REVISTA Globalización, Competitividad y Gobernabilidad | JOURNAL Globalization, Competitiveness & Governability | REVISTA Globalização, Competitividade e Governabilidade

GCG GEORGETOWN UNIVERSITY universia

An econometric Panel-MIDAS model of asset returns in the brazilian stock market

AREA: 2
TYPE: Aplication

Um modelo econométrico Painel-MIDAS dos retornos dos ativos do mercado acionário brasileiro
Un modelo econométrico Panel-MIDAS de los rendimientos de acciones del mercado bursatil brasileño

AUTHORS

Aline Moura Costa da Silva[1]
Universidade Federal Fluminense - UFF, Brasil
alinemoura@id.uff.br

Otávio Ribeiro de Medeiros
Universidade de Brasília - UnB, Brasil
otavio@unb.br

1. Corresponding author: Alameda Barros Terra, s/n – Campus UFF Valonguinho. Prédio 1. Faculdade de Administração, Ciências Contábeis. Centro, Niterói - RJ, 24020-150.

We present the specification, estimation and testing of an econometric model intended to explain and forecast individual returns of securities listed on the Brazilian stock market. The model's explanatory variables include macroeconomic, fundamental and behavioural variables sampled at different frequencies. The model uses the MIDAS regression methodology, which supports estimation of regressions with variables sampled at different frequencies. The sample includes non-financial institutions listed in the Brazilian stock exchange from 2010 to 2016. The results indicate that the model is robust in explaining and forecasting quarterly returns of individual shares listed on that market.

Apresentamos a especificação, estimação e análise de um modelo econométrico para explicar e projetar os retornos das ações do mercado acionário brasileiro. As variáveis explicativas do modelo incluem variáveis macroeconômicas, fundamentalistas e comportamentais amostradas em diferentes frequências. O modelo utiliza a metodologia de regressão MIDAS, que permite a estimação de regressões com variáveis mensuradas em diferentes frequências. A amostra contempla as ações das instituições não financeiras do mercado acionário brasileiro entre 2010 e 2016. Os resultados indicam que o modelo é robusto em explicar e projetar os retornos individuais das ações listadas naquele mercado.

Presentamos la especificación, la estimación y los análisis de un modelo econométrico para explicar y pronosticar los rendimientos de acciones del mercado bursatil brasileño. Las variables explicativas del modelo incluyen variables macroeconómicas, fundamentales y comportamentales muestreadas con diferentes frecuencias. El modelo utiliza la metodología de regresión MIDAS, que permite la estimación de regresiones con variables medidas en diferentes frecuencias. La muestra usada incluye acciones de instituciones no financieras del mercado accionario brasileño entre 2010 y 2016. Los resultados indican que el modelo es robusto explicando y pronosticando los rendimientos individuales de las acciones del mercado.

DOI	Received	Accepted
10.3232/GCG.2019.V13.N2.05	12.11.2018	01.03.2019

1. Introduction

Modern Financial Theory had its birth in the 1950s with Markowitz's (1952) portfolio theory, which provided mathematical and statistical support to the desire of financial theorists and market practitioners to develop tools to estimate and forecast financial asset prices (or returns). The following decade saw the first stock market pricing model come to life with CAPM. Within the next decades, attempts to improve CAPM and/or to find its anomalies gave rise to more comprehensive asset pricing models, some of them routinely used today by academics and market professionals.

However, other methods and techniques exist to estimate and forecast asset prices or to assess under or overpriced securities to guide investments. Traditionally, four methods are used to assess stock market prices and subsidize buying and selling orders by investors.

One of these is fundamental analysis, which consists in analysing a coherent set of financial ratios and accounting numbers to assess the financial soundness of a firm, leading to the choice of investment options in the market.

A second one is technical analysis which subsidises investors in choosing assets based on graphs of stock price movements through time.

A third one is the use of asset pricing models such as CAPM (1964), APT (1976), Fama and French's three (1993) and five (2015) factor models, Hou, Xue, e Zhang´s (2015) four factor model, among others.

A fourth possibility is the use of econometric models, such as multiple regression models, simultaneous equations models, VAR/VEC models, univariate or multivariate time series models, ARCH/GARCH models, and so on.

In the present paper, we make use of the fourth option to develop an econometric model aiming to explain individual stock returns of firms listed on the Brazilian stock market.

The model includes macroeconomic, fundamental and behavioural variables collected from data bases in different frequencies. The method developed for dealing with the estimation of regressions with variables in different frequencies is the Mixed Data Sampling (MIDAS), proposed by Ghysels, Santa-Clara, and Valkanov (2004, 2005). As such, MIDAS avoids the informational loss resulting from the conversion of higher into lower frequencies (Ghysels, Santa-Clara, and Valkanov 2004; Ghysels, Sinko, and Valkanov 2007; Andreou, Ghysels, and Kourtellos 2010,2013; Chambers 2016).

The paper innovates in several aspects: (i) it develops an econometric model upon which restrictions are not imposed a priori as asset pricing models do; (ii) macroeconomic variables are included in the model to capture the effect of the country's economic conditions on the market; (iii) microeconomic variables such as financial ratios and other fundamental variable are also included to capture the impact of idiosyncratic aspects; (iv) behavioural variables are also inserted model to capture the effect of social and psychological aspects on Brazilian the stock market; (v) since the variables are sampled in different frequencies (daily, monthly or quarterly),

KEY WORDS
Stock market; Econometric model; MIDAS; Brazil

PALAVRAS-CHAVE
Mercado acionário; Modelo econométrico; MIDAS; Brasil

PALABRAS CLAVE
Mercado bursatil; Modelo econométrico; MIDAS; Brasil.

CODIGOS JEL:
C51; C53; G10; G12

the model is specified using MIDAS, which finds the best estimation solution among a set of distributed lag functions.

2. Methods

2.1. Data and Sample

The sample consists of non-financial institutions with shares listed on the Brazilian stock market from 2010 to 2016 totalling of 192 companies. The sample period begins in 2010 because this is the first year the international accounting norms (IFRS) were fully and compulsorily adopted in Brazil[2]. We could have chosen an earlier beginning, but then possible changes due to the adoption of the new norms could occur in the middle of the sample period which could impact significantly some accounting variables such as accruals, entailing structural breaks which might introduce biases in the estimation of the model parameters. Besides, in August 2008, the world financial crisis broke out, affecting dramatically the balance sheets of most companies in Brazil as well as in most countries. This fact would also produce anomalous changes in accounting variables which would certainly generate estimation problems. The sample period ends in 2016 since this is the last fiscal year with all the required balance-sheet and economic data fully available by the time this research was concluded.

The data were collected from the following websites: Banco Central do Brasil (BCB), Institute of Applied Economic Research (Ipeadata), Getulio Vargas Foundation (FGV), Economic Policy Uncertainty (EPU), Transparency International (TI), Centre for Custody and Financial Settlement of Securities (Cetip), and Reuters' database.

2.2. The Econometric MIDAS Model

Table 1 presents the variables included, their definition, their parameters' expected signs, and their frequency. In the MIDAS model, the dependent variable must be sampled in a frequency equal to or lower to the highest frequency among the regressors. Therefore, since the regressors are sampled in daily, monthly and quarterly frequencies, the regressand, i.e. stock returns are sampled in the quarterly frequency.

Exhibit 1 - Variables included, definition, expected sign and frequency

Variable	*Definition*	*Expected Sign*	*Frequency*
RET	Stock return of individual stocks (Dependent variable)		Quarterly
ROA	Return on Assets = Operating Income divided by total assets	+	Quarterly
ROE	Return on Equity = Net income divided by Net Worth	+	Quarterly

NET	Net income minus operating cash flow divided by total assets of previous year	-	Quarterly
EBIT	Earnings before income tax divided by total assets	+	Quarterly
EBITN	Earnings before income tax divided by net revenues	+	Quarterly
NR	Net earnings divided by net revenues	+	Quarterly
NA	Net revenues divided by total assets	+	Quarterly
REA	ROE divided by ROA.	+/-	Quarterly
CFA	Operating cash flow divided by total assets	+	Quarterly
SIZE	Natural log of total assets	-	Quarterly
MB	Market-to-book ratio	-	Quarterly
MR	Return on the stock Market index	+	Daily
Δ%CDI	Percent change in the interbank deposit certificate, a proxy for the Brazilian economy interest base rate	-	Daily
Δ%XR	Percent change in the real BRL/USD exchange rate	+/-	Daily
Δ%USGDP	US real GDP growth rate	+	Quarterly
Δ%CPI	Percent change in the Brazilian Amplified Consumer Price Index (IPCA)	-	Monthly
Δ%EPU	Percent change in the Brazilian Economic Policy Uncertainty Index[3]	-	Monthly
COR	Corruption Perceptions Index for Brazil [4]	+	Quarterly
Δ%EMBI	Percent change in the Emerging Markets Bond Index for Brazil	-	Daily
VOL	Volatility of the Bovespa index	+	Daily
RSP500	Return on the S&P500 index	+	Daily
INEC	National Consumer Expectations Index is a *proxy* of the Brazilian stock market sentiment[5].	-	Monthly

2.3. Estimation of the MIDAS econometric model

The regression was estimated by panel-MIDAS, i.e. with a panel data configuration with variables in different frequencies estimated by the MIDAS procedure. For comparison, we estimated a conventional panel regression model, with the same variables of the MIDAS model. In this conventional model all variables have the same frequency and so the included variables are converted into the quarterly frequency, which is the lowest frequency among the variables.

The estimation of the conventional model was also performed with panel data. Such model can be estimated by three alternative methods: pooled regression, fixed effects, or random effects. The Breusch-Pagan, Hausman and Chow tests were used to help choosing the best model.

For both models (MIDAS and conventional), the following tests were performed to validate their results: unit root tests, residual serial correlation, heteroscedasticity, normality, multicollinearity and endogeneity.

2.4. MIDAS Regression

According to Ghysels, Santa-Clara, and Valkanov (2004), a simple MIDAS regression model, can be defined according to:

$$y_t=\beta_0+\beta_1 B(L^{1/m};\theta) X_t^{(m)} + \varepsilon_t \qquad (1)$$

where $B(L^{1/m};\theta) = \sum_{k=0}^{k\,max} B(k;\theta) L^{k/m}$ is a polynomial with extension k^{max} on the lag operator $L^{1/m}$ which impact the estimation of y_t: $L^{k/m} X_t^{(m)}=X_{t-k/m}^{(m)}$.

The term $B(k;\theta)$ is a weighting function, where the lag coefficients in $B(k; \theta)$ are parameterized by means of a small-parameter vector function θ; L is the lag operator; θ is the parameter vector of this function; m is the number of times the high-frequency variable repeats in each period t; $X_t^{(m)}$ is the independent variable with a frequency higher than that of the dependent variable, and $\beta 0$, $\beta 1$, θ are parameters estimated by nonlinear least squares.

The parameterization of the lagged coefficients of $B(k)$ is made in a parsimonious way, by means of some information criteria: Akaike, Schwarz, or Hannan-Quinn. There are the following parameterizations options: (1) Almon lag function; (2) Almon exponential lag function; (3) Beta polynomial function; and (4) step function (Ghysels, Sinko, and Valkanov 2007).

According to Kuzin, Marcellino, and Schumacher (2011), the MIDAS approach is a prediction tool, since it relates the dependent variable to the current and lagged independent variables, producing different prediction models for each horizon. The prediction model, considering a forecast horizon of h_q quarters with $h_q = h_m/3$, is determined according to:

$$y_{t_q} + h_q = y_{t_m} + h_m = \beta_0 + \beta_1 B(L_m;\theta) X^{(3)}_{t_{m+w}} + \varepsilon_{t_m} + h_m \qquad (2)$$

where $w = T^x_m - T^y_m$ and $B(L_m;\theta)$ is a polynomial in *lags*, as explained in **Equation 1**. The dependent variable $y_{t_q} + h_q$ is directly related to the indicator $X^{(3)}_{t_m+w}$ and its *lags*.

3. Previous Studies

The MIDAS modelling is used both in financial applications and macroeconomic time series forecasting. This occurs since variables sampled at a higher frequency contain potentially valuable information supposedly with greater predictive power when compared to lower frequency variables (Gao and Yang, 2017).

Li et al (2015) propose an integrated framework, which constructs a keywords base and extracts search data accordingly, and then incorporates the data into a mixed data sampling (MIDAS) model. Five groups of search data are extracted based on the keywords which are then used in the MIDAS model to forecast the Chinese consumer price index (CPI) from 2004 to 2012. The results show that the search data are strongly correlated with CPI. The MIDAS model outperforms the benchmark models, with an average reduction of the root mean square error (RMSE) of 33%.

Andreou (2016) aims at three objectives. First, she relates the standard OLS regression model with high frequency volatility predictors, with the corresponding Mixed Data Sampling Nonlinear LS (MIDAS-NLS) regression model and evaluates the properties of the regression estimators of these models. She also considers alternative high frequency volatility measures as well as various continuous time models using their corresponding relevant higher-order moments to further analyse the properties of these estimators. Second, she derives the relative MSE efficiency of the slope estimator in the standard LS and MIDAS regressions, and provides conditions for relative efficiency and present the numerical results for different continuous time models. Third, she extends the analysis of the bias of the slope estimator in standard LS regressions with alternative realized measures of risk such as Realized Covariance, Realized Beta and Realized Skewness when the true DGP is a MIDAS model. Overall, the MIDAS-NLS slope estimator turns out to be relatively more efficient than the standard LS estimator, under the various settings studied in the paper.

Zhao et al. (2018), discusses a hybrid of the mixed data sampling (MIDAS) regression model and BP (back propagation) neural network (MIDAS-BP model) to forecast carbon dioxide emissions. Such analysis uses mixed frequency data to study the effects of quarterly economic growth on annual carbon dioxide emissions. The forecasting ability of MIDAS-BP is remarkably better than MIDAS, ordinary least square (OLS), polynomial distributed lags (PDL), autoregressive distributed lags (ADL), and auto-regressive moving average (ARMA) models.

Pan et al. (2018) employ a mixed-frequency data sampling (MIDAS) approach to model the predictive relationship between monthly oil price and quarterly GDP. They used the vintage data of real GDP and consumer price index (CPI) from the Philadelphia Federal Reserve Bank real-time database. The quarterly GDP data consists of vintages for 1990Q3 through 2015Q3, each covering data extending back to 1974Q1. The monthly oil price and CPI data also cover the same sample period. They conclude that MIDAS can outperform a series of competing models including the OLS regression with quarterly oil price.

In an analysis of the Brazilian stock market, using intraday data for the most actively traded stocks of BOVESPA, Wink Jr. and Pereira (2011) considered two models available in the literature of estimation and forecasting realized volatility: the Heterogeneous Autoregressive Model of Realized Volatility (HAR-RV), developed by Corsi (2009), and the Mixed Data Sampling (MIDAS-RV), developed by Ghysels et al. (2004). Through statistical comparison of in-sample and out-of-sample forecasts, they found that superior results of the MIDAS-RV model occurred only for the in-sample forecasting. However, for out-of-sample forecasts, no statistically different results were found between the models.

4. Analysis of Results

Unit root tests (ADF-Fisher and PP-Fisher) for panel data were performed in all series. They were then differenced when necessary to make them *I(0)*. Following the estimation of a amplified model, where all variables in **Table 1** were inserted, sequential adjustments were made to obtain a parsimonious formulation. The ultimate econometric model became:

$$RET_{it} = \alpha_1 + \alpha_2 ROE_{it} + \alpha_3 REA_{it} + \alpha_4 CFA_{it} + \alpha_5 SIZE_{it} + \alpha_6 MB_{it} + \alpha_7 B(L^{1/90};\theta)\ MR_t^m + \alpha_8 B(L^{1/60};\theta)\ \Delta\%CDI_t^m + \alpha_9 B(L^{1/60};\theta)\ \Delta\%XR_t^m + \alpha_{10} \Delta\%USGDP_t + \alpha_{11} B(L^{1/3};\theta)\ \Delta\%CPI_t^m + \alpha_{12} B(L^{1/3};\theta)\ \Delta\%EPU_t^m + \alpha_{13} B(L^{1/90};\theta)\ \Delta\%EMBI_t^m + \alpha_{14} B(L^{1/60};\theta)\ RSP500_t^m + \alpha_{15} B(L^{1/3};\theta)\ INEC_t^m + \varepsilon_{it} \quad (3)$$

where $L^{1/90}$ and $L^{1/60}$ refer to a daily frequency, considering seven- and five-week days respectively, while $L^{1/3}$ refer to monthly data.

As mentioned above, for the sake of comparison, a conventional panel data regression model was also estimated. Based on the Breusch-Pagan (test-statistic=6.6593, p-value=0.0099), Hausman (test-statistic=52.5080, p-value=0.0000), Chow (test-statistic F=1.5741; p-value=0.0000, and Chi-square statistic=285.3857, p-value=0.0000), the best estimation method was the fixed effects panel. Therefore, the results of the conventional model estimated by fixed effects are compared to the results of the MIDAS model to find out which is best.

The diagnostic tests performed for the two models were: Breusch-Godfrey for residual autocorrelation, Breusch-Pagan-Godfrey for residual heteroscedasticity, Jarque-Bera for residual nonnormality, VIF test for multicollinearity; and Durbin-Wu-Hausman for endogeneity. The results of such tests for both models, point out that: (1) there is no evidence of residual autocorrelation; (2) there is no evidence of residual heteroscedasticity; (3) there is evidence of nonnormality; (4) there is no evidence of multicollinearity among regressors; and (5) there is no evidence of endogeneity.

Although the results of the Jarque-Bera test indicate that the residuals are non-Gaussian, it is valid to say that in large samples, based on the Central Limit Theorem, that the coefficients are asymptotically normal (Baltagi 2005).

The results of the estimation of the MIDAS and the conventional models (panel, fixed effects) are presented in **Table 1**.

Table 1 – Estimation results of the MIDAS and the conventional regression models

Dependent Variable: RET							
MIDAS regression–Almon				*Conventional regression (Fixed effects)*			
Variable	*Coef.*	*t-stat.*	*P-value*	*Variable*	*Coef.*	*t-stat.*	*P-value*
C	-0.0004	-0.1176	***0.9064***	*C*	-0.0130	-3.5867	0.0003***
$ROE_{(t)}$	0.0945	22.1308	0.0000***	$ROE_{(t)}$	0.0658	14.4763	0.0000***
$REA_{(t)}(-1)$	-0.0002	-8.0157	0.0000***	$REA_{(t)}(-3)$	-1.2E-05	-7.4674	0.0000***
$CFA_{(t)}$	0.6658	16.9077	0.0000***	$CFA_{(t)}$	0.6430	11.1180	0.0000***
$SIZE_{(t)}$	3.9717	8.0018	0.0000***	$SIZE_{(t)}$	3.1215	5.313123	0.0000***

$MB_{(t)}$	0.5031	57.9316	0.0000***	$MB_{(t)}$	0.3680	44.3076	0.0000***
$MR_{(d)}(-57)$	0.5579	4.6183	0.0000***	$MR(t)$	0.3395	9.8138	0.0000***
$\Delta\%VCDI_{(d)}(-21)$	-0.0466	-3.4544	0.0006***	$\Delta\%CDI_{(t)}(-3)$	-1.7865	-3.7898	0.0002***
$\Delta\%XR_{(d)}(-41)$	-0.3078	-1.9929	0.0463**	$\Delta\%XR_{(t)}(-1)$	-0.0676	-1.2718	0.2035
$\Delta\%USGDP_{(t)}$	-0.3550	-8.4618	0.0000***	$\Delta\%USGDP_{(t)}$	-0.1531	-3.2540	0.0011***
$\Delta\%CPI_{(m)}$	-6.0115	-8.1431	0.0000***	$\Delta\%CPI_{(t)}$	-4.2734	-4.9135	0.0000***
$\Delta\%EPU_{(m)}(-2)$	-0.019	-4.2312	0.0000***	$\Delta\%EPU_{(t)}$	-0.0134	-3.4942	0.0005***
$\Delta\%EMBI_{(d)}$	-0.2380	-2.4897	0.0128**	$\Delta\%EMBI_{(t)}(-1)$	-0.0377	-2.0769	0.0379**
$RSP500_{(d)}(-31)$	-0.3351	-1.9749	0.0483**	$RSP500_{(t)}(-2)$	0.1054	2.3165	0.0206**
$INEC_{(m)}$	-0.0033	-2.8422	0.0045***	$INEC_{(t)}(-2)$	0.0007	0.5417	0.5881
R²			0.6187	*R²*			0.4511
Adjusted R²			0.5762	*Adjusted R²*			0.4253
Regression standard error			0.1546	*Regression standard error*			0.1688
RSS			107.2539	*RSS*			117.0052
Log likelihood			2019.355	*Log likelihood*			1647.49
				F-stat			17.4856
				Prob. (F)			0.0000
Mean dep.var.			-0.0183	*Mean dep.var.*			-0.0175
S.E.dep.var.			0.2374	*S.E.dep.var.*			0.2227
Akaike I.C.			-0.89142	*Akaike I.C.*			-0.6760
Schwarz I.C.			-0.87003	*Schwarz I.C.*			-0.3888
Hannan-Quinn I.C.			-0.88388	*Hannan-Quinn I.C.*			-0.5746
				Durbin-Watson Stat.			1.9888
Sample (adjusted)			2010Q2 2016Q3	*Sample (adjusted)*			2010Q4 2016Q4
Included obs. after adjustments			4.497	*Included obs. after adjustments*			4.300

(d) data sampled in a daily frequency; (m) data sampled in a monthly frequency; (t) data sampled in a quarterly frequency.
***. **. *; significant at 1%. 5% and 10%, respectively.

Part A of **Table 1** shows the estimation results of the MIDAS model. These results indicate that all variables remaining in this final version are statistically significant at 1%, except Δ%XR, Δ%EMBI and RSP500 which were significant at the 5% level. Additionally, R^2 indicates that 61.87% of the changes in stock returns are explained by the model.

The set of fundamental variables that help explain the stock returns of listed companies in the Brazilian stock market were ROE, REA, CFA, SIZE, and MB. ROE, which represents the company's profitability, presented a positive relation with RET, i.e. the higher a company's profitability, the greater the return on its shares.

REA, which characterizes the level of corporate indebtedness, has shown a negative relation with stock returns. According to the Pecking Order Theory, companies follow a hierarchy when choosing financial resources, being: 1) internally generated resources, 2) issuance of debt, and 3) issuance of new shares. This is because debt indicates to the market that the company has a good reputation or favourable payment terms. However, there is a limit to indebtedness, which explains the relationship found here, according to the Static Trade-off Theory (Myers, 1984).

CFA, which represents the company's operating performance, has shown a positive relationship with the dependent variable, which shows that the higher the company's operating performance, the higher its return on equity.

The results for the variables ROE, REA and CFA, presented in **Table 1**, are in line with expectations and are supported by several studies (Chen and Zhang 2007).

On the other hand, the relationship between the SIZE variable and the stock returns does not reflect the expected relationship in the finance literature. Actually, several studies suggest the existence of the size effect (Banz 1981, Keim 1983, Fama and French 1992,1993, Jegadeesh and Titman 1993,2001, Rouwenhorst 1998). However, there are local studies that do not confirm the size effect in the Brazilian stock market (Machado and Medeiros 2011; Mussa, Fama, and Santos 2012; Martins, Paulo, and Albuquerque 2013). Therefore, the results presented here corroborate such studies, given the positive relationship found between size and return.

MB has shown a positive relationship with stock returns, which implies a negative relationship between its inverse, BM (book-to-market) and stock returns. Again, although the relationship between MB and returns should be negative according to literature, our results are consistent with the local empirical literature on the Brazilian market (Machado and Medeiros 2011; Martins, Paulo, and Albuquerque 2013).

The variables designed to capture the effect of the domestic economy on the Brazilian stock market, i.e. MR, Δ%CDI, Δ%XR, Δ%USGDP, Δ%CPI, Δ%EPU, Δ%EMBI, RSP500 and INEC have significant coefficients. MR, as expected, has a positive relationship with returns. Referring to the interest rate, the result has shown a negative relationship stock returns, as expected.

Δ%XR has shown a negative relationship with returns, which was also found by Hadhri and Ftitib (2017), for Chile and Tunisia. Tsai (2012) explains this result by stating that in a scenario where investors are more optimistic about a country's market, foreign investments in that market may increase due to speculative demand, indirectly causing the currency's appreciation.

Δ%USGDP and RSP500 have shown negative relation with stock returns. Contrary to the expected positive relationship, a plausible explanation for this could be the recession in the US, which extended beyond the 2008 crisis. In addition, a recession in the US economy or a devaluation in its currency could lead investors to flee to foreign markets such as Brazil.

The Brazilian economic instability, represented by Δ%CPI, the political uncertainty Δ%EPU and the country risk indicator Δ%EMBI, has shown a negative relationship with stock returns. It seems logical that the greater a country's economic, political and financial uncertainty, the lower the investor confidence in its market. Therefore, the trend is a decrease in investments and a fall in stock returns. These results match with the expectations and the findings of Hadhri and Ftitib (2017), Antonakakis, Gupta, and Tiwari (2017), and Christou et al. (2017).

Finally, the variable synthesizing investor sentiment, INEC, has shown a negative relation with stock returns, matching several studies (Brown and Cliff 2005; Yoshinaga, Castro Jr, and Liston 2016).

Part B of Table 1 shows the results of the estimation of the conventional model, estimated by a LSDV (Least Squares Dummy Variable) fixed effects method. As our purpose was limited to comparing the estimations of each model, in order to identify the most robust in explaining the stock returns in the Brazilian stock market, the necessary statistics for such comparison are the information criteria of Akaike, Schwarz and Hannan-Quinn, the adjusted R^2, the residual sum of squares (RSS), and the log-likelihood statistic. The results, also shown in Table 1, indicate that the Akaike, Schwarz and Hannan-Quinn informational criteria obtained from the estimation of the MIDAS regression (-0.8914; -0.8700; and -0.8839. respectively) are lower (-0.6760. -0.3888. and -0.5746. respectively) than those obtained from the conventional estimation.

The best model is the one with smaller values for such informational criteria (Akaike 1974, 1976; Schwarz 1978; Hannan-Quinn 1979). In addition, the MIDAS model presents an adjusted R^2 of 57.62% while the conventional model an adjusted R^2 of 42.53%. Besides, the RSS of the MIDAS model is smaller than that of the conventional model (107.2539 and 117.0052, respectively); and the log-likelihood statistics of the MIDAS regression is greater than that of the conventional regression (2019.3550 and 1647. 4900, respectively).

Thus, based on these results, it is possible to state that the model estimated by MIDAS presents superior performance when compared to the conventional model, i.e. the MIDAS model we propose provides a better fit to the data and can be considered more robust than the conventional one.

After the comparative analysis, forecasts of stock returns were performed. The out-of-sample forecasts were performed for the MIDAS and the conventional models, besides forecasts based on the historical average. This is because when the predictors are weak, their inclusion in the forecasting equation give low precision forecasts, which are overtaken by the historical average. Hence, when analysing the quality of forecasts from a model, several studies use the historical average as a benchmark (Campbell and Thompson 2008; Welch and Goyal 2008; Rapach, Strauss, and Zhou 2010).

The period from 2010 to 2015 was defined as the estimation window, and 2016 was chosen as the forecast window, to allow forecasts one step ahead of quarterly returns using a rolling window. To ascertain the accuracy of the models' forecasts, the mean squared error (MSE) and the mean absolute error (MAE) were computed, which are presented in Table 2.

Table 2 - Precision of stock return forecasts

Forecasted Variable	*Forecast period*	*Test*	*Model*		
			MIDAS	*Conventional (fixed effects)*	*Historical Mean*
RET	*h*=2016Q1	MSE	0.0431	0.0445	0.0824
		MAE	0.1396	0.1448	0.2101
	h=2016Q2	MSE	0.1084	0.1088	0.1466
		MAE	0.2719	0.2778	0.4003
	h=2016Q3	MSE	0.2067	0.2003	0.2294
		MAE	0.4302	0.4330	0.6190

Based on the results of **Table 2**, we can to see that the MIDAS model presents better forecasts than those obtained by the other methods, since the MSE and MAE accuracy tests, in general, revealed lower values for the forecasts of stock returns from the MIDAS regression. Lower MSE and MAE values indicate smaller forecast errors, although the differences between the results of the accuracy tests of the MIDAS and the conventional models are not very significant.

Additionally, to check whether the MIDAS model has a better predictive capacity that the conventional model, the Diebold and Mariano (1995) test was applied. The test was also applied do check whether the MIDAS model has a best predictive capacity than the historical mean. These results are in **Table 3**.

Table 3 - Diebold and Mariano (1995) test results

Loss Function	*t-statistics*	*P-value*
MIDAS-Conventional	-0.5436	0.5870
MIDAS-Historical mean	-6.2791	0.0000***

***. **. *; significant at 1%. 5% e 10%. respectively.

The results of the Diebold and Mariano (1995) test indicate that the MIDAS and the conventional models have the same predictive capacity, since the test's null could not be rejected (p-value = 0.5870). Thus, although the MIDAS model presents smaller prediction errors, as seen in **Table 2**, one cannot say it has a higher forecasting ability than the conventional one. However, the MIDAS model has a superior predictive capacity than the historical average, since the null of the Diebold and Mariano (1995) test is rejected at the 1% confidence level (p-value=0.0000). It should be mentioned that Wink Jr. and Pereira (2011) when making out-of-sample forecasts for five assets traded in the Brazilian stock market, found out that the MIDAS and the Heterogeneous Autoregressive Regression (HAR) models exhibited equivalent precision or forecasting capability.

Next, to assess the forecasts of the MIDAS model, two portfolios were built based on forecasts: (1) Portfolio 1, with companies with the highest forecast returns in each period, belonging to the first tercile; and (2) Portfolio 2, composed by companies with the lowest forecast returns in each period, belonging to the third tercile. In each forecast period, there are a total of 54 companies. The forecast average returns of these portfolios were compared with their respective average actual returns, to verify if the forecast returns, on average, are higher or lower than the actually observed returns. The results are in **Table 4**, where GSI is the Generalized Sharpe Index.

Table 4 – Portfolios built with forecast returns

Portfolio	*Tercile*	*Forecast mean Return*	*GSI (Forecast data)*	*Actual Mean Return*	*GSI*
1	1st(largest)	0.2574	0.8715	0.2769	1.3087
2	3rd(smallest)	-0.0924	-0.6749	-0.1019	-0.5856

Analysing **Table 4**, we see that the MIDAS forecasts differentiate companies with larger from those with smaller returns. The average return of forecasts of Portfolio 1 firms is 0.2574, with an average of 0.2769, while the average return of forecasts of Portfolio 2 firms and their average real return are, respectively, -0.0924 and -0.1019. According to the GSI results, the performance of the portfolio with companies with the highest forecast returns is 0.8715, while that of companies with the smallest forecast returns is -0.6749. Based on actual values, the performance of Portfolio 1 is 1.3087 and that of Portfolio 2, -0.5856.

As reported, the average return of the forecasts of companies included in the 1st tercile (0.2574) is lower than the average real return of these same companies (0.2769). The GSI based on forecasts (0.8715) was also lower than the GSI based on actual values (1.3067), that is, the average performance forecast by the MIDAS model was lower than the actual performance of the portfolio built. Since Kahneman and Tversky (1979) found that losses tend to have higher weights than gains, it is preferable that the actual performance of the constructed portfolio is higher than the forecast performance than the opposite.

4. Conclusions

An original econometric model was developed for the Brazilian stock market using the MIDAS methodology. The results show that the developed model, supported on variables representing the characteristics of the companies and on variables that characterize the economic environment in Brazil is statistically robust in explaining the returns of the shares listed on that market.

In addition, to obtain a robustness test of the MIDAS model, a conventional model was also estimated with the same variables. The estimation results indicate that the MIDAS model is statistically more robust than the conventional one.

In addition to explaining stock returns with the purpose of supporting investment decisions, we analyse the forecasts of these returns. In financial theory, forecasting stock returns is a fundamental issue, because it challenges the Efficient Market Hypothesis and because investment strategies and portfolio diversification are some of the main challenges faced by investors (Hadhri and Ftitib, 2017). In this paper, stock returns were forecast by the MIDAS model and by the conventional model, as well as by the historical average, for comparison purposes.

The results indicate that the MIDAS model achieves better forecasts, since its forecast errors are smaller than those of the other forecasts, although it is not possible to say that the MIDAS model is statistically more accurate than the conventional model, since they present the same predictive capacity. However, the forecasts elaborated by the MIDAS model are statistically more robust than those based on the historical average.

In addition, to analyse the forecasts of the MIDAS model, asset portfolios were built. The results of this analysis suggest that the forecasts performed by the MIDAS model is robust, being able to segregate companies with higher returns from those with lower returns.

According to the results, it is possible to say that the MIDAS model developed here is robust in explaining and forecasting stock returns of companies listed on the Brazilian stock market. Therefore, we believe the model can be used to set up asset portfolios for investment strategies, considering idiosyncratic aspects of the companies analysed and the economic environment in which companies are located.

We believe that the relevance of the present study and its contribution to the finance literature, especially as an application to an emerging market, is significant, since the model developed seems to be appropriate for the analysis and investment decision making in the politically and economically unstable Brazilian market.

References

Akaike, H. (1974), "A new look at the statistical model identification". IEEE Transactions on Automatic Control. 19(6):716-723. doi:10.1109/TAC.1974.1100705.

__________. (1976), "Canonical Correlation Analysis of Time Series and the Use of an Information Criterion". Mathematics in Science and Engineering. 126:27-96. doi:10.1016/S0076-5392(08)60869-3.

Andreou, E.(2016). "On the use of high frequency measures of volatility in MIDAS regressions". Journal of Econometrics, 193(2): 367–389.doi:10.1016/j.jeconom.2016.04.012.

Andreou, E.; Ghysels, E; Kourtellos, A. (2010), "Regression models with mixed sampling frequencies". Journal of Econometrics. 158(2):246-261. doi:10.1016/j.jeconom.2010.01.004.

________________________________. (2013), "Should Macroeconomic Forecasters Use Daily Financial Data and How?" Journal of Business & Economic Statistics. 31(2):240-251. doi:10.1080/07350015.2013.767199.

Antonakakis, N.; Gupta, R.; Tiwari, A.K. (2017), Has the correlation of inflation and stock prices changed in the United States over the last two centuries? Research in International Business and Finance. 42:1-8. doi:10.1016/j.ribaf.2017.04.005

Baltagi, B.H. (2005), "Econometric Analysis of panel data". 3. ed. Chichester: John Wiley&Sons.

Banz, R.W. (1981), "The relationship between return and market value of common stocks". Journal of Financial Economics. 9(1):3-18. doi:10.1016/0304-405X(81)90018-0.

Brown, G.W.; Cliff, M.T. (2005), "Investor sentiment and asset valuation". The Journal of Business. 78(2):405-440. doi:10.1086/427633.

Campbell, J.Y.; Thompson, S.B. (2008), "Predicting excess stock returns out of sample: can anything beat the historical average?" Review of Financial Studies. 21(4):1509–1531. doi:10.1093/rfs/hhm055
Chambers, M.J. (2016), "The estimation of continuous time models with mixed frequency data". Journal of Econometrics. 193(2):390-404. doi:10.1016/j.jeconom.2016.04.013.

Chen, P.; Zhang, G. (2007), "How do accounting variables explain stock price movements? Theory and evidence". Journal of Accounting and Economics. 43(2-3):219–244. doi:10.1016/j.jacceco.2007.01.001.

Christou, C.; Cunado, J.; Gupta, R.; Hassapis, C. (2017), "Economic policy uncertainty and stock market returns in Pacific Rim countries: Evidence based on a Bayesian panel VAR model". Journal of Multinational Financial Management. 40:92-102. doi:10.1016/j.mulfin.2017.03.001.

Diebold, F.X.; Mariano, R.S. (1995), "Comparing Predictive Accuracy". Journal of Business and Economic Statistics. 13:253-265. doi:10.1198/073500102753410444.

Fama, E.F.; French, K.R. (1992), "The cross-section of expected stock returns". Journal of Finance. 47(2):427-465. doi:10.2307/2329112.

___________________. (1993), "Common risk factors in the returns on stocks and bonds". Journal of Financial Economics. 33(1):3-56. doi:10.1016/0304-405X(93)90023-5.

Ghysels, E.; Santa-Clara, P; Valkanov, R. (2004), "The MIDAS touch: Mixed data sampling regression models". Cirano.Série scientifique.

_________________________________. (2005), "There is a risk-return tradeoff after all". Journal of Financial Economics. 76(3):509–548. doi:10.1016/j.jfineco.2004.03.008.

Ghysels, E.; Sinko, A.; Valkanov, R. (2007), "Midas regressions: Further results and new directions". Econometric Reviews. 26(1):53–90. doi:10.1080/07474930600972467.

Hadhri, S; Ftiti, Z. (2017), "Stock return predictability in emerging markets: Does the choice of predictors and models matter across countries?" Research in International Business and Finance.42:39–60. doi:10.1016/j.ribaf.2017.04.057.

Hou, K.; Xue, C.; Zhang, L. (2015), "Digesting anomalies: An investment approach". Review of Financial Studies. 28(3):650-705. doi:10.1093/rfs/hhu068.

Hannan, E.J.; Quinn, B.G. (1979), "The Determination of the Order of an Autoregression". Journal of the Royal Statistical Society. 41(2):190-195. Doi:177.142.61.153.

Jegadeesh, N.; Titman, S. (1993), "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency". The Journal of Finance. 48(1):65-91. doi:10.1111/j.1540-6261.1993.tb04702.x.

__________________. (2001), "Profitability of Momentum Strategies: An Evaluation of Alternative Explanations". The Journal of Finance. 56(2):699-720. doi:10.1111/0022-1082.00342.

Kahneman, D.; Tversky, A. (1979), "Prospect Theory: An analysis of decision under risk". Econometrica. 47(2):263-291. doi:10.2307/1914185.

Keim, D.B. (1983), "Size-related anomalies and stock return seasonality: Further empirical evidence". Journal of Financial Economics. 12(1):13-32. doi:10.1016/0304-405X(83)90025-9.

Kuzin, V.; Marcellino, M; Schumacher, C. (2011). "MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the euro area". International Journal of Forecasting. 27(2):529-542. doi:10.1016/j.ijforecast.2010.02.006.

Li, X.; Shang, W.; Wang, S.; Ma, J. (2015). "A MIDAS modelling framework for Chinese inflation index forecast incorporating Google search data". Electronic Commerce Research and Applications, 14(2):112–125.doi: 10.1016/j.elerap.2015.01.001.

Liston, D.P. (2016), "Sin stock returns and investor sentiment". The Quarterly Review of Economics and Finance. 59:63-70. doi:10.1016/j.qref.2015.08.004.

Machado, M.A.V.; Medeiros, O.R. (2011), "Modelos de precificação de ativos e o efeito liquidez: evidências empíricas no mercado acionário brasileiro". Revista Brasileira de Finanças. 9(3):383-412.

Marcellino, M.; Schumacher, C.C. (2010), "Factor-MIDAS for now-and forecasting with ragged-edge data: A model comparison for German GDP". Oxford Bulletin of Economics and Statistics. 72(4):518-550. doi:10.1111/j.1468-0084.2010.00591.x.

Martins, O.S.; Paulo, E.; Albuquerque, P.H.M. (2013), "Negociação com informação privilegiada e retorno das ações na BM&FBovespa". RAE. 53(4):350-362.

Markowitz, H. (1952), "Portfolio Selection". The Journal of Finance. 7(1):77-91.

Mussa, A.; Famá, R., Santos, J.O. (2012). "A adição do fator de risco momento ao modelo de precificação de ativos dos três fatores de Fama & French aplicado ao mercado acionário brasileiro". REGE. 19(3):431-447. doi:10.5700/issn.2177-8736.rege.2012.49925.

Myers, S.C. (1984), "The capital structure puzzle". Journal of Finance. 39(3):575-592. doi:10.1111/j.1540-6261.1984.tb03646.x

Pan, Z.; Wang, Q.; Wang, Y.; Yang, L.(2018). "Forecasting U.S. real GDP using oil prices: A time-varying parameter MIDAS model". Energy Economics. 72:177–187.doi: /10.1016/j.eneco.2018.04.008.

Rapach, D.E.; Strauss, J; Zhou, G.Z. (2010), "Out-of-sample equity premium prediction: Combination forecasts and links to the real economy". Review of Financial Studies. 23(2): 821–862. doi:0.1093/rfs/hhp063.

Rouwenhorst, G.K. (1998), "International Momentum Strategies". The Journal of Finance. 53(1):267-284. 1998. doi:10.1111/0022-1082.95722.

Schwarz, G. (1978), "Estimating the Dimension of a Model". Annals of Statistics. 6(2):461-464.

Welch, I.; Goyal, A. (2008). "A comprehensive look at the empirical performance of equity premium prediction". Review of Financial Studies. 21:1455–1508. doi:10.1093/rfs/hhm014.

Wink Jr., M.V.; Pereira, P.L.V. (2011) "Modeling and Forecasting of Realized Volatility: Evidence from Brazil". Brazilian Review of Econometrics, 31(2):315–337.doi: 10.12660/bre.v31n22011.4056.

Yoshinaga, C.E.; Castro Jr. F.H.F. (2012), "The Relationship between Market Sentiment Index and Stock Rates of Return: a Panel Data Analysis". BAR. 9(2):189-210.

Zhao, X.; Han, M.; Ding, L.; Calin, A.C.(2018). "Forecasting carbon dioxide emissions based on a hybrid of mixed data sampling regression model and back propagation neural network in the USA". Environmental Science and Pollution Research. 25(3):2899–2910.doi: 10.1007/s11356-017-0642-6.