, K. G. (1999). Local Return Factors and Turnover in Emerging Stock Markets. *Journal of Finance 54*(4), 1439–1464.

Sampaio, F. S. (2002). Existe equity premium puzzle no brasil. *Finanças aplicadas ao Brasil 1*, 87–117.

Shanken, J. (1992). On the estimation of beta-pricing models. *Review of Financial Studies 5*(1), 1–55.

Varga, G. and R. Brito (2015). The Cross-Section of Expected Stock Returns in Brazil.