# Stock Index Futures Arbitrage: Evidence from a Meta-analysis

Jędrzej Białkowski[1]
Department of Economics and Finance,
University of Canterbury, New Zealand

Devmali Perera[2]
Department of Economics and Finance,
University of Canterbury, New Zealand

Revised: 03-09-2018

Abstract

A number of empirical studies have focused on examining the stock index futures arbitrage. The reported results were not consistent and depended on a number of factors. Our study aims to review the literature on stock index arbitrage using meta-regression techniques. In particular, it aims to synthesize estimates on the existence of mispricing and on the relationship between mispricing and time to maturity. This study does not find strong evidence on the publication bias in reported estimates on mispricing and estimates on the effect of time to maturity on mispricing. Finally, this study tests whether characteristics of data and publication determine the heterogeneity in the reported estimates on mispricing and time to maturity. The results suggest that these characteristics could explain these differences significantly.

*Keywords*: Cost-of-carry model, Index arbitrage, Meta-analysis, Publication bias

JEL Classification: G13, G14, B41

[1] Corresponding Author: Department of Economics and Finance, UC Business School, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand. Phone: +64 3 369 4060.
Email: jedrzej.bialkowski@canterbury.ac.nz

[2] Department of Economics and Finance, UC Business School, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand. Phone: +64 22 165 1494. Email: devmali.perera@pg.canterbury.ac.nz

## 1. Introduction

Examining the availability of arbitrage opportunities and arbitrage profits between stock index prices and stock index futures prices has been the subject of a number of empirical studies. In an efficient market with no frictions, the theoretical futures price is supposed to be equal to the market price of a futures contract. The traditional approach of calculating the theoretical futures price is using the cost-of-carry model proposed by Cornell and French (1983). The mispricing is defined as the difference between the market futures price and the theoretical price given by the cost-of-carry formula. Stoll and Whaley (1986) argue that there will be no arbitrage opportunities if the mispricing lies within the upper and lower boundaries of the transaction cost.

Some studies report the existence of mispricing in the U.S. market (Chung, 1991; Klemkosky and Lee, 1991; MacKinlay and Ramaswamy, 1988; Neal, 1996). In the case of European markets, Yadav and Pope (1990) report mispricing in U.K., Bühler and Kempf (1995) for the German market, Stulz, Wasserfallen, and Stucki (1990) for the Swiss market, Puttonen (1993) for the Finnish market, Bialkowski and Jakubowski (2008) and Marcinkiewicz (2016) for the Polish market, Andreou and Pierides (2008) and Fassas (2011) for the Greek market and Berglund and Kabir (2003) for the Dutch market. In the case of Asian markets, there is evidence of mispricing in the Japanese market (Brenner, Subrahmanyam, and Uno, 1989a; 1989b; Chung, Kang, and Rhee, 2003; Lim, 1992), in the Hong Kong market (Fung and Draper, 1999; Fung and Lam, 2004), in the Taiwanese market (Lin, Lee, and Wang, 2013; Wang, 2010) and in the Indian market (Vipul, 2005).

The research findings on the existence of mispricing, profitability of mispricing and factors affecting the level of mispricing are inconclusive. Some studies find that the mispricing on average is negative (Brailsford and Hodgson, 1997; Cornell and French, 1983; Figlewski, 1984; Gay and Jung, 1999; Modest and Sundaresan, 1983), whereas there is some evidence on positive mean mispricing as well (Bhatt and Cakici, 1990; MacKinlay and Ramaswamy, 1988). There is another argument that derivatives markets are less efficient at their early trading history and the magnitude and frequency of mispricing diminishes over time when the market develops (Bialkowski and Jakubowski, 2008; Buhler and Kempf, 1995; Dwyer, Locke and Yu, 1996).

Furthermore, previous studies have examined the factors that determine the mispricing in stock index futures contracts. The mispricing is caused by differences in tax rates applicable to stocks and futures (Cornell and French, 1983), dividend uncertainty (Peters, 1985), unequal borrowing and lending rates (Gould, 1988), imperfect replication of the underlying index, differences in execution lags and transaction costs involved (Chung, 1991; MacKinlay and Ramaswamy, 1988) and short-selling availability (Fung and Draper, 1999; Pope and Yadav, 1994). The studies in different countries report that the degree of importance of these sources of mispricing varies due to the institutional framework and the structure of the respective market as well.

When the findings on a research theme are contrasting, it is essential to review the literature to discuss the reasons for this heterogeneity in results. Traditional narrative reviews of literature often suffer from a researcher bias because these reviews synthesize and summarize previous findings descriptively rather than quantitatively. The meta-analysis approach overcomes this issue by providing a systematic and objective process of conducting the literature review along with some statistical techniques to synthesize and summarize the existing findings. This methodology is widely adopted in medical science research studies and is now gaining in popularity among finance[3] researchers as well.

[3] See the studies Arestis, Chortareas, and Magkonis (2015); Asongu (2015); Astakhov, Havranek, and Novak (2017); Ewijk, de Groot, and Santing (2012); Fidrmuc and Lind (2018); Rusnak, Havranek, and Horvath (2013) and Zigraiova and Havranek (2016) for recent examples of the application of meta-analysis in finance.

Accordingly, this study aims to review the literature on stock index futures arbitrage using meta-analysis techniques to reach a single conclusion about the existence of mispricing and the relationship between mispricing and time to maturity of a contract. We collect mean mispricing data reported in different markets, data on factors affecting the mispricing and data that describes differences in markets such as transaction costs, availability of short selling and the level of development of the market. To be precise, this study examines whether there is a publication bias regarding stock index arbitrage and analyses the reasons for the heterogeneity in results reported across studies.

The remainder of the paper is organized as follows. Section 2 discusses the research design, with details about the selection of papers and coding of data, and also provides the findings of a preliminary analysis. Section 3 explains the methodology of testing for publication bias. Section 4 describes the meta-regression method used to analyze the factors affecting the heterogeneous results. Section 5 presents and discusses the findings on publication bias and the results of meta-regression analysis. Section 6 summarizes and concludes the paper.

## 2. Research Design

### *2.1. Selection of studies*

A meta-analysis is a statistical analysis of estimates collected from multiple studies that are concerned with examining a similar relationship. This study synthesizes estimates from the research studies analyzing the presence of stock index futures arbitrage and the factors affecting the level of mispricing. We used the key words: “ARBITRAGE WITH FUTURES”, “INDEX FUTURES ARBITRAGE”, “STOCK INDEX FUTURES” and “MISPRICING OF INDEX FUTURES” to search for papers relevant for this theme. We searched literature electronically in Google Scholar, Web of Sciences, Research Gate, SSRN, EconLit, ProQuest, JSTOR, Scopus, RePEc and EBSCO. We identified a list of 90 relevant papers but selected only 36 for this study. We closed our literature search on 30th April 2018. The final sample of studies selected are listed in Appendix 1.

We selected studies based on the following inclusion criteria. First, studies should calculate the mispricing using the cost-of-carry model with or without transaction costs. Second, studies should consider index arbitrage between stock index and index futures only. Third, studies should consider buy-and-hold arbitrage strategy only. Fourth, we included studies written in English only. Finally, we selected papers that report a mean mispricing estimate along with its standard error (or any other statistics that help to calculate the standard error) and papers that report the estimated beta coefficient of time to maturity and its standard error from a regression of mispricing on time to maturity.

We excluded papers considering put-call futures arbitrage strategy, index futures and index funds arbitrage strategy and early winding or rollover arbitrage strategy. Furthermore, we excluded multiple papers of the same authors providing the same results (exactly same numerical values for the same sample of data) by keeping only the first published paper in the sample). Finally, we excluded the studies that analyze the stock index arbitrage using other different techniques, such as autocorrelation and co-integration techniques. In addition, it is worth noting that there was not a sufficient number of papers in each of these exclusion criteria groupings to be considered as a subsample in this study.

Finally, we could not analyze the evidence on arbitrage profits reported in the studies because only 11 studies in the sample report arbitrage profit values and even out of those only four studies provide standard errors of arbitrage profits or other relevant information (standard deviations, number of observations and *t* statistics) required to calculate the standard error. Hence, we do not analyze the overall significance of arbitrage profitability in this study.

### 2.2. Coding of data

We read each paper thoroughly and coded a set of study characteristics and statistical estimates reported in each study. The coding was done independently by two coders[4] with a careful reconciliation of any inconsistencies. All search and coding procedures conducted are in line with the Meta-analysis of Economics Research Reporting Guidelines of Stanley et al. (2013).

We collected mispricing estimates from studies which use the cost-of-carry model (Cornell and French, 1983) to calculate the theoretical price of a futures contract.

$$F_t = S_t e^{(r_t - \delta_t)T}, \qquad (1)$$

where $F_t$ is the theoretical futures price at *t*, $S_t$ is the spot index price at *t*, $r_t$ is the interest rate at *t*, $\delta_t$ is the dividend yield at *t* and *T* is the time to maturity of the futures contract.

We collected the mean mispricing percentages reported in each study along with the respective standard error or any other statistics (such as standard deviation, *t* statistic and number of observations) that help to calculate standard error. The selected studies calculate the relative mispricing percentage using the following two definitions.

$$R_t = \frac{f_t - F_t}{F_t}, \qquad (2)$$

where $f_t$ is the actual market price of a futures contract at *t* and $F_t$ is the theoretical futures price at *t*.

$$R_t = \frac{f_t - F_t}{S_t}, \qquad (3)$$

where $f_t$ is the actual market price of a futures contract at *t*, $F_t$ is the theoretical futures price at *t* and $S_t$ is the spot index price at *t*.

Due to this inconsistency in the definition of relative mispricing, we transform the reported mispricing estimates into Partial Correlation Coefficients (PCCs). This is the standard

[4] The two coders are Devmali Perera and Ha Hai Hoang, a Masters' student recruited as a research assistant. The coding process was conducted with the guidance and support of Associate Professor Jedrzej Bialkowski.

practice in meta-analysis. The PCC is a unitless measure of the strength and direction of the association between two variables, holding other factors constant. We calculated the PCC using the following formula suggested by Stanley and Doucouliagos (2012).

$$PCC = \frac{t}{\sqrt{t^2 + df}}, \quad (4)$$

where *t* is the *t* statistic of the reported mispricing and the *df* denotes the degrees of freedom (i.e. *n-1*) used for the estimation in each paper. The corresponding standard errors of the PCC are also calculated as follows (Stanley and Doucouliagos, 2012).

$$SE(PCC) = \sqrt{\frac{(1 - PCC^2)}{df}} \quad (5)$$

The second objective of this study is to analyze the linear relationship between mispricing and time to maturity. We have collected the beta coefficients of time to maturity (beta (TTM)) and respective standard errors or any other statistics (such as standard deviation, *t* statistic and number of observations) that help to calculate the standard error. We plotted beta (TTM) on a scatter plot (not provided in this study) and found that there are three outlier estimates having a large positive value. We removed these three values from the analysis, and the study then has only 23 beta (TTM) estimates remaining in the sample.

In addition, we have coded some additional data that depict the different characteristics of each study. Appendix 2 provides the list of these variables and their definitions.

### *2.3. Preliminary analysis of data*

This section summarizes the characteristics of the sample of studies selected and presents the descriptive statistics of the data collected. Figure 1 depicts the composition of the sample of studies by country. We have papers representing 15 countries and the highest number of papers are on U.S. and Japanese markets, respectively.

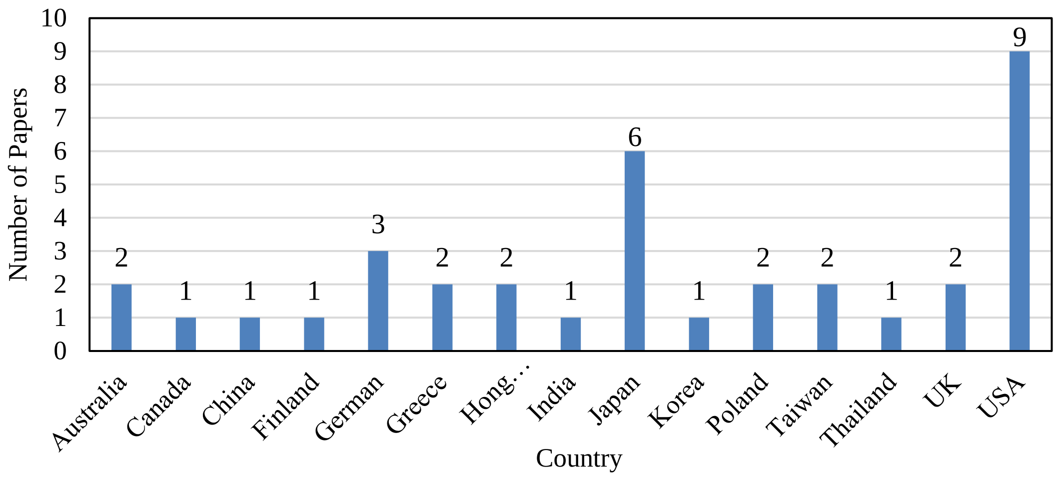


Figure 1: Composition of the sample by country

Figure 2 depicts the composition of the sample by publication year. This study includes papers published during the period from 1987 to 2016. The highest number of papers were published in 1997.

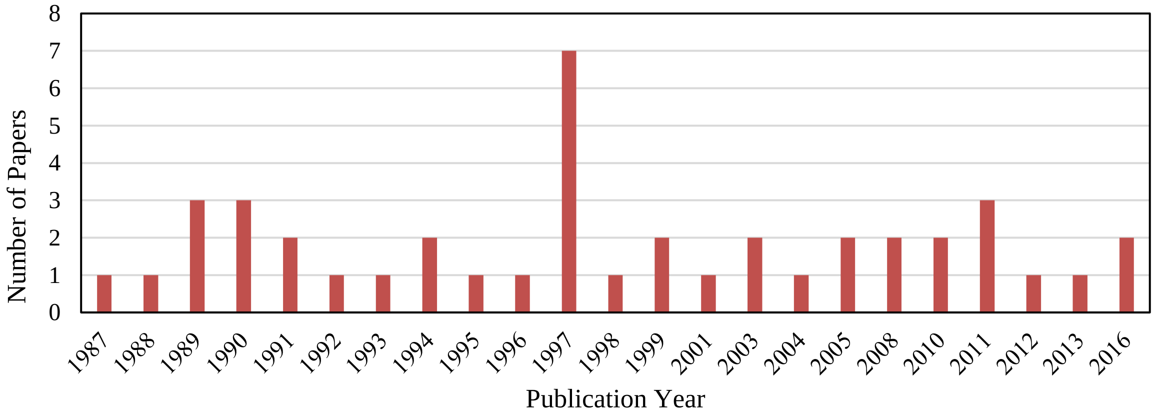


Figure 2: Composition of the sample by publication year

Figure 3 presents the composition of the sample by journal rank. The 25 papers out of 36 selected were published in A and A* journals[5] (i.e. 69% of the sample). There are only six studies which were either not published or published in a journal which does not have a journal rank based on the ABDC ranking system.

---

[5] Journal ranks are based on the ABDC Journal Ranking System published in 2017. Retrieved from http://www.abdc.edu.au/master-journal-list.php

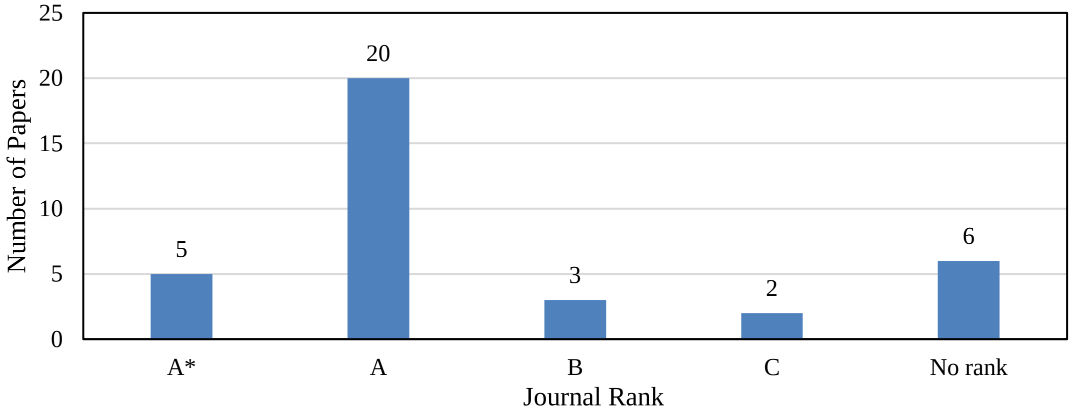


Figure 3: Composition of the sample by journal rank

As a preliminary study on the data, we calculated the descriptive statistics of some selected variables and Table 1 summarizes these descriptive statistics. Accordingly, the average mispricing is 4.38 basis points across the sample of 357 mispricing estimates collected from the sample. The mean PCC is 0.0695. The negative mispricing (approximately 65%) is relatively common compared with positive mispricing (approximately 30%). The mean mispricing across the studies has a standard deviation of 71 basis points.

Table 1: Descriptive statistics

| Variable | Est. | Mean | STD | Min | Max |
|---|---|---|---|---|---|
| Mean Mispricing (%) | 357 | 0.0438 | 0.7120 | -4.5148 | 4.7122 |
| Partial Correlation Coefficient | 357 | 0.0695 | 0.7215 | -2.3040 | 1.4755 |
| Total Mispricing Frequency | 316 | 2416 | 6028 | 3 | 26588 |
| Positive Mispricing Proportion | 147 | 0.2958 | 0.2866 | 0.0000 | 1.0000 |
| Negative Mispricing Proportion | 147 | 0.6461 | 0.3012 | 0.0000 | 1.0000 |

Note: This table provides descriptive statistics of mean mispricing percentage, Partial Correlation Coefficient (PCC), total mispricing frequency, proportions of positive and negative mispricing observations. We report the number of estimates (Est.), mean, standard deviation (STD) and the minimum and maximum values of each variable.

Furthermore, we present the distribution of PCC and the distribution of the beta (TTM) graphically in Figure 4 and 5. The histograms show that PCC and the beta (TTM) are not normally distributed and heterogeneous.

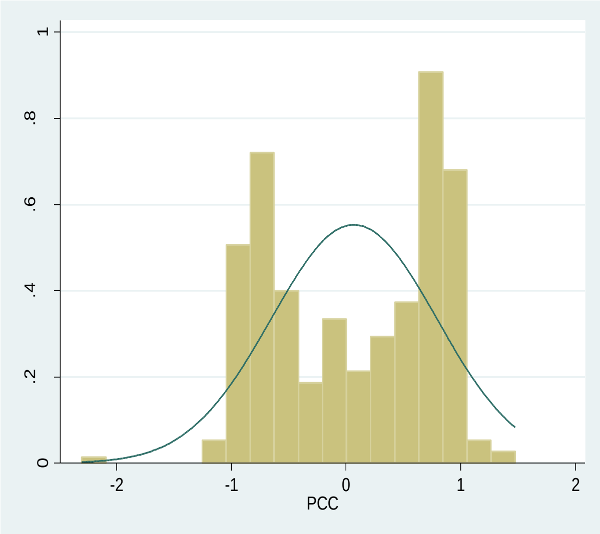


Figure 4: Histogram of PCC

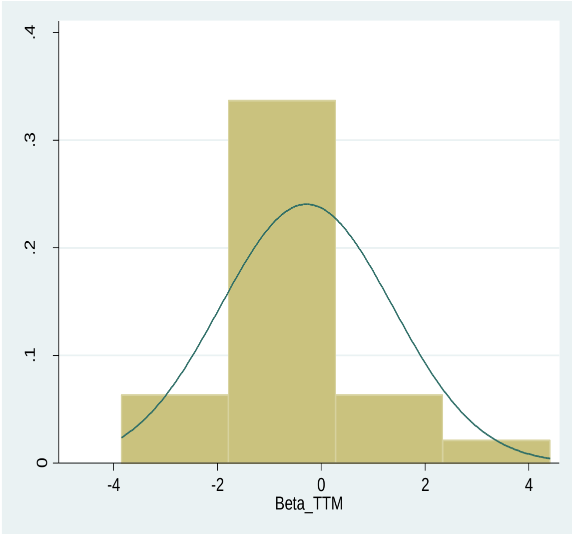


Figure 5: Histogram of beta (TTM)

## 3. Testing for publication bias

### *3.1. Publication bias in mispricing estimates*

First, we test for the publication bias of reported mean mispricing estimates. The publication selection bias arises when an estimate's probability of being reported depends on its sign or statistical significance. This implies that researchers may either hide or not publish the estimates that are insignificant or have a different sign from what is expected. This is a problem, as it distorts the true distribution of the effect. In addition, Stanley and Doucouliagos (2012) find that most studies in the area of empirical economics suffer from publication bias. We first test the publication bias using a visual test called Funnel Plot (Egger et al., 1997) and then test statistically using the Funnel Asymmetry Test (FAT) (Card and Krueger, 1995).

FAT analyses the relationship between PCC and its standard errors. This test suggests that the reported estimates are correlated with their standard errors in the presence of publication bias. Hence, we estimate the following regression in order to test the publication bias using FAT.

$$PCC_i = \beta_0 + \beta_1 SE(PCC_i) + \varepsilon_i, \quad (6)$$

where $PCC_i$ is the PCC of the mean mispricing estimate $i$, $SE(PCC_i)$ is the standard error of the PCC of the mean mispricing estimate $i$, $\beta_0$ measures the overall mispricing corrected for the potential publication bias, $\beta_1$ measures the extent of publication bias and $\varepsilon_i$ is the disturbance term. We test the nul hypothesis that $H0$: $\beta_1$ = 0 or the null hypothesis of no publication bias in the FAT. If the results reject the null hypothesis, that implies there is a publication bias. The direction of the publication bias depends on the sign of $\beta_1$.

Moreover, we test the null hypothesis that $H0$: $\beta_0$ = 0 or the null hypothesis that there is no overall effect. This test is known as the Precision Effect Test (PET) and it measures the statistical significance of the overall effect after controlling for the publication bias (Shemilt et al., 2011; Stanley and Doucoulliagos, 2012). If the results reject the null hypothesis, that implies the model has an overall effect.

First, we estimate the regression in equation (6) using an Ordinary Least Square (OLS) estimator with robust standard errors. However, since this model suffers from heteroscedasticity issues, the common practice in meta-analysis is to weight the model by precision (i.e. using the inverse of the standard error). Following Stanley (2005; 2008), we estimate the model in equation (6) using Weighted Least Square (WLS) to give more weight to more precise studies. Hence, the regression model now becomes as follows.

$$t_{PCC_i} = \beta_1 + \beta_0 \frac{1}{SE(PCC_i)} + \nu_i \;, \tag{7}$$

where $t_{PCC_i}$ equals to the PCC of the mean mispricing estimate $i$ divided by the standard error of $PCC_i$, $\frac{1}{SE(PCC_i)}$ is the inverse of the standard error of $PCC_i$ and $\nu_i$ is the error term. We test the hypotheses $H0$: $\beta_0$ = 0 and $H0$: $\beta_1$= 0 as mentioned above. However, in this regression, the new constant $\beta_1$ indicates the publication bias and the slope $\beta_0$ indicates the existence of a genuine effect after controlling for publication bias. If there is publication bias, we reject the null hypothesis that $\beta_1 = 0$. If there is an overall effect, we reject the null hypothesis that $\beta_0 = 0$. The WLS regression includes robust standard errors.

The OLS estimation of equation (6) includes a bias arising from the correlation between the estimates of the same study. To avoid this issue, we estimate the above regression using a Fixed Effect (FE) model as well. The FE model in meta-analysis assumes that there is an identical true effect size which underlies all the studies in the sample, and these estimates can differ only due to the sampling error. We estimate the FE regression using two weighting schemes: "Weight 1" and "Weight 2".

As with the above WLS estimation model, we use the inverse of the standard error of PCC as the weighting factor. This weighting scheme allocates the same weight to each estimate. Hence, a study including more than one estimate would be weighted more than a study including a single estimate, holding other factors constant. This is the standard weighting scheme and it is denoted by "Weight 1" in this study. The coefficients estimated under WLS and FE with the "Weight 1" model would be same, as we are using the inverse of the standard error as the weighting scheme under both models. However, the standard errors of these two models will be different because WLS includes robust standard errors, whereas FE model with Weight 1 includes robust standard errors clustered by the Study_ID.

As an alternative, we multiply the set of weights given in "Weight 1" by the inverse of the number of estimates reported per study (i.e. by $1/N_i$). In this alternative scheme, we assign each study the same weight but different weights to each estimate. This alternative weighting scheme is denoted by "Weight 2" in this study. FE model with Weight 2 also includes robust standard errors clustered by the Study_ID.

*3.2. Publication bias related to mispricing and time to maturity relationship*

In order to test the publication bias in the estimates of the relationship between mispricing and time to maturity, we use the same techniques mentioned above. First, we plot the Funnel Plot with beta (TTM) on the horizontal axis and the inverse of the standard errors on the vertical axis. Second, we estimate the following regression model with an OLS estimator.

$$TTM_i = \beta_0 + \beta_1 SE(TTM_i) + \varepsilon_i, \quad (8)$$

where $TTM_i$ is the $i^{th}$ beta coefficient of time to maturity, $SE(TTM_i)$ is the standard error of the $i^{th}$ beta coefficient of time to maturity and $\varepsilon_i$ is the error term. As mentioned above, $\beta_0$ measures the overall effect between mispricing and time to maturity corrected for the potential publication bias (i.e. PET) and $\beta_1$ measures the extent of publication bias (i.e. FAT). We test the hypotheses that $H0$: $\beta_1 = 0$ and $H0$: $\beta_0 = 0$ and interpret the results in the same way as discussed above.

In addition, we estimate the regression in equation (8) using a WLS estimator to address the issue of heteroscedasticity. The weighting factor is the inverse of standard errors of the beta (TTM). Accordingly, the regression model will be as follows.

$$t_{TTM_i} = \beta_1 + \beta_0 \frac{1}{SE(TTM_i)} + \nu_i, \quad (9)$$

where $t_{TTM_i}$ equals to the $i^{th}$ beta coefficient of time to maturity divided by the respective standard error of time to maturity, $\frac{1}{SE(TTM_i)}$ is the inverse of the standard error of $i^{th}$ beta coefficient of time to maturity and $\nu_i$ is the error term. We test the same hypotheses $H0$: $\beta_0 = 0$ and $H0$: $\beta_1 = 0$. In this model, $\beta_1$ indicates the publication bias (FAT) and $\beta_0$ indicates the existence of a genuine effect after controlling for publication bias (PET).

Finally, we estimate the above regression in equation (8) using an FE model as well. This FE regression also adopts the same weighting schemes mentioned in the previous subsection: “Weight 1” and “Weight 2”.

## 4. Meta-regression analysis

The aim of the meta-regression analysis is to identify the factors that would cause the reported mean mispricing to differ across studies. To search for these determinants, we run the regression in equation (6) and (7), with additional control variables added to describe the characteristics of data and the publication. We specify here only the revised regression model for equation (7).

$$t_{PCC_i} = \beta_1 + \beta_0 \frac{1}{SE(PCC_i)} + \sum_{k=1}^{K} \frac{\delta_k Z_{ik}}{SE(PCC)_i} + \nu_i, \quad (7)$$

where $t_{PCC_i}$ equals to the PCC of the mean mispricing estimate *i* divided by the standard error of $PCC_i$, $\frac{1}{SE(PCC_i)}$ is the inverse of the standard error of $PCC_i$, $Z_{ik}$ are the control variables added to the regression model and $\nu_i$ is the error term. The choice of these control variables depends on the research question of the present study and the characteristics of the studies in the sample.

Accordingly, we added three variables that characterize the differences in data. We examine whether the level of development in the stock market[6] (i.e. Developed versus Emerging), the availability of short selling and the adjustment of transaction costs determine the heterogeneity in the reported mispricing estimates.[7] The dummy variable (Developed_Dummy) is used to distinguish developed and emerging markets by 1 and 0, respectively.[8] The availability of short selling is also represented by a dummy variable

[6] Buhler and Kempf (1995) and Dwyer, Locke and Yu (1996) document that emerging derivatives markets are less efficient, and the magnitude of mispricing diminishes over time.

[7] There are previous studies concluding that transaction costs and availability of short selling affect the extent of stock index futures mispricing (Andreou and Pierides, 2008; Brenner, Subrahmanyam and Uno, 1989a; 1990; Chung, Kang and Rhee, 2003; Fassas, 2011; Fung and Draper, 1999).

[8] Stock markets are classified as developed and emerging based on the Morgan Stanley Capital International, Inc. (MSCI) 2017 classification. Retrieved from https://www.msci.com/market-classification

(SS_Dummy), where the value is one if short selling is available and zero if it is restricted. The transaction cost adjustment is also a dummy variable (TC_Dummy) which takes the value of one if transaction costs are adjusted in the study and zero if not.

We added three control variables that characterize the differences in the publications as well. These variables are publication year, journal rank and the impact factor of the journal. The publication year of the study is coded from each study. Journal rank is based on the ABDC Journal Raking System of 2017. We create a dummy variable for journal rank (Rank_Dummy) which takes the value of one if the rank is A* or A and zero if the rank is B or C. The impact factors of the published journals are the impact factors published on https://ideas.repec.org downloaded on 23rd May 2018. We then estimate the regression in equation (10) using WLS model with “Weight 1” and using FE model with “Weight 1” and “Weight 2”. The null hypothesis tests whether the coefficient of each explanatory variable equals to zero separately.

We estimate the same meta-regression model with the beta (TTM) as well. The objective is to identify the factors that influence the difference in the relationship between the mispricing and time to maturity. We regress the beta (TTM) on the inverse of the standard errors of the beta (TTM), transaction cost dummy, short selling dummy, developed dummy, publication year, journal rank dummy and the impact factor. Finally, we test the significance of the coefficients of each explanatory variable.

## 5. Results

### *5.1. Publication bias of mispricing*

The Funnel Plot in Figure 6 depicts the PCC of the mean mispricing estimates on the horizontal axis and the precision (i.e. the inverse of standard error of PCCs) on the vertical axis. The Funnel Plot provides a visual test of publication bias. We can observe that PCCs are relatively symmetrically distributed around the mean PCC. This suggests that there will not be any publication bias with regard to mean mispricing estimates.

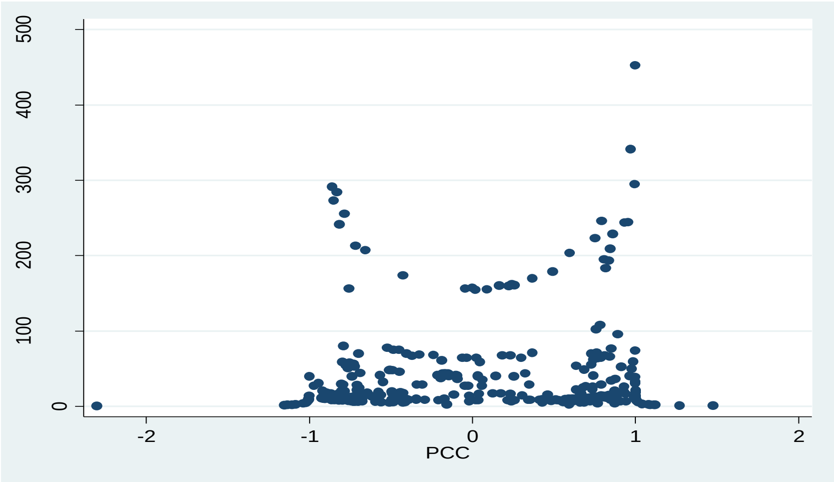


Figure 6: Funnel Plot of PCC

Table 2 below summarizes the results of FAT and PET estimated with OLS, WLS and FE with Weight 1 and FE with Weight 2. None of these estimated $\beta_1$ coefficients are significant except the $\beta_1$ estimated using the WLS estimator. Under WLS, we reject the null hypothesis of no publication bias at the 10% significance level. Hence, we conclude that there is publication bias in favour of positive mispricing estimates of stock index futures arbitrage only under the WLS method. This might be due to the fact that either researchers more prefer to publish positive mispricing estimates than negative mispricing estimates or reviewers in the ranked journals disproportionately report positive mispricing results against negative mispricing.

The $\beta_0$ values under OLS, WLS and FE models report the overall mean mispricing estimate, after controlling for the publication bias. The $\beta_0$ is positive and significant at the 10% significance level only under the OLS model. It indicates that the mean mispricing is positive on average and according to Doucouliagos (2011)[9] guidelines of interpreting PCC, this estimate is a medium sized effect.

Table 2: Funnel Asymmetry Test and Precision Effect Test results of mispricing

| | OLS | WLS | FE (Weight 1) | FE (Weight 2) |
|---|---|---|---|---|
| Publication Bias (FAT) $\beta_1$ | -0.3320 | **0.2943*** | 0.2943 | 0.4114 |
| | (0.5211) | (0.1750) | (0.3888) | (0.4139) |
| Genuine Effect (PET) $\beta_0$ | **0.1012*** | -5.8971 | -5.8971 | -8.4691 |
| | (0.0563) | (4.4992) | (8.7636) | (11.5420) |
| Number of Estimates | 357 | 357 | 357 | 357 |
| Adjusted R Squared | 0.0020 | 0.0999 | 0.0999 | 0.1646 |

Note: This table reports coefficients of $\beta_1$ (FAT) and $\beta_0$ (PET), respectively. In both cases, the top value is the coefficient estimate and the value in parentheses is the associated standard error. The first column provides the Ordinary Least Square (OLS) estimates, second column provides Weighted Least Square (WLS) estimates (weighted by the inverse of standard error) and the last columns provide Fixed Effect (FE) estimates with weight 1 (each estimate is given an equal weight) and with weight 2 (each study is given an equal weight). OLS and WLS estimations have robust standard errors and FE estimates have robust standard errors clustered by Study_ID. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

### *5.2. Meta-regression analysis of the determinants of mispricing*

Next, we test whether the mispricing varies systematically depending on the characteristics of data and publication. Table 3 presents the results of the meta-regression analysis on mispricing. The results suggest that there is no publication bias in reported mispricing under any model, after controlling for these other explanatory variables.

Under FE models with Weight 1 and Weight 2, a financial market with short selling availability reports approximately 0.06 larger mean mispricing compared to a market which does not allow short selling. Kempf (1998) and Puttonen (1993) point out that short selling restriction significantly influence the mispricing to be negative. The TC_Dummy variable shows that the mean mispricing is lower by 0.37 and 0.53 respectively, under OLS and WLS models when a study adjusts for the transaction costs compared for not adjusting transaction

[9] Doucouliagos (2011) answer the key question when the partial correlation coefficient can be treated as large. According to his guideline, a PCC less than ±0.07 can be regarded as small (even if it is statistically significant) and a PCC greater than ±0.33 can be regarded as large.

costs. The existence of transaction cost implies that mispricing within the upper and lower bounds of transaction cost considered as consistent with market efficiency. Hence, a study that adjusts for transaction costs is expected to report a lower level of mean mispricing compared to a study that does not adjust for transaction costs. Finally, the Developed_Dummy reports a negative and statistically significant coefficient under all four models. This reinstates the fact that developed financial markets are comparatively efficient than emerging financial markets and hence report a lower level of mean mispricing estimate than emerging markets.

Table 3: Meta-regression analysis of mispricing

| Variable | OLS | WLS | FE (Weight1) | FE (Weight 2) |
|---|---|---|---|---|
| SE of PCC | -0.5848 | 28.3197 | 28.3197 | 93.3546 |
| | (0.4403) | (58.1636) | (69.4941) | (86.2408) |
| SS_Dummy | 0.3526 | 0.0614 | **0.0614***** | **0.0556***** |
| | (0.2524) | (0.2038) | (0.0184) | (0.0154) |
| TC_Dummy | **-0.3678***** | **-0.5273**** | -0.5273 | -0.5592 |
| | (0.0955) | (0.2105) | (0.4175) | (0.4726) |
| Developed_Dummy | **-0.9605***** | **-1.4244***** | **-1.4244**** | **-2.0351*** |
| | (0.2354) | (0.4087) | (0.4880) | (0.9934) |
| Rank_Dummy | **-0.4454***** | **-11.9135***** | -11.9135 | -11.6038 |
| | (0.1076) | (3.4146) | (6.7760) | (8.3384) |
| Pub_Year | **-0.0155**** | -0.0135 | -0.0135 | -0.0458 |
| | (0.0061) | (0.0289) | (0.0345) | (0.0428) |
| Impfac | **0.0280***** | **0.1862***** | 0.1862 | 0.2957 |
| | (0.0068) | (0.0574) | (0.1107) | (0.2179) |
| Constant | **32.1588**** | 0.5487 | 0.5487 | -6.9736 |
| | (0.1236) | (5.2745) | (6.4967) | (7.1865) |
| Number of Observations | 194 | 194 | 194 | 194 |
| Adjusted R Squared | 0.2073 | 0.4124 | 0.4124 | 0.3783 |

Note: This table provides the regression coefficients (first row) and standard errors (in parentheses) of the given variables calculated under Ordinary Least Square (OLS), Weighted Least Square (WLS) and Fixed Effect (FE) models with Weight 1 and 2. The dependent variable is PCC of mispricing for OLS and the *t* statistic of PCC (PCC/SE(PCC)) for other models. The OLS and WLS estimations have robust standard errors and FE estimates have robust standard errors clustered by Study_ID.*, ** and *** indicate significance at the 10%, 5% and 1% level, respectively.

Furthermore, only OLS results indicate that there is a negative relationship between publication year and mispricing. This implies that the mispricing declines with the passage of time. The negative coefficients of the Rank_Dummy under OLS and WLS (i.e. -0.44 and

-11.91) suggest that the mean mispricing estimates of the published papers in high ranked journals are lower than the mean mispricing estimates of the published papers in low ranked journals or unpublished papers. In addition, the impact factor shows a positive relationship with mispricing under OLS and WLS models. The higher the impact factor, higher would be the mean mispricing reported in those journals and vice versa. Hence, we conclude that the transaction cost, level of development of the stock market, availability of short selling, publication year, journal rank and impact factor explain the observed heterogeneity in mispricing.

Finally, the results show that there is a positive overall mispricing effect (positive $\beta_0$), after controlling for publication bias and all other explanatory factors as well, under OLS model only. This finding approves the previous finding of PET under OLS in Table 2 in terms of the sign of the relationship.

However, with more precise FE models, we do not find any significant evidence for publication bias or overall effect of mispricing. Under FE models, only the level of stock market development and short selling availability explain the heterogeneity in mispricing significantly.

### *5.3. Publication bias in the relationship between the mispricing and time to maturity*

This section examines whether there is any publication bias in the reported estimates on the relationship between mispricing and time to maturity. The average beta (TTM) is negative 0.28 and the standard deviation is 1.66. Figure 7 depicts the Funnel Plot of beta (TTM) on horizontal axis and the precision (i.e. 1/ standard error of beta (TTM)) on vertical axis. The distribution of beta (TTM) values shows a negative skewness as per the funnel plot. If the funnel plot is skewed to a particular direction, it is assumed that the reported results would also have a publication bias into that direction. Accordingly, in this study, we can expect a negative publication bias with regard to the relationship between the mispricing and time to maturity.

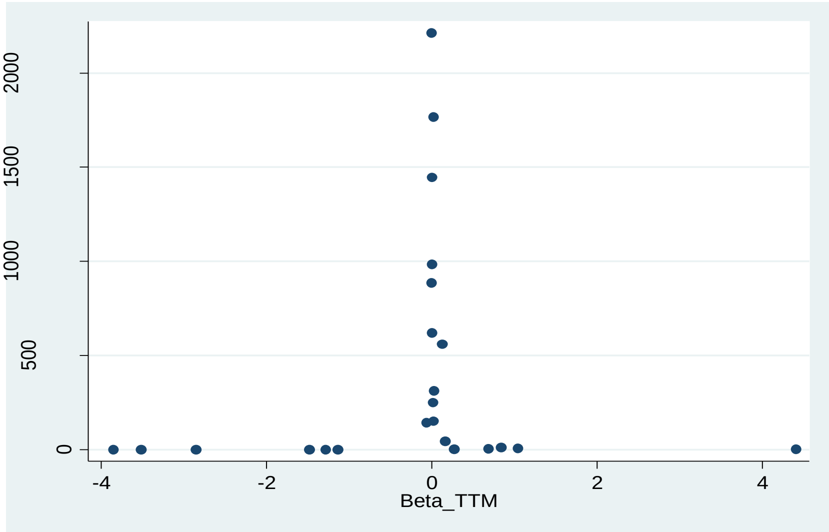


Figure 7: Funnel Plot of time to maturity

Thereafter, we test the publication bias using statistical methods. The results of the regressions based on equations (8) and (9) are presented in Table 4. Theoretically, we expect a mispricing to be disappeared when the time to maturity decreases.

Table 4: Funnel Asymmetry Test and Precision Effect Test results of time to maturity

| Variable | OLS | WLS | FE (Weight 1) | FE (Weight 2) |
|---|---|---|---|---|
| Publication Bias (FAT) $\beta_1$ | **-1.9656***** | 0.0100 | 0.0100 | 0.0086 |
| | (0.4576) | (0.0056) | (0.0064) | (0.0068) |
| Genuine Effect (PET) $\beta_0$ | **0.4309*** | 3.6800 | 3.6800 | 7.1061 |
| | (0.2211) | (2.9647) | (3.6051) | (5.3171) |
| Number of Estimates | 23 | 23 | 23 | 23 |
| Adjusted R Square | 0.3141 | 0.0968 | 0.0968 | 0.0145 |

Note: This table reports coefficients of $\beta_1$ (FAT) and $\beta_0$ (PET), respectively. In both cases, the top value is the coefficient estimate and the value in parentheses is the associated standard error. The first column provides the Ordinary Least Square (OLS) estimates, second column provides Weighted Least Square (WLS) estimates (weighted by the inverse of standard error) and the last columns provide Fixed Effect (FE) estimates with weight 1 (each estimate is given an equal weight) and with weight 2 (each study is given an equal weight). OLS and WLS estimations have robust standard errors and FE estimates have robust standard errors clustered by Study_ID. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

We reject the hypothesis that there is no publication bias at 1% significance level only under OLS results. Accordingly, OLS results suggest that there is a negative publication bias related to the relationship between the mispricing and time to maturity. It indicates that either researchers are more prone to publish negative effect of time to maturity on mispricing or peer-reviewed journals tend to prefer publishing papers providing evidence on a negative effect of time to maturity on mispricing. Alternatively, in this dataset, this negative effect of time to maturity could be driven by the emerging markets. This study finds that emerging markets report a negative beta (TTM) value of -1.06 on average, whereas developed markets report a positive beta (TTM) value of 0.56 on average.

The previous literature provides evidence of positive effect of time to maturity on mispricing in developed markets specifically in USA (MacKinlay and Ramaswamy, 1988), in Germany (Buhler and Kempf, 1995), in UK (Yadav and Pope, 1990), in Switzerland (Stulz et al., 1990) and in Australia (Cummings and Frino, 2011).  In contrast, Gay and Jung (1999) reports a large negative effect of time of maturity on mispricing in Korean market (which is an emerging market). In the next section, we added Developed_Dummy to test whether the effect of time to maturity differs between developed and emerging markets.

In addition, the PET shows that there is a positive overall effect from time to maturity on mispricing under OLS model only. This finding confirms the theoretical expectation about the sign of the effect of time to maturity on mispricing.

### *5.4. Meta-regression analysis of the factors determining the relationship between mispricing and time to maturity*

Table 5 summarizes the results of meta-regression analysis, including six other variables that characterize the differences in data and publication. Except OLS estimates, all other estimators provide evidence on positive publication bias after controlling for other explanatory variables. This suggests that there is a tendency to publish studies with positive beta coefficients on time to maturity. However, it is noteworthy to mention that this positive and significant bias is created statistically due to introducing other explanatory variables into this regression.

Furthermore, there is no evidence that there is an overall effect of time to maturity on mispricing because the value of the constant is positive but not statistically significant.

Table 5: Meta-regression analysis of the relationship between mispricing and time to maturity

| Variable | OLS | WLS | FE (Weight1) | FE (Weight 2) |
|---|---|---|---|---|
| SE of TTM | 1.7128 | **16.1355***** | **16.1355***** | **16.1276***** |
| | (2.1769) | (1.1623) | (1.1762) | (1.2055) |
| SS_Dummy | -0.1168 | **-0.1097***** | **-0.1097***** | **-0.1095***** |
| | (0.2460) | (0.0153) | (0.0165) | (0.0169) |
| TC_Dummy | -0.1973 | **-0.0240**** | **-0.0240**** | **-0.0239**** |
| | (0.6452) | (0.0091) | (0.0096) | (0.0099) |
| Developed_Dummy | 4.2711 | **5.5066*** | **5.5066*** | **5.4750*** |
| | (2.6356) | (2.6090) | (2.7310) | (2.7993) |
| Rank_Dummy | **-4.3922*** | **-6.1845**** | **-6.1845**** | **-6.1557**** |
| | (2.1588) | (2.4360) | (2.4917) | (2.5540) |
| Pub_Year | -0.0183 | **-0.0076***** | **-0.0076***** | **-0.0076***** |
| | (0.0356) | (0.0007) | (0.0007) | (0.0007) |
| Impfac | -0.0599 | **-0.0762***** | **-0.0762***** | **-0.0761***** |
| | (0.2340) | (0.0057) | (0.0061) | (0.0063) |
| Constant | 37.0265 | 2.8135 | 2.8135 | 2.7848 |
| | (72.0443) | (2.2991) | (2.4785) | (2.5405) |
| Number of Observations | 16 | 16 | 16 | 16 |
| Adjusted R Squared | 0.5929 | 0.9578 | 0.9578 | 0.9483 |

Note: This table provides the regression coefficients (first row) and standard errors (in parentheses) of the given variables calculated under Ordinary Least Square (OLS), Weighted Least Square (WLS) and Fixed Effect (FE) models with Weight 1 and 2. The dependent variable is PCC of mispricing for OLS and the *t* statistic of PCC (PCC/SE(PCC)) for other models. The OLS and WLS estimations have robust standard errors and FE estimates have robust standard errors clustered by Study_ID.*, ** and *** indicate significance at the 10%, 5% and 1% level, respectively.

According to these results, all the additional explanatory variables introduced into this regression are significant under WLS and FE models. Hence, we can interpret each variable as follows. The markets with short selling availability will report a 0.1097 lower effect of TTM on mispricing on average than a market with short selling restriction. The studies adjusting for transaction cost will report a 0.02 lower effect of TTM on mispricing on average than a study which does not adjust for transaction cost. A developed market will report a higher effect of TTM on mispricing than an emerging market. This could be driven by the fact that developed markets in general report a positive beta (TTM) and emerging markets in general report a negative beta (TTM) in this case.

Furthermore, this study finds that high ranked journals report a lower effect of TTM on mispricing compared to low ranked journals or journals without a rank. With the passage of time (in this study with the passage of the publication year), studies show a lower effect of TTM on mispricing. Finally, the impact factor shows a negative relationship with TTM. It means that higher the impact factor of a journal, lower the effect of TTM on mispricing they report. In conclusion, the findings suggest that these control variables included to capture the characteristics of data and publication proved to be good explanatory variables of the heterogeneity in beta (TTM).

## 6. Conclusion

This study aims to understand the overall significance of mispricing on futures markets on stock indices. We review past studies and analyze them using meta-analysis techniques. Since previous literature finds mixed evidence on the existence of mispricing among different countries, it is vital to conduct a meta-regression analysis on stock index arbitrage to reach to a single conclusion about these findings. We collect mean mispricing estimates and beta (TTM) from 36 studies. The sample includes studies from developed and emerging markets, studies published in both high ranked and low ranked journals, studies published in journals with a high impact factor or a low impact factor and studies covering the period from 1987 to 2016.

The results of our study indicate that there is no publication bias when reporting the empirical results of the significance of mispricing, except in the case of WLS estimations. Under WLS, there is a significant positive publication bias in mispricing. It suggests that there is a high chance that the studies reporting positive mispricing estimates would be published frequently. The OLS results show that the overall effect of mispricing is medium in size, positive and significant. However, after controlling for the characteristics of data and publication, there is no evidence on any publication bias related to mispricing estimates. Conversely, with more precise FE models, only short selling availability and the level of development of the stock market significantly explain the variation in mispricing estimates.

The results on the relationship between mispricing and time to maturity suggest that there is negative publication bias meaning that there is a tendency to disproportionately publish studies reporting a negative effect of time to maturity on mispricing. However, the overall effect of time to maturity on mispricing is positive under OLS estimations. This finding confirms the theoretical expectation about the relationship between time to maturity and mispricing. Nonetheless, this result is not that convincing, as better models of regression using WLS and FE estimators do not prove this positive publication bias or the overall effect. Furthermore, we conclude that the characteristics of data and publication are good explanatory variables regarding the differences in reported beta values of time to maturity.

Finally, it is noteworthy to mention that there are certain limitations in our meta-analysis study. First, there are only a few studies analyzing stock index arbitrage. Second, even those studies do not have a consistency in reporting the statistical findings. This ultimately lead to a very limited sample of comparable studies available for our analysis. This problem seems to be common for academic research in the area of finance, as each study on a given topic is different and direct comparison is often challenging.

**Acknowledgements**
We are grateful to Maurice Peat, John Goodell, Jarko Fidrmuc and participants at the Symposium on Meta-Analysis and Systematic Reviews in International Finance held on 9th June 2018 at Poznan, Poland, for their valuable comments and suggestions. In addition, we appreciate the support rendered by Bob Reed at the University of Canterbury, New Zealand, in order to understand the meta-analysis methodology.

## References


Andreou, P.C., & Pierides, Y.A. (2008). Empirical investigation of stock index futures market efficiency: The case of the Athens Derivatives Exchange. *European Journal of Finance*, 14(3), 211–223. http://doi.org/10.1080/13518470801890768

Arestis, P., Chortareas, G., & Magkonis, G. (2015). The financial development and growth nexus: A meta-analysis. *Journal of Economic Surveys*, 29(3), 549–565.

Asongu, S.A. (2015). Finance and growth: new evidence from meta-analysis. *Managerial Finance,* 41(6), 615-639. https://doi.org/10.1108/MF-09-2013-0249

Astakhov A., Havranek T., Novak J. (2017). Firm size and stock returns: A meta-analysis. IES Working Paper 14/2017. IES FSV. Charles University.

Berglund, T., & Kabir, R. (2003). What explains the difference between the futures' price and its "fair" value? Evidence from the Euronext Amsterdam. *Research in Banking and Finance*, 3,1−11.

Bhatt, S., & Cakici, N. (1990). Premiums on stock index futures - some evidence. *The Journal of Futures Markets*, *10*(4), 367–375. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/fut.3990100405/abstract

Bialkowski, J., & Jakubowski, J. (2008). Stock index futures arbitrage in emerging markets: Polish evidence. *International Review of Financial Analysis*, *17*, 363–381.

Brailsford, T., & Hodgson, A. (1997). Mispricing in stock index futures: A re-examination using the SPI. *Australian Journal of Management*, *22*(1), 21–45.

Brenner, M., Subrahmanyam, M.G., & Uno, J. (1989a). Arbitrage opportunities in the Japanese stock and futures markets. *Financial Analysts Journal*, 46(2), 14-24.

Brenner, M., Subrahmanyam, M.G., & Uno, J. (1989b). The behavior of prices in the Nikkei spot and futures market. *Journal of Financial Economics*, 23, 363–383.

Brenner, M., Subrahmanyam, M.G. & Uno, J. (1990). Arbitrage opportunities in the Japanese stock and futures markets. *Financial Analysts Journal*, 2, 14–24.

Bühler, W., & Kempf, A. (1995). DAX index futures: Mispricing and arbitrage in German markets. *Journal of Futures Markets*, 15, 833−859.

Card, D., & Krueger, A.B. (1995). Time-series minimum-wage studies: A meta-analysis. *The American Economic Review*, 852, 238-243.

Chung, Y.P. (1991). A transactions data test of stock index futures market efficiency and index arbitrage profitability. *The Journal of Finance*, 46(5), 1791-1809.

Chung, Y.P., Kang, J.K., & Rhee, S.G. (2003). Index futures arbitrage in Japan. In The Japanese Finance: Corporate Finance and Capital Markets in ... Published online: 09 Mar 2015; 173-197.

Cornell, B., & French, K. (1983). The pricing of stock index futures. *Journal of Futures Markets,* 3, 1−14.

Cummings, J.R., & Frino, A. (2011). Index arbitrage and the pricing relationship between Australian stock index futures and their underlying shares. *Accounting and Finance*, 51, 661-683.

Doucouliagos, H. (2011). How large is large? Preliminary and relative guidelines for interpreting partial correlations in economics. Working Papers eco_2011_5, Deakin University, Department of Economics.

Dwyer, G.P., Locke, P., & Yu, W. (1996). Index arbitrage and nonlinear dynamics between the SP500 futures and cash. *Review of Financial Studies*, 9, 301–332.

Egger, M., Smith, G.D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315, 629-634.

Ewijk, C.V., de Groot, H.L.F., & Santing, A.J.C. (2012). A meta-analysis of the equity premium. *Journal of Empirical Finance*, 19, 819-830.

Fassas, A.P. (2011). Mispricing in stock index futures markets: The case of Greece. *Investment Management and Financial Innovations*, 8(2), 103–109.

Fidrmuc, J. & Lind, R. (2018). Macroeconomic impact of Basel III: Evidence from a meta-analysis. *Journal of Banking and Finance*, 1-14.

Figlewski, S. (1984). Hedging performance and basis risk in stock index futures. *Journal of Finance*, 39, 657−669.

Fung, J.K.W., & Draper, P. (1999). Mispricing of index futures contracts and short sales constraints. *Journal of Futures Markets*, 19(6), 695–715.

Fung, J.K.W., & Lam, K. (2004). Overreaction of index futures in Hong Kong. *Journal of Empirical Finance*, 11, 331-351.

Gay, G.D., & Jung, D.Y. (1999). A further look at transaction costs, short sale futures market efficiency: The case of Korean stock index futures. *The Journal of Futures Markets*, *19*(2), 153–174.

Gould, F.J. (1988). Stock index futures: The arbitrage cycle and portfolio insurance. *Financial Analyst Journal*, 48,48−62.

Kempf, A. (1998). Short selling, unwinding, and mispricing. *The Journal of Futures Markets*, 18(8), 903–923.

Klemkosky, R.C., & Lee, J.H. (1991). The intraday ex-post and ex-ante profitability of index arbitrage. *The Journal of Futures Markets*, 11(3), 291–311.

Lim, K. (1992). Arbitrage and price behaviour of the Nikkei stock index futures. *Journal of Futures Markets*, 12, 151−162.

Lin, E., Lee, C.F., & Wang, K. (2013). Futures mispricing, order imbalance, and short-selling constraints. *International Review of Economics and Finance*, 25, 408–423.

MacKinlay, A.C., & Ramaswamy, K. (1988). Index futures arbitrage and the behavior of stock index futures prices. *The Review of Financial Studies*, 1(2), 137–158.

Marcinkiewicz, E. (2016). Short sale and index futures mispricing: Evidence from the Warsaw Stock Exchange. *Prague Economic Papers, 25*(5), 547–559. http://doi.org/10.18267/j.pep.579

Modest, D.M., & Sundaresan, M. (1983). The relationship between spot and futures prices in the stock index futures market: Some preliminary evidence. *Journal of Futures Markets*, 3, 15−41.

Neal, R. (1996). Direct tests of index arbitrage models. *Journal of Financial and Quantitative Analysis*, 31, 541−562.

Peters, E. (1985). The growing efficiency of index futures markets. *Journal of Portfolio Management*, 52−56.

Pope, P.F., & Yadav, P.K. (1994). The impact of short sales constraints on stock index futures prices: Evidence from FT-SE 100 futures. *Journal of Derivatives*, 15-26.

Puttonen, V. (1993). Stock index futures arbitrage in Finland: Theory and evidence in a new market. *European Journal of Operational Research*, 68, 301–317.

Vipul. (2005). Temporal variation in futures mispricing. *Vikalpa*, 30(4), 25-38. https://doi.org/10.1177/0256090920050403

Rusnak, M., Havranek, T., & Horvath, R. (2013). How to solve the price puzzle? A meta-analysis. *Journal of Money, Credit and Banking*, 45 (1), 37–70.

Shemilt, I., Mugford, M., Vale, L., Marsh, K. & Donaldson, C. (2011). *Evidence-based decisions and economics: Health care, social welfare, education, and criminal justice*, 2nd edition. New York: John Wiley & Sons.

Stanley, et al. (2013). Meta-analysis of economics research reporting guidelines. *Journal of Economic Surveys*, 27(2), 390-394.

Stanley, T.D. (2005). Beyond publication bias. *Journal of Economic Surveys*, 19(3), 309-345.

Stanley, T.D. (2008). Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection. *Oxford Bulletin of Economics and statistics*, *70*(1), 103-127.

Stanley, T.D., & Doucouliagos, H. (2012). *Meta-regression analysis in economics and business*. London: Routledge.

Stoll, H.R., & Whaley, R.E. (1986). Expiration day effects of index options and futures. Monograph 1986-3, Salomon Brothers Center for the Study of Financial Institutions, Graduate School of Business Administration, New York University.

Stulz, R., Wasserfallen, W., & Stucki, T. (1990). Stock index futures in Switzerland: Pricing and hedging performance. *Review of Futures Markets*, 9(3), 576−592.

Wang, J. (2010). Short selling and index arbitrage profitability: Evidence from the SGX MSCI and TAIFEX Taiwan index futures markets. *Emerging Markets Finance and Trade*, 46(5), 48–66. http://doi.org/10.2753/REE1540-496X460504

Yadav, P.K., & Pope, P.F. (1990). Stock index futures arbitrage: International evidence. *The Journal of Futures Markets*, *10*(6), 573–603. http://doi.org/10.1002/fut.3990100603

Zigraiova, D. & Havranek, T. (2016). Bank competition and financial stability: Much ado about nothing? *Journal of Economic Surveys*, 30(5), 944-981.

**Appendix 1.**
**List of selected papers**


Andreou, P.C., & Pierides, Y.A. (2008). Empirical investigation of stock index futures market efficiency: The case of the Athens Derivatives Exchange. *European Journal of Finance*, 14(3), 211–223. http://doi.org/10.1080/13518470801890768

Bhatt, S., & Cakici, N. (1990). Premiums on stock index futures - some evidence. *The Journal of Futures Markets*, 10(4), 367–375. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/fut.3990100405/abstract

Bialkowski, J., & Jakubowski, J. (2008). Stock index futures arbitrage in emerging markets: Polish evidence. *International Review of Financial Analysis*, 17, 363–381. http://doi.org/10.1016/j.irfa.2006.10.004

Brailsford, T., & Hodgson, A. (1997). Mispricing in stock index futures: A re-examination using the SPI. *Australian Journal of Management*, 22(1), 21–45. http://doi.org/10.1177/031289629702200102

Brenner, M., Subrahmanyam, M.G., & Uno, J. (1989). Stock index futures arbitrage in the Japanese markets. *Japan and The World Economy*, 1(3), 303–330. http://doi.org/10.1016/0922-1425(89)90005-4

Brenner, M., Subrahmanyam, M.G., & Uno, J. (1989). The behavior of prices in the Nikkei spot and futures market. *Journal of Financial Economics*, 23, 363–383.

Brenner, M., Subrahmanyam, M.G. & Uno, J. (1990). Arbitrage opportunities in the Japanese stock and futures markets. *Financial Analysts Journal*, 2, 14–24. http://doi.org/10.2469/faj.v46.n2.14

Buhler, W., & Kempf, A. (1995). DAX index futures: Mispricing and arbitrage in German markets. *Journal of Futures Markets*, 15(7), 833–859.

Chen, A., Chianglin, C., & Chung, H. (2001). Establishing an index arbitrage model by applying neural networks method - A case study of Nikkei 225 Index. *International Journal of Neural Systems*, 11(5), 489–496. http://doi.org/10.1002/9780470015902.a0003656.pub2

Chung, Y.P., Kang, J.K., & Rhee, S.G. (2003). Index futures arbitrage in Japan. *In The Japanese Finance: Corporate Finance and Capital Markets in ... Published online:* 09 Mar 2015; 173-197.

Chung, Y.P. (1991). A transactions data test of stock index futures market efficiency and index arbitrage profitability. *The Journal of Finance*, 46(5), 1791-1809.

Cummings, J.R., & Frino, A. (2011). Index arbitrage and the pricing relationship between Australian stock index futures and their underlying shares. *Accounting and Finance*, 51, 661-683.

Deaves, R. (1994). A note on speculative versus arbitrage opportunities from index futures mispricing: Some Canadian evidence. *International Review of Economics and Finance*, 3(3), 319–325. http://doi.org/10.1016/1059-0560(94)90014-0

Fassas, A.P. (2011). Mispricing in stock index futures markets: The case of Greece. *Investment Management and Financial Innovations*, 8(2), 103–109.

Fung, J.K.W., & Lam, K. (2004). Overreaction of index futures in Hong Kong. *Journal of Empirical Finance*, 11, 331-351.

Fung, J.K.W., & Draper, P. (1999). Mispricing of index futures contracts and short sales constraints. *Journal of Futures Markets*, 19(6), 695–715. http://doi.org/10.1002/(SICI)1096-9934(199909)19:6<695::AID-FUT4>3.0.CO;2-H

Gay, G.D., & Jung, D.Y. (1999). A further look at transaction costs, short sale futures market efficiency: The case of Korean stock index futures. *The Journal of Futures Markets*, 19(2), 153–174.

Henker, T., & Martens, M. (2005). Index futures arbitrage before and after the introduction of sixteenths on the NYSE. *The Journal of Empirical Finance*, 12, 353–373. http://doi.org/10.1016/j.jempfin.2004.04.006

Kempf, A. (1998). Short selling, unwinding, and mispricing. *The Journal of Futures Markets*, 18(8), 903–923.

Klemkosky, R.C., & Lee, J.H. (1991). The intraday ex-post and ex-ante profitability of index arbitrage. *The Journal of Futures Markets*, 11(3), 291–311.

Lim, K. (2000). Arbitrage and price behavior of the Nikkei stock index futures. *The Journal of Futures Markets*, 12(2), 151–161.

Lin, E., Lee, C.F., & Wang, K. (2013). Futures mispricing, order imbalance, and short-selling constraints. *International Review of Economics and Finance*, 25, 408–423. http://doi.org/10.1016/j.iref.2012.08.001

MacKinlay, A.C., & Ramaswamy, K. (1988). Index futures arbitrage and the behavior of stock index futures prices. *The Review of Financial Studies*, 1(2), 137–158.

Marcinkiewicz, E. (2016). Short Sale and Index Futures mispricing: Evidence from the Warsaw Stock Exchange. *Prague Economic Papers*, 25(5), 547–559. http://doi.org/10.18267/j.pep.579

Merrick, J.J. (1987). Volume determination in stock and stock index futures markets: An analysis of arbitrage and volatility effects. *The Journal of Futures Markets*, 7(5), 483–496.

Merrick, J.J. (1989). Early unwindings and rollovers of stock index futures arbitrage programs: Analysis and implications for predicting expiration day effects. *The Journal of Futures Markets*, 9(2), 101–111. http://doi.org/10.1002/fut.3990090203

Nam, S.O., Kim, H.K., & Kim, B.C. (2010). An alternative approach to evaluating the agreement between financial markets. *Journal of International Financial Markets, Institutions and Money*, 20(1), 13–35. http://doi.org/10.1016/j.intfin.2009.11.001

Neal, R. (2016). Direct tests of index arbitrage models. *The Journal of Financial and Quantitative Analysis*, 31(4), 541–562.

Puttonen, V. (1993). Stock index futures arbitrage in Finland: Theory and evidence in a new market. *European Journal of Operational Research*, 68, 301–317.

Tu, A.H., Hsieh, W.G., & Wu, W. (2016). Market uncertainty, expected volatility and the mispricing of S&P 500 index futures. *Journal of Empirical Finance*, 35, 78–98. http://doi.org/10.1016/j.jempfin.2015.10.006

Tungsong, S., & Srijuntongsiri, G. (2011). Index arbitrage and futures pricing efficiency: Evidence from Thailand (August 23, 2011). 24th Australasian Finance and Banking Conference 2011 Paper. Available at SSRN: https://ssrn.com/abstract=1915282. or http://dx.doi.org/10.2139/ssrn.1915282

Vipul (2005). Temporal variation in futures mispricing. *Vikalpa*, 30(4), 25-38. https://doi.org/10.1177/0256090920050403

Wang, J. (2010). Short selling and index arbitrage profitability: Evidence from the SGX MSCI and TAIFEX Taiwan index futures markets. *Emerging Markets Finance and Trade*, 46(5), 48–66. http://doi.org/10.2753/REE1540-496X460504

Yadav, P.K., & Pope, P.F. (1990). Stock index futures arbitrage: International evidence. *The Journal of Futures Markets*, 10(6), 573–603. http://doi.org/10.1002/fut.3990100603

Yadav, P. K., & Pope, P. F. (1994). Stock index futures mispricing: Profit opportunities or risk premia? *Journal of Banking and Finance*, 18, 921–953.

Zhuo, W., Zhao, X., Zhou, Z., & Wang, S. (2012). Study on Stock index futures' mean reversion effect and arbitrage in China based on high frequency data. *iBusiness*, 4, 78–83. http://doi.org/10.4236/ib.2012.41009

**Appendix 2. List of variables coded from the selected studies**

| Variable | Definition |
|---|---|
| Study_ID | Identification number allocated for each study |
| Est_ID | Identification number allocated for each estimate in each study |
| Pub_Year | Publication year |
| J | Published journal |
| Rank | Journal rank (Based on ABDC ranking 2017) |
| Rank_Dummy | Rank_Dummy = 1 if A and A* or = 0 if B and C |
| Impfac | Impact factor of the journal obtained from https://ideas.repec.org |
| DS | Source of data |
| E | Exchange |
| Developed | Developed = 1 if developed or = 0 if emerging (Based on MSCI 2017 Market Classification) |
| BI | Benchmark index |
| SP-From | Sample period – From |
| SP-To | Sample period – To |
| SS_Dummy | SS_Dummy= 1 if short selling is available or = 0 if restricted |
| TC_Dummy | TC_Dummy = 1 if transaction cost is adjusted or = 0 if not adjusted |
| Mean (%) | Mean mispricing percentage |
| ABS_Mean (%) | Mean absolute mispricing percentage |
| Mean+ (%) | Mean positive mispricing percentage |
| Mean-(%) | Mean negative mispricing percentage |
| SD_Mean% | Standard deviation of mean mispricing percentage |
| SD_AbsMean% | Standard deviation of mean absolute mispricing percentage |
| SD_Mean+% | Standard deviation of mean positive mispricing percentage |
| SD_Mean-% | Standard deviation of mean negative mispricing percentage |
| SE_Mean% | Standard error of mean mispricing percentage |
| SE_AbsMean% | Standard error of mean absolute mispricing percentage |
| SE_Mean+% | Standard error of mean positive mispricing percentage |
| SE_Mean-% | Standard error of mean negative mispricing percentage |
| T_Val | *t* statistic of mean mispricing percentage |
| Mis+Freq | Number of positive mispricing observations |
| Mis-Freq | Number of negative mispricing observations |
| Tot Freq | Total number of mispricing observations |
| β(TTM) | Beta coefficient of time to maturity |
| SE(TTM) | Standard error of the beta coefficient of time to maturity |
| SD(TTM) | Standard deviation of the beta coefficient of time to maturity |
| t(TTM) | *t* statistic of the beta coefficient of time to maturity |