Global Economy and Finance Journal
Vol. 4. No. 1. March 2011 Pp. 32-43

Cash Flow, Earnings Ratio and Stock Returns in Emerging Global Regions: Evidence from Longitudinal Data

Luiz Paulo Fávero[1] and Patrícia Belfiore[2]

Some investors prefer to use cash flow instead of earnings per share to evaluate stocks' current prices, and they argument that, while the first is not easily manipulated, the same cannot be said for earnings. Based on a sample from Compustat Global, including 3,567 stocks of companies from 35 emerging countries, covering 118 months (1998-2007), totalizing 218,530 observations, this study applies panel data models with different estimators to verify that price to cash flow ratio is more significant to influence returns over time, with more efficient estimators for the fixed effect model. Moreover, different regressors are obtained for each global region.

JEL Codes: C01, C23, G15.

1. Introduction

Differences in stock price behaviors among developed countries have been extensively researched and documented (Bondt, 2008). Investigations of this nature for emerging economies are, however, much less comprehensive. Important works include Islam (1995), Kwon et. al. (1997), Lee (1997), Tsoukalas (2003), Karim and Guan (2004), Al-Khazali and Pyun (2004), Makina (2008), Valadkhani and Chancharat (2008), Al-Khazali (2008), Rutledge et al. (2008), Valadkhani et al. (2009) and Huang et al. (2009).

This study applies some panel data models to investigate how monthly returns of stock prices are influenced by price to cash flow ratio and price to earnings ratio in emerging economies over time. Both ratios are used by investors to evaluate the investment attractiveness, from a value standpoint, of a company's stock, but some analysts prefer to consider cash flow over earnings because it is not easily manipulated is not affected by depreciation and other non-cash factors.

According to Kennon (2010), price to cash flow ratio provides a better idea of the amount of money available to management for further research and development, marketing support, debt reductions, dividends, share repurchases, and more. On the other hand, price to earnings ratio is not useful to compare companies in different industries, as each industry has much different growth prospects.

Data models with different estimators are utilized to attempt to answer the following two questions: (i) how do returns of stock prices of companies in emerging

[1] School of Economics, Business and Accountancy, University of São Paulo. Av. Prof. Luciano Gualberto, 908, FEA 3, Cidade Universitária, São Paulo, SP, Brazil. Zip Code: 05508-900.
E-mail: lpfavero@usp.br

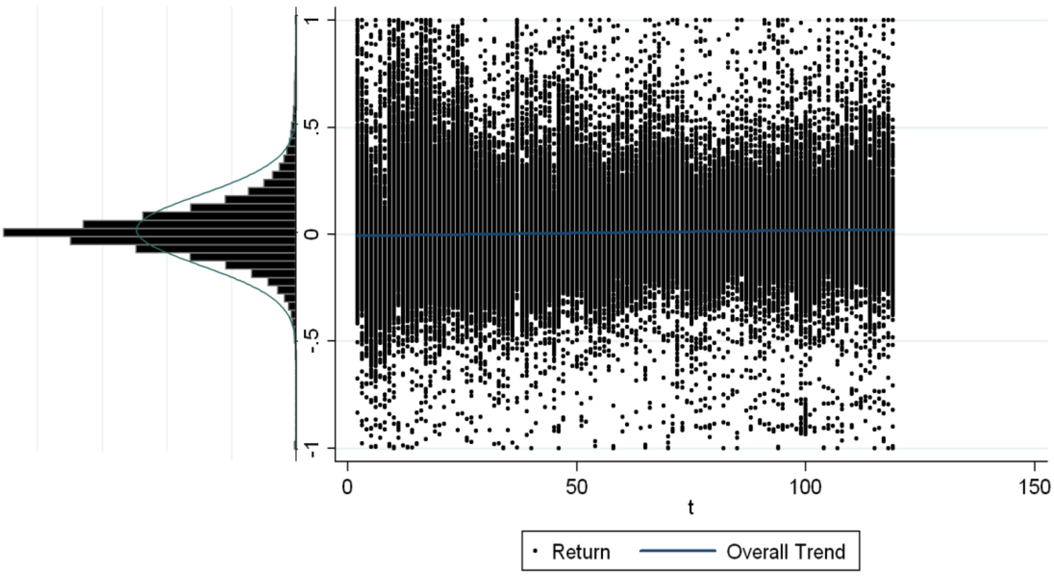
[2] Department of Management Engineering, Federal University of ABC. Av. Prof. Luciano Gualberto, 908, FEA 3, Cidade Universitária, São Paulo, SP, Brazil. Zip Code: 05508-900.
E-mail: patricia.belfiore@pq.cnpq.br

countries react to price to cash flow ratio and to price to earnings ratio over time?; and (ii) is there a “region effect” on returns of stock prices? For this purpose, a *Compustat Global* sample of 3,567 stocks of companies from 35 countries was statistically examined from a temporal perspective, through a period of 118 months (1998-2007).

Section 2 presents a diagnosis of the returns of stock prices of companies in these 35 emerging countries, focusing on variations over time. Section 3 presents a review of panel data models and its estimators. Section 4 presents and discusses the empirical results. Conceptual inferences and suggestions for further empirical works are presented in the last section.

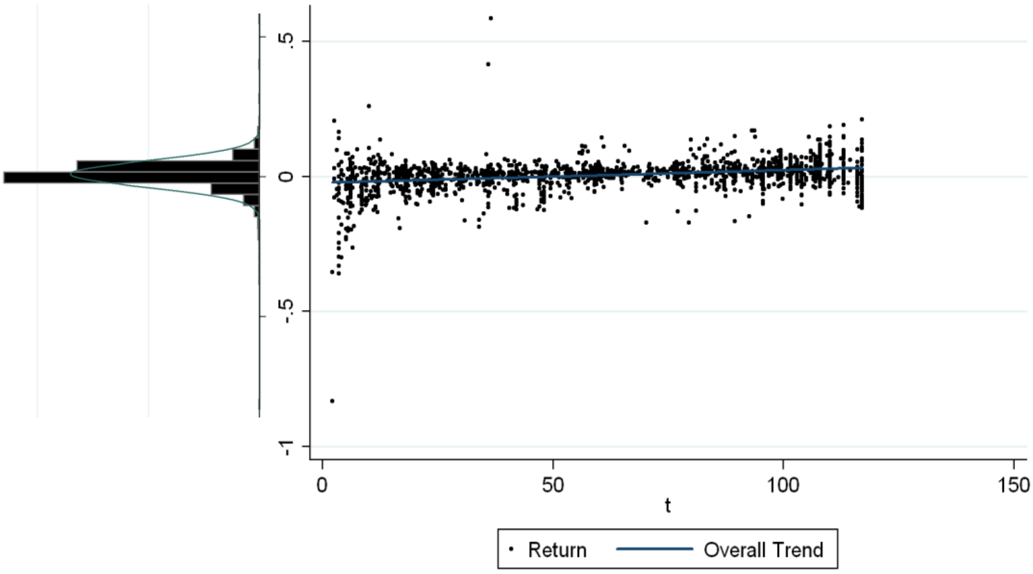
2. Returns of Stock Prices in Emerging Countries

In the period 1998-2007, many companies presented expressive growth rates in stock prices during one or more months. However, as graph 1 reveals, we considered only returns below 100% in one month. It’s possible to verify that these returns of stock prices reveal similar behavior over time, despite occasional differences in means and inclinations among individual series.



Graph 1: Evolution of Returns of Stock Prices in Emerging Countries

Each point in graph 1 represents a stock return-month pair. This behavior suggests the elaboration of longitudinal models and, according to Cameron and Trivedi (2009), many of these microeconometric methods emphasize the existence of correlation over time for individual data, with mutual independence. However, in different cases, like models that consider data from companies active in the same countries, correlations can occur between companies, as shown in graph 2. Hence, it is usually necessary to elaborate corrections in the models’ standard errors through the application of more consistent panel data estimators.



Graph 2: Between Variation of Returns of Stock Prices

The justification to adopt panel data models in this research is also related to the fact that part of return can be invariable over time, while another part, as an occasional general trend, can be invariable among companies. A representative part of return, though, can vary among companies and over time, and this fact can generate different behaviors among global regions, as one can see through graph 3. Section 4 presents the within and between variations of each of the regressors this research will take into account. As discussed further ahead and considered in the models with different types of estimators, the panel data models can permit endogenous regressors, due to the existence of correlation with an error component that is invariable over time (fixed effects), or assume that the regressors are completely exogenous (random effects). Both estimators will be considered in this article.

3. Panel Data Models

Many different models can be used for panel data, and the basic distinction is the existence of fixed or random effects. The term "fixed effects" gives a mistaken idea of the model as, for both cases, the individual (in this case, stock) level effects are random. Hence, according to Cameron and Trivedi (2009), the fixed effects models present the additional complication that the regressors are correlated with the effects of the individual-level effects and, therefore, a consistent estimate of the model parameters demands the elimination or control of fixed effects. Hence, a model that takes into account the individual's specific effects for a dependent variable y_{it} specifies that:

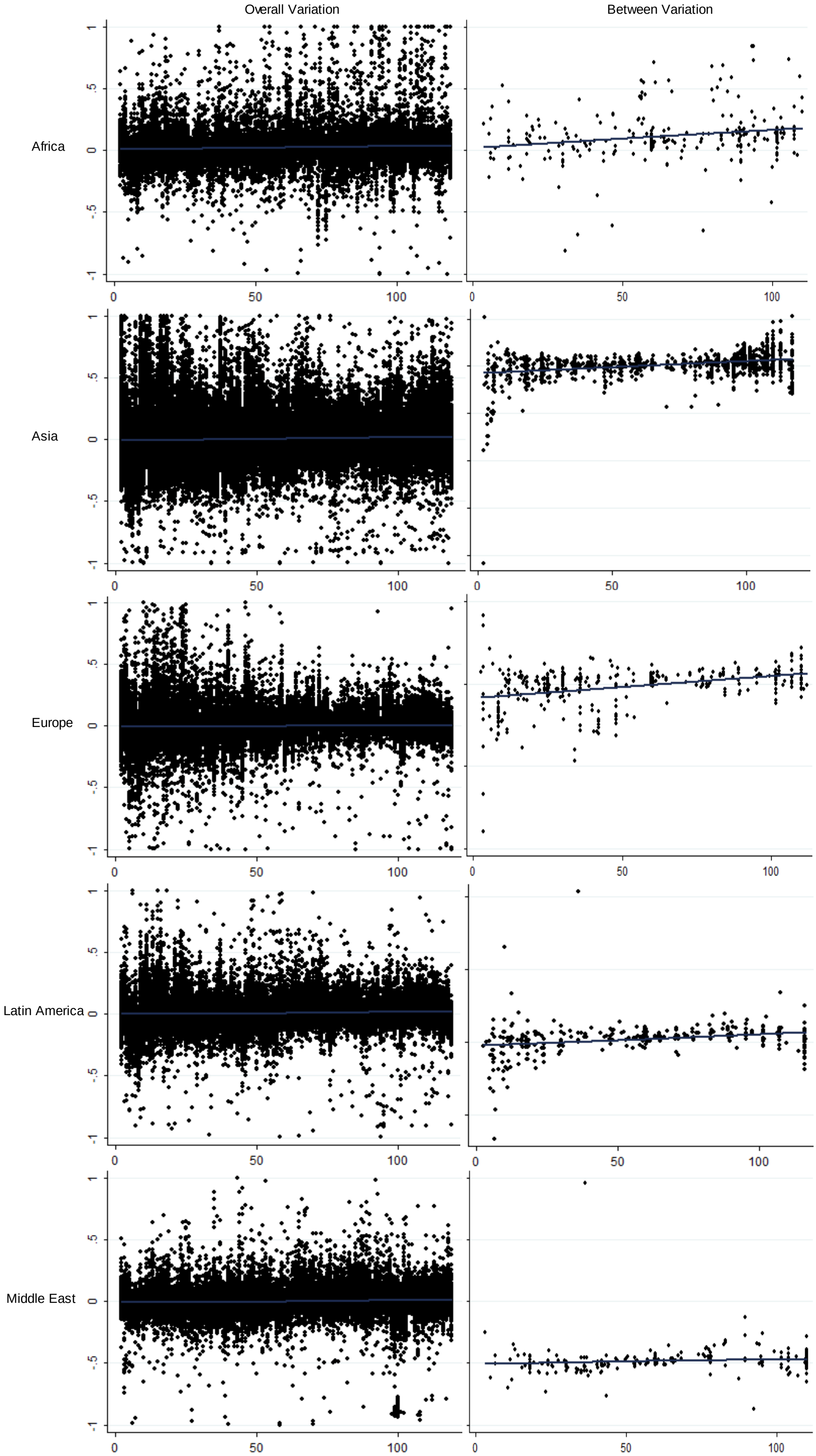
$$y_{it} = \beta_{0i} + x'_{it}\,\beta_1 + \varepsilon_{it} \quad (1)$$

in which x_{it} are regressors, β_{0i} are the individual specific random effects and ε_{it} represents the idiosyncratic error.

Two different models can derive from expression (1). In the fixed effects model, the β_{0i} terms can be correlated with the regressors x_{it}, which permits a limited form of endogeneity. Making the error term $\mu_{it} = \beta_{0i} + \varepsilon_{it}$ and allowing x_{it} to be correlated with the error term invariable over time (β_{0i}), it is assumed that x_{it} is not correlated with the idiosyncratic error ε_{it}. The fixed effects model implies that $E(y_{it}|\beta_{0i}, x_{it}) = \beta_{0i} + x'_{it}\beta_1$, assuming that $E(\varepsilon_{it}|\beta_{0i}, x_{it}) = 0$, so that $\beta_j = \partial E(y_{it}|\beta_{0i}, x_{it})/\partial x_{j,it}$. The advantage of the fixed effects model is that a consistent estimator of the marginal effect of the j-eth regressor of $E(y_{it}|\beta_{0i}, x_{it})$ can be obtained, given that $x_{j,it}$ varies over time, even if the regressors were endogenous (with a limited form of endogeneity).

In the random effects model, on the other hand, it is assumed that β_{0i} is purely random, that is, that it is not correlated with the regressors. Therefore, the estimation is elaborated with an FGLS (feasible generalized least squares) estimator. The advantage of the random effects model is that it estimates all coefficients, even of regressors that do not vary over time and, hence, the marginal effects. Moreover, $E(y_{it}|x_{it})$ can be estimated. The main advantage, however, is that these estimators are inconsistent if the fixed effects model is more appropriate.

While the variation over time or for a given individual is known as within variance, the variation between individuals is called between variance. According to Wooldridge (2002), in the fixed effects model, the coefficient of a regressor with low within variance will be estimated imprecisely and will not be identified if there is no within variance whatsoever. Hence, distinguishing between these variances is of fundamental importance to define the best panel data model.



Graph 3: Overall and Between Variation of Returns per Global Region

According to Cameron and Trivedi (2009), the estimators of the β_1 parameters of a fixed effects model for equation (1) eliminate the fixed effects β_{0i}, that is, a transformation within is elaborated through the distinction among means. Thus, a within estimation elaborates an OLS (ordinary least squares) model with data distinguished around the mean and, as the observations of an invariable regressor over time with differentiation around the mean are zero, the coefficient of a regressor that does not vary over time cannot be estimated. On the other hand, as a within estimator offers a consistent estimate for the fixed effects model, it is frequently called fixed effects estimator.

Hence, the fixed effects β_{0i} in equation (1) can be eliminated by subtracting the means of each individual $\bar{y}_i = \bar{x}'_i \beta_1 + \bar{\varepsilon}_i$ in the corresponding model, resulting either in the within model or the model of differences of means:

$$(y_{it} - \bar{y}_i) = (x_{it} - \bar{x}_i)'\beta_1 + (\varepsilon_{it} - \bar{\varepsilon}_i) \qquad (2)$$

where $\bar{x}_i = T_i^{-1} \sum_{t=1}^{T_i} x_{it}$ and the within estimator is the OLS estimator of this model. Due to the fact that β_{0i} has been eliminated, according to Cameron and Trivedi (2009), the OLS estimator offers consistent estimates of β_1, even if β_{0i} is correlated with x_{it}, like in the case of the fixed effects model.

The between estimator uses only cross-sections variation and is the OLS estimator of a regression of $\bar{y}_i$ in function of $\bar{x}_i$, presented next (equation 3). Due to considering only cross-section data variations, the coefficients of any regressor invariable among individuals cannot be identified.

$$\bar{y}_i = \beta_0 + \bar{x}'_i \beta_1 + (\beta_{0i} - \beta_0 + \bar{\varepsilon}_i) \qquad (3)$$

The consistency of this estimator requires no correlation between the error term $(\beta_{0i} - \beta_0 + \bar{\varepsilon}_i)$ and x_{it}, which occurs when β_{0i} is a random effect, but not when it is a fixed effect.

The random effects estimator, on the other hand, is an FGLS estimator in equation (1), based on the hypothesis that the random effect β_{0i} and the idiosyncratic error ε_{it} are independents and identically distributed. Hence, the random effects model is the individual effects model:

$$y_{it} = x'_{it} \beta_1 + (\beta_{0i} + \varepsilon_{it}) \qquad (4)$$

with β_{0i} ~ (β_0, σ^2_α) and ε_{it} ~ (0, σ^2_μ). Thus, the error term $\mu_{it} = \beta_{0i} + \varepsilon_{it}$ is correlated over time t, for a given observation i, with correlation:

$$\text{corr}(\mu_{it}, \mu_{is}) = \sigma^2_\alpha / (\sigma^2_\alpha + \sigma^2_\varepsilon), \text{ for any } s \neq t \qquad (5)$$

The random effects estimator is the FGLS estimator of β_1 in equation (4), given the error correlations in expression (5). In models with heteroscedastic errors, the FGLS estimator can be calculated as an OLS estimator in a transformed model, so as to obtain non-correlated homoscedastic errors. Hence, the random effects estimator can be obtained through the OLS estimation of the following transformed model:

$$(y_{it} - \hat{\theta}_i \bar{y}_i) = (1 - \hat{\theta}_i)\beta_0 + (x_{it} - \hat{\theta}_i \bar{x}_i)'\beta_1 + \{(1 - \hat{\theta}_i)\beta_{0i} + (\varepsilon_{it} - \hat{\theta}_i \bar{\varepsilon}_i)\} \quad (6)$$

in which $\hat{\theta}_i$ is a consistent estimate of:

$$\theta_i = 1 - \sqrt{\sigma_\varepsilon^2 / (T_i \sigma_\alpha^2 + \sigma_\varepsilon^2)}$$

The random effects estimator uses within and between variances for the data model, in which within estimation is a special case when $\hat{\theta}_i$ = 1. This estimator will be consistent and completely efficient if the random effects model is appropriate, but will be inconsistent if the fixed effects model is appropriate, as the correlation between x_{it} and β_{0i} results in a correlation between the regressors and the error term in expression (6).

Following Cameron and Trivedi (2009), if there are no fixed effects but the errors demonstrate correlation inside the panel, then the random effects estimator will be consistent but inefficient and, therefore, an estimate with robust clusterized standard errors will have to be obtained.

A starting point for the panel data model is the application of the POLS (pooled ordinary least squares) model, which assumes that regressors are exogenous and that the error is μ_{it}, instead of the decomposition $\alpha_i + \varepsilon_{it}$. Hence:

$$y_{it} = \beta_0 + x'_{it}\beta_1 + \mu_{it} \quad (7)$$

This model's parameters are estimated through OLS, but the inference requires control of error μ_{it}'s within correlation for a given individual, to be elaborated by using robust standard errors grouped at individual level.

This study applies 6 different panel data models, to obtain a better understanding of the different types of estimators, and also to present the best model to explain the return of stock prices of companies in emerging countries.

4. Model and Results

As discussed above, 6 panel data models are elaborated, with different considerations about the estimators and error terms. The general model is as follows:

$$\text{return}_{it} = \beta_{0i} + \beta_1.(\text{pcf})_{it} + \beta_2.(\text{pe})_{it} + \varepsilon_{it} \quad (8)$$

in which β_1 and β_2 represent the change in return when one unit in price to cash flow ratio (pcf) or in price to earnings ratio (pe) happens, respectively, *ceteris paribus*.

The panel data model is based on the existence of overall, within and between variances in each regressor, as discussed earlier. Table 1 displays the variance decomposition for each of the regressors.

Table 1: Panel Statistics and Decomposition of Variance

Variable	Decomp.	Mean	Standard Deviation	Minimum	Maximum	Obs.
stock	overall between within		0.000			N.T = 218,530 N = 3,840
month	overall between within		35.210 33.315 26.489	1.000 1.500 -5.657	118.000 118.000 147.977	N.T = 218,530 N = 3,840
return	overall between within	0.010	0.161 0.041 0.159	-1.000 -0.338 -1.189	1.000 0.584 1.085	N.T = 218,530 N = 3,840
pcf	overall between within	26.856	1,150.177 455.778 1,106.678	-21,356.35 -10,170.97 -17,646.37	286,258.9 16,673.07 280,734.9	N.T = 166,750 N = 3,214
pe	overall between within	62.284	1,967.290 750.558 1,849.308	-29,855.40 -5,721.959 -29,624.32	306,869.5 29,721.18 297,903.8	N.T = 210,630 N = 3,567

Since the stock is invariable over time, its within variation is equal to zero. The time variable (month), however, is not invariable among individuals, because this is a strongly unbalanced panel. The within variation of the panel regressors is higher than the between variation, what is the first indicator that the between estimate can result in an efficiency loss in this case. Table 1 also offers further foundations to adopt the panel data models and apply different estimators. Columns "Minimum" and "Maximum" present, respectively, the minimum and maximum values of x_{it} for the "overall" line, $\bar{x}_i$ for the "between" line and $x_{it} - \bar{x}_i + \bar{x}$ for the "within" line.

Table 2 shows the results of each of the six proposed panel data models.

Table 2: Panel Data Models

Variable	POLS Robust Error	Between Estimator	Fixed Effects	Fixed Effects Robust Error	Random Effects	Random Effects Robust Error
pcf	$3.69\text{x}10^{-6}$* ($1.46\text{x}10^{-6}$)	$8.33\text{x}10^{-6}$* ($1.99\text{x}10^{-6}$)	$3.02\text{x}10^{-6}$* ($4.74\text{x}10^{-7}$)	$3.02\text{x}10^{-6}$* ($1.31\text{x}10^{-6}$)	$3.24\text{x}10^{-6}$* ($4.57\text{x}10^{-7}$)	$3.24\text{x}10^{-6}$* ($1.14\text{x}10^{-6}$)
pe	$-1.23\text{x}10^{-7}$ ($3.37\text{x}10^{-7}$)	$5.04\text{x}10^{-7}$ ($8.10\text{x}10^{-7}$)	$-1.66\text{x}10^{-7}$ ($2.62\text{x}10^{-7}$)	$-1.66\text{x}10^{-7}$ ($2.89\text{x}10^{-7}$)	$-1.19\text{x}10^{-7}$ ($2.47\text{x}10^{-7}$)	$-1.19\text{x}10^{-7}$ ($2.13\text{x}10^{-7}$)
constant	0.011* ($4.39\text{x}10^{-4}$)	0.006* ($8.19\text{x}10^{-4}$)	0.011* ($4.03\text{x}10^{-4}$)	0.011* ($2.97\text{x}10^{-5}$)	0.009* ($7.06\text{x}10^{-4}$)	0.009* ($7.21\text{x}10^{-4}$)
N	166,743	166,743	166,743	166,743	166,743	166,743
R^2	$6.0\text{x}10^{-4}$					
R^2 (overall)		$6.0\text{x}10^{-4}$	$6.0\text{x}10^{-4}$	$6.0\text{x}10^{-4}$	$6.0\text{x}10^{-4}$	$6.0\text{x}10^{-4}$
R^2 *(*between)		$7.7\text{x}10^{-3}$	$7.2\text{x}10^{-3}$	$7.2\text{x}10^{-3}$	$7.4\text{x}10^{-3}$	$7.4\text{x}10^{-3}$
R^2 *(*within)		$4.0\text{x}10^{-4}$	$4.0\text{x}10^{-4}$	$4.0\text{x}10^{-4}$	$4.0\text{x}10^{-4}$	$4.0\text{x}10^{-4}$
F	3.79	12.45	30.30	2.97		
sig. F	0.022	0.000	0.000	0.049		
Wald χ^2					76.39	9.27
sig. χ^2					0.000	0.009

Obs.: Standard errors between parentheses.
*   sig. < 0.05.

As it can be observed, estimated coefficients can vary from model to model, which reflects the existence of different results if within or between variations are used.

First, higher standard errors exist in the fixed effects and random effects models with robust clusterized standard errors when compared with the respective models without this consideration. Moreover, the estimated regressors in the POLS and between models result in even higher standard errors, despite the existence of significant regressor coefficients in all models (sig. < 0.05).

As to the adequacy of the models themselves, statistical significance of the set of variables is observed in all cases (sig. F for POLS, between and fixed effects models and sig. Wald χ^2 for random effects models). As R^2 statistics are considerably low in the obtained models, these aren't relatively adequate for prediction purposes.

Breusch-Pagan's LM test, applied after random effects modeling, helps to reject the null hypothesis of the POLS model's adequacy in relation to the random effects model, as χ^2 = 31.94 (sig. χ^2 = 0.000). Next, using Chow's F test, the null hypothesis about the existence of equal intercepts and inclinations for all companies (POLS) was rejected. Therefore, these parameters differ from cross-sections parameters (fixed effects), as F = 1.37 (sig. F = 0.000). Finally, Hausman test for fixed effects, applied after fixed and random effects modeling, helps to reject the null hypothesis that the random effects model offers consistent parameter estimates as, in this case, χ^2 = 16.13 (sig. χ^2 = 0.000).

The regressors explain the returns of stock prices in emerging countries, which sustains the initial study hypothesis. But, through table 2, only price to cash flow ratio is significant to explain these returns in the presence of price to earnings ratio, confirming the argument of some analysts who have preferred to consider cash flow over earnings.

According to Islam (1995), the main utility of panel data modeling is its ability to permit differences among individuals in each country, leading to results that significantly differ from those obtained through isolated regressions for each country. The same can be said to the global regions. In Table 3, regression coefficients are shown for each of the global regions considered in this study.

Table 3: Coefficients per Region

Region	pcf	constant
Africa	$2.86\text{x}10^{-6}$	0.026
Asia	$4.93\text{x}10^{-6}$	0.011
Europe	$3.58\text{x}10^{-6}$	0.003
Latin America	$1.78\text{x}10^{-5}$	0.013
Middle East	$5.06\text{x}10^{-6}$	0.006

Obs.: Dependent variable: Monthly return of stock price.

Although price to cash flow ratio is more significant to explain the return of stock prices in emerging countries, table 3 reveals the existence of different influences. The existence of different coefficients for the regressors and for the values of the constant express the importance of considering panel data and call for further research about the economic reasons why global regions present different behaviors of their companies' stock prices over time.

One can see that all regions offer positive slopes of price to cash flow ratio, but Latin America and Middle East present the highest ones. If greater returns over time could be seen as a proxy for equity capital market development, perhaps as an indicator of market depth, this financial development is, *ceteris paribus*, a contributor "per se" to higher stock prices and gains in shareholder wealth.

5. Conclusions

Panel data models have been increasingly frequent in the assessment of stock markets returns. These models permit one to test the relationship between some performance variable and some predictive variables, allowing the determination of the influences of temporal and regional differences on return evolution.

Previous studies have adopted longitudinal modeling to attempt to explain significant variability in performance variables among companies, countries and over time, with empirical researches developed in different areas (Makino et al. 2004), but it's not common to find longitudinal studies applied to stock markets, considering the differences among global regions. The analysis of the contribution of the price to cash flow ratio and the price to earnings ratio over return arouses a discussion about how are the behaviors of the stocks in countries with less developed financial markets.

This research contribution is also related to the possible assessment of different estimators when panel data modeling is used, with a view to determining the most adequate model, in function of the characterization of a short panel.

Further, the consideration of developed countries, such as US, UK, France or Australia, could bring one the possibility of comparison of returns of stock prices among countries in a broader perspective.

References

Al-Khazali, OM 2008, 'The impact of thin trading on day-of-the-week effect', *Review of Accounting and Finance*, vol. 7, no. 3, pp. 270-284.

Al-Khazali, OM & Pyun, CS 2004, 'Stock prices and inflation: new evidence from the Pacific-Basin countries', *Review of Quantitative Finance and Accounting*, vol. 22, no. 2, pp. 123-140.

Bondt, GJ 2008, 'Determinants of stock prices: new international evidence', *Journal of Portfolio Management*, vol. 34, no. 3, pp. 81-92.

Cameron, AC & Trivedi, PK 2009, *Microeconometrics using Stata*, Stata Press, College Station, Texas.

Huang, YS, Liu, DY & Lin, FY 2009, 'Short-sales constraints and stock price behavior: evidence from the Taiwan stock exchange', *The Business Review*, vol. 13, no. 1, pp. 93-98.

Islam, N 1995, 'Growth empirics: a panel data approach', *The Quarterly Journal of Economics*, vol. 110, no. 4, pp. 1127-1170.

Karim, MZA & Guan, TB 2004, 'Stock prices, foreign opportunity cost, and money demand in Malaysia: a cointegration and error correction model approach', *Jurnal Ekonomi Malaysia*, vol. 38, pp. 29-62.

Kennon, J 2010, *Price to cash flow ratio*, viewed 07 August 2010, <http://beginnersinvest.about.com/od/financialratio/a/pricetocashflow.htm>.

Kwon, CS, Shin, TS & Bacon, FW 1997, 'The effect of macroeconomic variables on stock market returns in developing markets', *Multinational Business Review*, vol. 5, no. 2, pp. 63-70.

Lee, U 1997, 'Stock market and macroeconomic policies: new evidence from Pacific Basin countries', *Multinational Finance Journal*, vol. 1, no. 4, pp. 273-289.

Makina, D 2008, 'Financial liberalization and time-varying stock price volatility: firm level evidence from South Africa', *The Business Review*, vol. 11, no. 1, pp. 309-316.

Makino, S, Isobe, T & Chan, CM 2004, 'Does country matter?', *Strategic Management Journal*, vol. 25, no. 10, pp. 1027-1043.

Tsoukalas, D 2003, ‘Macroeconomic factors and stock prices in the emerging Cypriot equity market’, *Managerial Finance*, vol. 29, no. 4, pp. 87-92.

Valadkhani, A & Chancharat, S 2008, ‘Dynamic linkages between Thai and international stock markets’, *Journal of Economic Studies*, vol. 35, no. 5, pp. 425-441.

Valadkhani, A, Chancharat, S & Havie, C 2009, ‘Analysing the impact of international stock markets and macroeconomic variables on the Thai Stock Market’, *The Business Review*, vol. 12, no. 2, pp. 50-56.

Wooldridge, JM 2002, *Econometric analysis of cross section and panel data*, MIT Press, Cambridge, Massachusetts.