A Work Project, presented as part of the requirements for the Award of a Master Degree in Economics from the NOVA – School of Business and Economics and Insper Institute of Education and Research.

DO INDUSTRIES LEAD STOCK MARKETS IN BRAZIL?

RAPHAEL BENO CITRON

A Project carried out on the Master in Economics Program, under the supervision of: Martijn Boons and Ricardo D. O. Brito.

January 6, 2017

Do Industries Lead Stock Markets In Brazil?

Abstract

Hong et al. (2007) claim that a number of industry returns in U.S. and in eight largest non-U.S. stock markets can forecast market returns using monthly data. Tse (2015) reexamines their results in U.S. with updated data and extended period. He finds evidence that the market can predict industries more significantly than can the reverse. I investigate these relationships for the Brazilian market, adding a study to check if the causality from industries to the market is independent of the chosen model. Data are from August of 1994 to September of 2016. I found that is not possible to conclude that industry returns cause stock market returns for Brazil. I also show that the forecasting power that some models may present is not robust to a sub-sample analysis. The results are compatible with the efficient market hypothesis.

Keywords: Asset Pricing, Information and Market Efficiency, Financial Markets and Macroeconomy, International Financial Markets

JEL: G12, G14, E44

Executive Summary

In the Stock Exchange there are listed companies with diverse economic activity. It's possible to classify them accordingly to similarities. There are companies that sell consumer goods, others that specialize in mining, banks and etcetera. The similarities that these companies have between them, make it possible that some factors affect some sectors in a way and have a different impact on others. The oil price affects the oil companies more than the companies that sell insurance, the mining companies are affected by the GDP slowdown of emerging economies, like China, more than the companies that sell telecommunications services. In this thesis I analyze the relations of these sectors with the stock market return and economic activity for Brazil. I have based myself in the study of Hong, Tourous and Valkanov (2007), they did a similar analysis to the U.S..

In the economic cycle, companies are somehow related. When a TV is sold, the company that produced the TV have to buy packing material before it sells the TV. A signal that the economy is booming will appear first in the results of companies that sell paper to pack, as an example. The same could be true for other activities, when a car is sold, the car manufacturer have to buy steel first. It's possible to check if there are leading or lagging sectors. Sectors where the price of the companies that compound it can move before or after the aggregate stock return, as whole, move. If this relationship exists, it's due to the relationship of the sectors with economic activity, this will be evident analyzing data related to economic activity, like the Industrial Production Growth.

The efficient market hypothesis states that prices fully and quickly reflect all available information. In other words, when the steel company that are selling steel to a car manufacturer have its profits increasing, and consequently the stock price, the price of the car manufacturer share will move accordingly, at the same time. The leading behavior of a sector contradicts this hypothesis and there are gradual diffusion of information. The gradual diffusion of information can exist because of the limited information-processing capacity for asset prices. Investors have a limited ability of paying attention to the many sources of information, much less rationalize their impact on the prices of the assets that they trade.

To formally claim that a sector have a leading behavior, this sector must cause the stock market return and the stock market return can't cause the sector return. In other words, the sector return today will impact next month stock market return and the stock return today can't impact the sector return in the next month. To test this I have to control other variables that may impact the analysis, like the inflation, which I did.

Knowing if data is leading may help to criticize the efficient market hypothesis, it can also help portfolio managers and traders to improve their performance.

I found that, for Brazil, from August of 1994 to September of 2016, there is no strong evidence to claim that industries cause the stock market. I tested this with multiples models and model configurations. The leading behavior of the vast majority of the sectors does not hold when changing the period of analysis. There are a small amount of evidence of the leading behavior of the Utilities industry and the Electricity sector. I found evidence that lagged economic data explains industries returns more than the reverse, proportionally. Economic data has a clear correlation with market return. In this case the economic information is absorbed by the market participants and, consequently, the market return, during the same month. The overall results are consistent with the efficient market hypothesis.

1. Introduction

The efficient market hypothesis states that prices fully and quickly reflect all available information. This study is based on theories that explore the implications of limited information-processing capacity for asset prices. Investors are better characterized as being boundedly rational, Shiller (2000) and Sims (2001), quoted by Hong et al. (2007). In other words, investors have a limited ability of paying attention to the many sources of information, much less rationalize their impact on the prices of the assets that they trade.

I repliacate part of the paper “Do Industries Lead Stock Markets? Hong, Tourous and Valkanov (2007)”, for the Brazilian market (from now on HTV). In the paper, the authors test the hypothesis that industry portfolios are able to predict the movement of stock markets and economic indicators. For the United States, HTV find that some industry portfolios forecast the stock market by up to two months. The predictive power is due to the correlation of the industry portfolios with economic activity indicators such as industrial production growth. The purpose of this thesis is test the ability of Brazilian industries portfolios to lead the stock market and some economic indicators. This present study will also include a causality investigation as suggested by Tse (2015) in “Do industries lead stock markets? A reexamination” (from now on TSE).

The predictability of industry portfolios can be useful in many ways. Traders and portfolios managers would have the ability to improve their portfolios with a higher Sharpe ratio, i.e. hedge funds that have in their mandate market timing would benefit from knowing which industries have forecasting power and the exact time of the impact. It can also help economist investigate economic activity, since industry portfolios can be used as instrumental variable.

I found that, for Brazil, is not possible to conclude that industry portfolios returns Granger causes stock market returns. Like TSE and different than HTV I analyze the causality in the opposite direction, from the market to industries. The predictability that some estimated regressions may present is not robust to a sub-sample analysis. I also did not find any evidence that the predictability of industry portfolios is correlated with economic activity indicators.

In Section 2, I review some of the literature, specially HTV and the work of TSE, who criticized the way that HTV tested causality. In Section 3, I present the model and methodology, first I present the theoretical model that can make predictability in industries portfolios to exist, and then I present the empirical models that can be estimated to test the propositions. In Section 4, I explain the data that were used and its sources. In Section 5, I present the results of the many models, the causality between market returns and industry portfolios and a correlation study between economic indicators and the coefficients of lagged industries returns. In Section 6, I check if the results of the previous section are robust to sub-sample analysis. In Section 7, I conclude and compare the results with HTV and TSE.

2. Literature Review

One possible explanation for stock return predictability is given by Hong and Stein (1999), quoted by HTV. “They develop a dynamic model where there is only one asset, the information gradually diffuses across investors, and market participants are unable to perform the rational expectations mechanism of extracting information from prices. The consequence is an under-reation of the price to information, making the existence of stock return predictability possible". Merton (1987), quoted by HTV, "created a static model of numerous stocks in which investors only have information about a restricted number of stocks and only deal with those they have information. As a result, companies that are less recognized (small research coverage and trading volume) by investors have a diminuted investor's base and trade at a greater discount because of limited risk-sharing."

The hypothesis that guided HTV is that the gradual diffusion of information across asset markets leads to cross-asset return predictability. "Some investors, such as those that specialize in trading the market index, such as the S&P Future, receive delayed information from individual industries. This hypothesis relies on two key assumptions. The first is that important information that is generated in one asset market reaches traders in other markets only with a lag. The second assumption is that because of limited information-processing capacity, many investors may not pay attention or be able to extract the information from asset prices of markets that they do not have deep knowledge."

HTV find that in the U.S., "apparel, print, petroleum, leather, real estate, mines, metal, transportation, utilities, retail, money or financial and services have t-statistics of the corresponding lagged industry return coefficient that are greater than 1.96 in absolute value (significant at 5% level). Two additional industries, non-metallic minerals and television, have a t-statistics of about 1.7 (significant at 10% level). There are a total of fourteen industries that can significantly predict the market." The regressions make economic sense, i.e. the lagged returns of petroleum and metal industry portfolios are negatively related to the next period's market return. This happens because a commodity price shock led to an economic downturn.

HTV find that in regressions with longer horizons, with market return lagged in more than one month, only four industries are significant at 10% level and none at 5% level. This is consistent with information taking about two months to be completely incorporated from industries into the broad market index.

Previous studies, quoted by HTV, like Fama and French (1989), Fama and Schwert (1977) and Campbell and Shiller (1988), have demonstrated the predictive nature of inflation, the default spread, the market dividend yield and the term spread. These variables are related to time varying risk. HTV suggest that lagged industry portfolios are not a proxy to that and capture different factors.

HTV find that twelve of thirty-four industries have predictive power over Industrial Production. HTV explains that this result is in line with Lamont (2001), "who finds that portfolios formed from industry returns can lead various economic variables like industrial production growth, inflation and consumption growth. Industries that forecast the market also forecast industrial production growth." Similar results were obtained using Stock and Watson Coincident Index. Eight industries are able to forecast market fundamentals at 10% significance level and five at 5% significance level.

When extending the empirical analysis to the rest of the world, HTV find that across eight countries the U.K. and Japan have the smallest proportion of significant industries (8 out of 34). Australia and the Netherlands have the most (18 out of 31 and 16 out of 29, respectively). The numbers for IPG (Industrial Production Growth) are comparable, which lead to a similar conclusion of the U.S. case.

A recent study by Tse (2015) argues that "using an extend period, 1946-2013, and data, 48 industries, only between one to seven industries have significant predictive ability for the stock market, depending on the significance level and the model specifications. The author argues in favor of the opposite predictive direction, from the stock market to industries." In this way, results are consistent with the efficient market hypothesis.

Different than HTV, TSE uses Granger causality to investigate the relation between the market and industries with the Newey-west correction for heteroscedasticity and autocorrelation.

3. Model and Methodology

Theoretically, HTV proposes a model consisting of the pricing of two assets (stocks) in a three-date economy, t = 0, 1, 2, they assume that the risk free rate is zero. The two assets, X and Y, have terminal values at t=2 given by D_x and D_y, which are jointly normal with means of zero and variances ${\sigma^2}_{X,D}$ and ${\sigma^2}_{Y,D}$ and covariance $\sigma_{XY,D}$.

Investors participate in either market X or market Y. This limited participation is due to taxes and regulations.

"At t=1, investors in market X receive signal $S_X = D_x + \varepsilon_{X,S}$ about the terminal value of X, investors in market Y receive signal $S_Y = D_Y + \varepsilon_{Y,S}$ about the terminal value of Y, and these signals are known to all participants at t=2. This is the gradual-information-diffusion assumption. The noise in the signals is normally distributed and the errors are independent of each other and of all other shocks in the economy. The supply of assets is assumed to be Q_x and Q_y shares outstanding for assets X and Y, respectively." HTV.

HTV assumes that investors have CARA preferences with a risk aversion coefficient of a. Given the price function $P_{k,t}$, the investor in asset market k ($k = X, Y$) solves the following optimization problem:

Max $E_{k,0}$[-exp(-a$W_{k,2}$)],   k=X,Y,
$\{\theta_k\}$
s.t $W_{k,t} = W_{k,t-1} + \theta_{k,t-1}(P_{k,t} - P_{k,t-1})$,

where $W_{k,t}$ and $\theta_{k,t}$ are the wealth and the share holdings of a representative investor in asset market k at time t and $P_{k,2} = D_k$.

The equilibrium price in market k is given by:

$$P_{k,t} = E_{k,t}[D_k] - b_{k,t}Q_k, \quad k = X, Y,$$

Where $E_{k,t}[D_k]$ is the conditional expectation of the terminal payoff of asset k at time t, $b_{k,t} > 0$ is the standard risk discount at time t, and Q_k is the supply of the asset.

Returns are given by:

$$R_{k,t} = \frac{P_{k,t} - P_{k,t-1}}{P_{k,t-1}}$$

Given the equilibrium prices from the model above, two propositions arose.

"Proposition 1: The own serial return correlations are zero, i.e., Corr($R_{k,2}, R_{k,1}$) = 0 for K = X, Y. The cross-serial return correlations, Corr($R_{Y,2}, R_{X,1}$) and Corr($R_{X,2}, R_{Y,1}$) are non-zero and can be positive or negative depending on the sign of the covariance of asset payoffs, $\sigma_{XY,D}$." HTV.

"Proposition 2: Even if there are arbitrageurs who trade in both markets to exploit the cross-predictability, as long as there are limits to arbitrage, some cross-predictability will remain in equilibrium." HTV.

Empirically, HTV test two specific predictions that are implied by the model. Thinking on the broad market index as asset Y and an industry portfolio that is informative of market fundamentals as asset X.

"Prediction 1: The broad market index can be predicted by the returns of industry portfolios, controlling for lagged market returns and well-known predictors such as inflation, the default spread, and the dividend yield." HTV.

"Prediction 2: The ability of an industry to forecast the market is related to its ability to forecast changes in market fundamentals such as industrial production growth or changes in other indicators of economic activity." HTV.

Predictive Regressions Involving Industry and Market Returns for the U.S. Stock Market (Prediction 1):

$$RM_t = \alpha_i + \lambda_i R_{i,t-1} + A_i Z_{t-1} + e_{i,t} \qquad (1)$$

where RM_t is the excess return of the market in month *t*, $R_{i,t-1}$ is the excess return of industry portfolio *i* lagged in one month and Z_{t-1} is a vector of additional market predictors such as lagged excess market, inflation, the default spread and the market dividend yield.

Industry Returns and Market Fundamentals (Prediction 2):

$$X_t = \eta_i + \gamma_i R_{i,t-1} + C_i ZX_{t-1} + \nu_{i,t} \qquad (2)$$

Where X_t is the month *t* realization of the indicator of economic activity, $R_{i,t-1}$ is the previous month's return of industry *i* and ZX_{t-1} is the same as Z_{t-1} on the previous equation, except that three monthly lags of economic activity are include.

The peculiarities of HTV, related to the extension of the analysis to the rest of the world, is that in (1), a lagged monthly market return is included as a control, due to lack of data other market predictors were excluded, the market return is raw, there is no interest rate discount on it. In (2) a lag of the monthly market return and three monthly lags of Industrial Production Growth are include. In this thesis, the results use a similar U.S. specification that is in HTV.

TSE, to check the causal relation form the market to industries, in which HTV is silent, estimate equation (3):

$$IND_t = \alpha_i + \beta_i RM_{i,t-1} + A_i Z_{t-1} + e_{i,t} \qquad (3)$$

Doing so will lead us to infer if industry Granger cause the stock market. In this study I will compare how many $\beta_{i=1\ to\ n}$ in equation (3) are relevant comparable to the $\lambda_{i=1\ to\ n}$ in equation (1).

4. Data

The period of analysis range from August of 1994 until September of 2016, using monthly data. For industries portfolios, as in HTV, I use the data available in Eikon (Former Datastream). The Structure on these indexes uses four different levels, first there is the Industry, followed by the Supersector then the Sector and finally the Subsector. The Industry level is the more general one and the Subsector one is the most specific. The full description of the levels are available in the industry classification benchmark website[1]. In HTV there is no explanation on the methodology of choice of the industry portfolios, there are Supersectors, Sectors and Subsectors. In Table 1, summary statistics of the Brazilian industries portfolios, they are classified by the four industries levels. There were a total of 73 industry portfolios in 4 different levels for the period studied. I use the return of this series minus the interest rate (described below).

The interest rate is the Brazil Cetip DI Interbank Deposit Rate, since data are in yearly terms, a transformation to make it monthly is done. There are daily data for the Brazilian interest rate, but the series only starts in 1999. For the market return I use TOTMKBR, a value weighted index constituted of one hundred Brazilian companies that is the same that HTV used for market returns in the eight countries analyzed, apart from the U.S.. The market return (RM), is the TOTMKBR minus the Interest rate.

For the inflation rate (INF) the IPCA (Broad National Consumer Price) in a monthly basis is used, this index is calculated by IBGE and it covers eleven urban areas, these consumer price data is collected from the first to the last day of the reference month. The data were extracted from the Bloomberg platform.

The Default Spread (DSPR) in HTV is the Spread between BAA rated and AAA rated corporate bonds, there is no proxy for Brazil. The EMBI calculated by JP Morgan is used, this index is the difference between a portfolio sovereigns Brazilian bonds and the U.S. treasury, in this way a country effect is included. A high correlation between the corporate debt spread index and the EMBI is expected, but is clear that they measure totally different things.

The Market Dividend Yield (MDY) is the dividend yield of the market index TOTMKBR available in Eikon.

The Market volatility (MVOL) is the standard deviation of the last thirty days of TOTMKBR. For each month a new standard deviation is calculated with the last thirty days values.

The Industrial Production Growth (IPG) is the Brazilian Industrial production percentage volume variation from month to month. It includes mining, extraction, industrials and utilities production volume. This data are calculated by IBGE and were extracted from Bloomberg.

As the Default Spread (DSPR) is different in essence from HTV, I present, in Table 2, a covariance analysis with the purpose of investigate multicolinearity. The forecasting variables seem to be linearly independent.

[1] www.icbenchmark.com/ICBDocs/Structure_Defs_English.pdf

Table 1: Summary statistics for Brazilian Industries portfolios, data are from August of 1994 to September of 2016.

Summary statistics for industries portfolios, values are presented in percentage. The returns are in excess of the risk free rate (Brazil Cetip DI). For the mean, there are a persistency of negative values, the standard deviation is high.

INDUSTRIES	Mean	Median	Max.	Min.	Std. Dev.	Obs.	Description
BMATRBR	-0.50	-0.99	26.92	-33.48	8.94	266	Basic Mats
CNSMGBR	0.42	0.22	49.46	-33.39	8.58	266	Consumer Gds
CNSMSBR	0.45	1.11	45.71	-31.91	8.92	175	Consumer Svs
FINANBR	-0.14	0.07	23.49	-29.18	7.66	266	Financials
HLTHCBR	0.08	-0.61	31.81	-29.46	8.97	106	Health Care
INDUSBR	0.21	-0.14	34.57	-25.87	8.08	266	Industrials
OILGSBR	0.32	-0.12	46.31	-29.13	11.34	266	Oil & Gas
TECNOBR	0.62	-0.67	41.06	-25.93	8.68	125	Technology
TELCMBR	-0.58	-1.66	40.61	-44.67	8.95	266	Telecom
UTILSBR	-0.61	-0.04	27.94	-47.52	8.63	266	Utilities

SUPERSECTORS	Mean	Median	Max.	Min.	Std. Dev.	Obs.	Description
BRESRBR	-0.44	-0.86	32.46	-37.19	9.50	266	Basic Resource
FDBEVBR	0.66	0.55	54.99	-39.99	9.51	266	Food & Bev
FINSVBR	0.40	0.28	38.13	-24.05	9.56	192	Financial Svs(3)
INDGSBR	2.26	-0.58	338.32	-84.07	26.63	266	Inds Gds & Svs
INSURBR	0.41	0.34	21.31	-25.48	8.44	141	Insurance
PERHHBR	-0.05	0.21	48.96	-26.16	8.45	266	Pers & H/H Gds
RLESTBR	0.87	-1.00	55.62	-42.02	13.94	134	Real Estate
RTAILBR	1.12	-0.20	152.84	-42.64	17.25	266	Retail

SECTORS	Mean	Median	Max.	Min.	Std. Dev.	Obs.	Description
AUTMBBR	1.23	-1.41	596.40	-52.11	37.42	266	Auto & Parts
BANKSBR	-0.09	0.08	26.60	-31.70	7.96	266	Banks
CHMCLBR	-0.17	-1.39	56.20	-39.10	10.67	266	Chemicals
CNSTMBR	-0.26	-0.50	70.59	-42.94	12.39	107	Con & Mat
AERSPBR	2.17	-1.60	338.32	-90.64	28.74	266	Aero/Defence
BEVESBR	0.67	0.56	25.03	-37.90	8.13	266	Beverages
ELECTBR	-0.66	-0.46	27.94	-47.51	8.82	266	Electricity
ELTNCBR	1.46	0.55	67.09	-23.99	10.17	224	Eltro/Elec Eq
FDRGRBR	0.42	0.27	33.45	-58.15	8.47	249	Fd & Drug Rtl
FNSVSBR	0.22	0.26	38.13	-24.05	9.32	192	Financial Svs(4)
FOODSBR	1.30	0.44	94.52	-68.71	15.34	266	Fd Producers
FSTPABR	-0.13	-0.26	111.37	-42.13	12.91	266	Forestry & Pap
GNINDBR	1.06	0.45	33.55	-24.00	9.56	176	General Inds
GNRETBR	0.35	-0.34	107.76	-87.87	15.12	266	Gen Retailers
GWMUTBR	-0.24	0.21	42.92	-50.48	10.69	231	Gs/Wt/Mul Util
HCEQSBR	0.08	-0.61	31.81	-29.46	8.97	106	H/C Eq & Svs
INDMTBR	-0.29	-0.44	40.99	-39.31	11.75	266	Ind. Met & Mines
INDTRBR	0.66	-0.26	45.73	-31.90	9.89	175	Inds Transpt
NLINSBR	0.52	0.56	25.13	-25.49	8.52	141	Nonlife Insur
OILGPBR	0.47	0.85	49.57	-28.29	12.12	266	Oil & Gas Prod
PERSGBR	0.25	0.22	21.71	-20.53	7.99	147	Personal Goods
RLISVBR	0.96	-0.78	55.62	-42.02	13.96	134	Real Est Inv,Svs
SFTCSBR	0.62	-0.67	41.05	-25.93	8.68	125	S/W & Comp Svs
TELFLBR	-0.76	-1.84	41.85	-48.19	9.08	266	Fxd Line T/Cm
TELMBBR	-0.03	-0.94	64.98	-41.41	13.57	266	Mobile T/Cm

Table 1 (Continued): Summary statistics for Brazilian Industries portfolios, data are from August of 1994 to September of 2016.
At the end, values for other variables and the number of industries portfolios.

SUBSECTORS

	Mean	Median	Max.	Min.	Std. Dev.	Obs.	Description
AEROSBR	2.17	-1.60	338.32	-90.64	28.74	266	Aerospace
ALTELBR	-0.34	-0.35	34.34	-47.52	10.11	266	Alt. Electricity
APRETBR	2.35	-1.28	245.53	-63.02	23.52	211	Apparel Rtl
AUPRTBR	1.24	-1.42	596.47	-52.12	37.43	266	Auto Parts
BDRETBR	0.29	-0.78	107.97	-87.87	15.29	266	Broadline Rtl
BMATSBR	0.05	-1.33	70.61	-42.92	13.82	107	Build Mat/Fixt
BREWSBR	0.67	0.56	25.08	-37.91	8.15	266	Brewers
CHEMSBR	0.12	-2.58	72.04	-36.27	14.35	266	Commodity Chem
CNFINBR	0.33	0.24	18.12	-17.16	7.65	109	Consumer Fin
CNVELBR	-1.04	-0.76	48.06	-40.26	9.08	266	Con. Electricity
CONPKBR	1.06	0.45	33.55	-24.00	9.56	176	Cont & Pack
DGRETBR	2.66	-1.40	422.23	-31.15	29.59	244	Drug Retailers
ELEQPBR	0.98	0.41	26.12	-23.99	7.79	224	Elec Compo/Eq
FDPRDBR	1.30	0.44	98.07	-68.71	15.50	266	Food Products
FDRETBR	-0.25	-0.01	25.49	-28.07	6.47	249	Fd Rtl & W
FLINSBR	0.52	0.57	25.13	-25.48	8.52	141	Full Lin Insur
FOOTWBR	-0.22	-2.05	30.27	-25.11	9.39	142	Footwear
GASDSBR	1.32	-1.26	119.04	-28.25	13.98	209	Gas Dst
HCPROBR	0.08	-0.61	31.82	-29.46	8.97	106	H/C Providers
INVSVBR	0.25	-0.44	43.22	-34.84	11.53	192	Investment Svs
OILINBR	0.44	0.65	53.22	-50.55	13.02	266	Int Oil & Gas
PAPERBR	0.11	-0.54	106.36	-42.13	14.33	266	Paper
PRSNLBR	0.27	0.90	27.24	-19.52	8.74	147	Personal Prd
RLDEVBR	0.97	-0.82	55.62	-42.02	14.02	134	R/E Hld & Dvlp
SOFTWBR	0.63	-0.67	41.05	-25.94	8.68	125	Software
SPCSVBR	1.71	1.73	50.21	-52.84	11.71	135	Spec Cons Svs
SPFINBR	0.30	0.46	20.37	-32.24	7.85	266	Speciality Fin
STEELBR	-0.26	-0.44	40.98	-39.31	11.86	266	Iron & Steel
TRNSVBR	0.92	0.31	45.73	-31.90	10.06	175	Transpt Svs
WATERBR	-0.40	-0.10	42.92	-50.48	11.31	231	Water

OTHER VARIABLES

	Mean	Median	Max.	Min.	Std. Dev.	Obs.	Description
RM	-0.22	0.04	21.15	-33.63	7.29	266	Market minus int.
INF	0.65	0.52	6.84	-0.51	0.64	266	Inflation
DSPR	0.02	0.00	17.61	-16.64	3.49	266	Embi
MDY	1.27	-0.28	78.39	-38.28	13.43	266	Mkt Div Yield
MVOL	1.39	1.21	5.69	0.34	0.76	266	Mkt Vol
IPG	0.11	0.21	6.76	-12.78	1.99	266	Ind Prod Growht

Number of Industries Portfolios

INDUSTRIES	10
SUPERSECTORS	8
SECTORS	25
SUBSECTORS	30
TOTAL	**73**

Table 2: Covariance and correlation analysis, data are from August of 1994 to September of 2016.

Variables are not correlated. The higher correlation is between RM and MDY, which is -0.46.

Covariance Correlation	**RM**	**MVOL**	**MDY**	**INF**	**DSPR**
RM	0.01 1				
MVOL	-0.01 -0.18	0 1			
MDY	-0.01 -0.46	0.01 0.15	0.02 1		
INF	0 0.14	0 0.08	0 -0.11	0 1	
DSPR	0 -0.03	0 0.06	0 0.02	0 -0.01	0.01 1

5. Results

In Table 3 are the results of various regressions that establish the predictive ability of the Oil & Gas Producers industry portfolio, market return is the endogenous variable. I choose this industry because is a relevant industry in the Brazilian economy. In the first column (A), a forecasting regression of market returns on a constant and the lagged values of the OILGP portfolio. In column B, a forecasting regression of market returns on a constant and lagged market returns. In the third column (C), a forecasting regression of market returns on a constant, the lagged values of the OILGP portfolio, and RM. The coefficient on lagged OILGP is -0.1692 and is statistically significant. As HTV, I can see that this coefficient is still statistically significant even after I control for other predictors such as INF, DSPR, and MDY in the regression D. In regression E, the lagged market volatility is included and the coefficient is still significant. The standard errors include a Newey-West serial correlation and heteroskedasticity correction with twelve monthly lags. Note that the signal on OILGP coefficient is negative, this is in line with the economic rationale since an increase in oil price, a commodity shock, historically lead to downturns.

Table 3: Predictive regressions between Oil & Gas Producers and market portfolios (Brazil), data are from August of 1994 to September of 2016. Market Return is the Endogenous variable. Coefficients are in percentage (In regression A, lagged OILGP coefficient is -2.21%)
Regression E is the same as Predictive Regression (1).

	A	B	C	D	E
C	-0.29	-0.28	-0.15	-0.21	-0.18
p-value	*0.51*	*0.53*	*0.73*	*0.75*	*0.84*
OILGPBR(-1)	**-2.21**		**-16.93**	**-17.01**	**-17.01**
p-value	***0.55***		***0.01****	***0.02***	***0.02 ***
RM(-1)		9.19	31.20	26.00	25.97
p-value		*0.13*	*0.01 ****	*0.09 **	*0.09 **
INF(-1)				18.78	19.08
p-value				*0.74*	*0.72*
DSPR(-1)				5.25	5.28
p-value				*0.66*	*0.70*
MDY(-1)				-6.11	-6.10
p-value				*0.15*	*0.13*
MVOL(-1)					-2.47
p-value					0.95
R^2	0.00	0.01	0.04	0.05	0.05

Note: (*) significant at 10%, (**) significant at 5%, (***) significant at 1%.

In Table 4, I present forecasts of the market return using various industries and forecast of various industries using market returns. In the left side of the table, the number of significant industries, λ from predictive regression 1, that lead stock market. In the right side of the table the number of industries that are led by the market, β from predictive regression 3, with this, I include the TSE critics of testing both directions. The time lag of variables is specified by *i*. The purpose of Table 4 is not to find the optimal lag or a better model, but to find the same patterns in the nine models estimated. Table 4 is constructed to better illustrate these patterns.

HTV and TSE use different arguments to make different conclusions. HTV try to check if industries lead the stock market by more than one month. “Literature on stock market predictability says that being able to predict next month return is already an outstanding achievement.” HTV. Valkanov (2003) and Tourous, Valkanov and Yan (2004), quoted by HTV, argue that “previous findings on long-horizon predictability are an artifact of not properly adjusting standard errors for the near random walk behavior of various predictors.” In Table 4 I increase the lag length of exogenous variables and we can see that the number of variables with predictability power increases. I suspect that this is due to the standard errors being wrong, but as HTV did not focus on that and my purpose is not to find the optimum lag length, I will look on one month predictability to infer Granger causality. Besides that, I have tried multiple models with different lag lengths, in the left side of Table 4 I see how many industries lead the stock market, I’ve studied four different industries aggregations levels, displayed are the number of industries with relevant coefficients with 1%, 5% and 10% significance level. In parenthesis is the percentage of relevant coefficients in relation to the total of each level. In the right side of Table 4 is the number of industries that are led by the stock market. There are a total of 10 Industries, 8 Supersectors, 25 Sectors and 30 Subsectors.

Analyzing the second type of regressions (M2 in Table 4), with exogenous coefficients lagged in one month (i=1) and with NW correction of 12 lags, I can see that both industries lead and are led by the stock market by the same proportion. In the left side, row Total, of the table 21.9% (10% signf.), 15.1% (5% signf.) and 4.1% (1% signf.) lead the stock market and in the right side, row Total, 20.5% (10% signf.), 17.8% (5% signf.) and 5,5% (1% signf.) are led by the stock market.

Table 4 lead me to suspect that, different than HTV and TSE, causality does not come from industries to market and does not come from market to industries. However, I can see that there is a gradual diffusion of information in almost every model and in both directions.

Table 4: Forecast of RM and Industries involving various models

This table presents forecasts of the market return using various industries and forecast of various industries using market returns. In the left side of the table the number of significant industries (λ from pred. regress. 1) that lead stock market, and in the right side of the table the number of industries that are led by the market (β form pred. regress. 3). The time lag of variables is specified by *i*. The percentages are related to the total of industries portfolios, in Model 1, at 10%, 25% of the Supersectors portfolios are significant (2/8) and 17.8% of the industries portfolios are significant (13/73). The exogenous variables for all models except Model 8 and Model 9 are MDY, MVOL, DSPR, INF and RM (the same as Table 3, Regression E). I've tried two Newey-West correction specifications, with 3 and 12 lags, as TSE did.

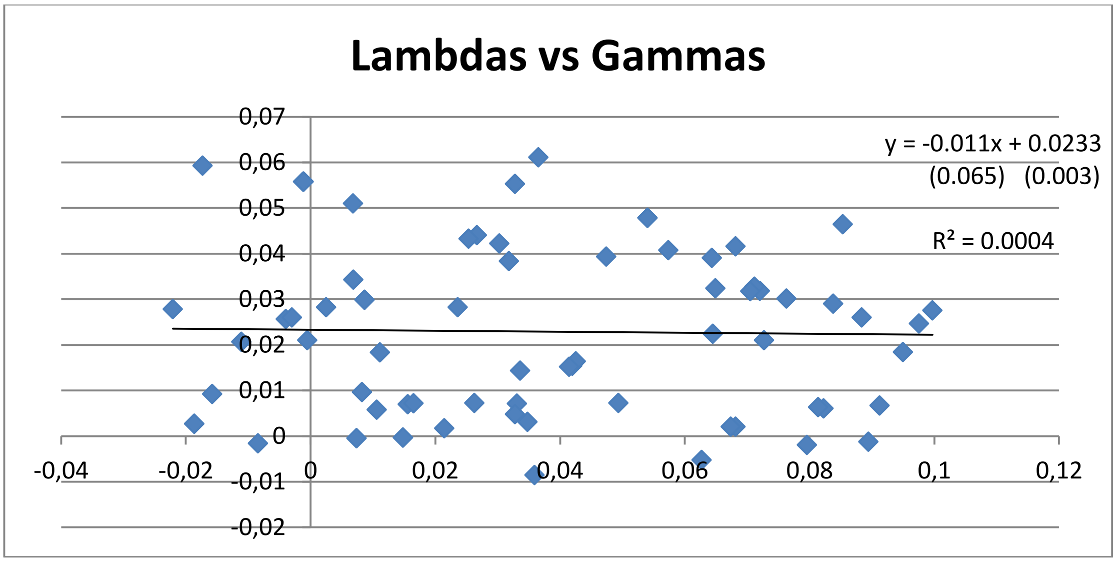
		No of Significants λ(-i)			No of Significants β(-i)		
		10%	5%	1%	10%	5%	1%
M1	Newey-West with 3 lags i=1						
	Industry	1 (10.0%)	1 (10.0%)	1 (10.0%)	1 (10.0%)	1 (10.0%)	0 (0.0%)
	Supersector	2 (25.0%)	1 (12.5%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	Sector	6 (24.0%)	4 (16.0%)	1 (4.0%)	1 (4.0%)	1 (4.0%)	1 (4.0%)
	Subsector	4 (13.3%)	2 (6.7%)	0 (0.0%)	6 (20.0%)	2 (6.7%)	1 (3.3%)
	Total	**13 (17.8%)**	**8 (11.0%)**	**2 (2.7%)**	**8 (11.0%)**	**4 (5.5%)**	**2 (2.7%)**
M2	Newey-West with 12 lags, i=1						
	Industry	2 (20.0%)	1 (10.0%)	0 (0.0%)	2 (20.0%)	1 (10.0%)	1 (10.0%)
	Supersector	2 (25.0%)	2 (25.0%)	0 (0.0%)	2 (25.0%)	1 (12.5%)	0 (0.0%)
	Sector	7 (28.0%)	5 (20.0%)	2 (8.0%)	4 (16.0%)	4 (16.0%)	1 (4.0%)
	Subsector	5 (16.7%)	3 (10.0%)	1 (3.3%)	7 (23.3%)	7 (23.3%)	2 (6.7%)
	Total	**16 (21.9%)**	**11 (15.1%)**	**3 (4.1%)**	**15 (20.5%)**	**13 (17.8%)**	**4 (5.5%)**
M3	*NW with 3 lags, i=2*						
	Industry	3 (30.0%)	2 (20.0%)	1 (10.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	Supersector	2 (25.0%)	2 (25.0%)	2 (25.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	Sector	7 (28.0%)	4 (16.0%)	1 (4.0%)	2 (8.0%)	1 (4.0%)	0 (0.0%)
	Subsector	8 (26.7%)	5 (16.7%)	1 (3.3%)	2 (6.7%)	1 (3.3%)	0 (0.0%)
	Total	**20 (27.4%)**	**13 (17.8%)**	**5 (6.8%)**	**4 (5.5%)**	**2 (2.7%)**	**0 (0.0%)**
M4	*NW with 12 lags, i=2*						
	Industry	3 (30.0%)	3 (30.0%)	2 (20.0%)	2 (20.0%)	2 (20.0%)	1 (10.0%)
	Supersector	3 (37.5%)	2 (25.0%)	2 (25.0%)	2 (25.0%)	2 (25.0%)	0 (0.0%)
	Sector	11 (44.0%)	7 (28.0%)	3 (12.0%)	5 (20.0%)	2 (8.0%)	2 (8.0%)
	Subsector	11 (36.7%)	9 (30.0%)	5 (16.7%)	5 (16.7%)	2 (6.7%)	1 (3.3%)
	Total	**28 (38.4%)**	**21 (28.8%)**	**12 (16.4%)**	**14 (19.2%)**	**8 (11.0%)**	**4 (5.5%)**

Table 4 (Continued): Forecast of RM and Industries involving various models
In Model 7 I include 3 lags of RM and in all other variable i=1. In Model 8, i=1, constant and industries are exogenous, Model 9 is equal Model 8 with RM (-1) included.

		No of Significants λ(-i)			No of Significants β(-i)		
		10%	5%	1%	10%	5%	1%
	NW with 3 lags, i=3						
	Industry	3 (30.0%)	1 (10.0%)	0 (0.0%)	1 (10.0%)	0 (0.0%)	0 (0.0%)
M5	*Supersector*	3 (37.5%)	2 (25.0%)	1 (12.5%)	1 (12.5%)	1 (12.5%)	1 (12.5%)
	Sector	8 (32.0%)	4 (16.0%)	3 (12.0%)	4 (16.0%)	2 (8.0%)	2 (8.0%)
	Subsector	8 (26.7%)	5 (16.7%)	0 (0.0%)	4 (13.3%)	2 (6.7%)	1 (3.3%)
	Total	**22 (30.1%)**	**12 (16.4%)**	**4 (5.5%)**	**10 (13.7%)**	**5 (6.8%)**	**4 (5.5%)**
	NW with 12 lags, i=3						
	Industry	6 (60.0%)	5 (50.0%)	3 (30.0%)	4 (40.0%)	2 (20.0%)	1 (10.0%)
M6	*Supersector*	5 (62.5%)	4 (50.0%)	2 (25.0%)	2 (25.0%)	2 (25.0%)	1 (12.5%)
	Sector	11 (44.0%)	10 (40.0%)	4 (16.0%)	10 (40.0%)	7 (28.0%)	4 (16.0%)
	Subsector	12 (40.0%)	8 (26.7%)	4 (13.3%)	7 (23.3%)	5 (16.7%)	4 (13.3%)
	Total	**34 (46.6%)**	**27 (37.0%)**	**13 (17.8%)**	**23 (31.5%)**	**16 (21.9%)**	**10 (13.7%)**
	3 lags of dependend Variables						
	Industry	1 (10.0%)	1 (10.0%)	0 (0.0%)	1 (10.0%)	0 (0.0%)	0 (0.0%)
M7	*Supersector*	2 (25.0%)	2 (25.0%)	0 (0.0%)	1 (12.5%)	0 (0.0%)	0 (0.0%)
	Sector	7 (28.0%)	4 (16.0%)	2 (8.0%)	2 (8.0%)	1 (4.0%)	0 (0.0%)
	Subsector	4 (13.3%)	2 (6.7%)	1 (3.3%)	6 (20.0%)	3 (10.0%)	1 (3.3%)
	Total	**14 (19.2%)**	**9 (12.3%)**	**3 (4.1%)**	**10 (13.7%)**	**4 (5.5%)**	**1 (1.4%)**
	Rm End, Const and Indust. Exg						
	Industry	2 (20.0%)	1 (10.0%)	0 (0.0%)	5 (50.0%)	2 (20.0%)	1 (10.0%)
M8	*Supersector*	2 (25.0%)	0 (0.0%)	0 (0.0%)	1 (12.5%)	1 (12.5%)	0 (0.0%)
	Sector	9 (36.0%)	3 (12.0%)	1 (4.0%)	8 (32.0%)	4 (16.0%)	2 (8.0%)
	Subsector	6 (20.0%)	2 (6.7%)	1 (3.3%)	9 (30.0%)	5 (16.7%)	2 (6.7%)
	Total	**19 (26.0%)**	**6 (8.2%)**	**2 (2.7%)**	**23 (31.5%)**	**12 (16.4%)**	**5 (6.8%)**
	Rm End, C and Indust.,rm(-1) Exg						
	Industry	1 (10.0%)	1 (10.0%)	1 (10.0%)	5 (50.0%)	4 (40.0%)	3 (30.0%)
M9	*Supersector*	1 (12.5%)	1 (12.5%)	0 (0.0%)	2 (25.0%)	2 (25.0%)	1 (12.5%)
	Sector	7 (28.0%)	5 (20.0%)	4 (16.0%)	8 (32.0%)	6 (24.0%)	2 (8.0%)
	Subsector	6 (20.0%)	4 (13.3%)	2 (6.7%)	10 (33.3%)	7 (23.3%)	4 (13.3%)
	Total	**15 (20.5%)**	**11 (15.1%)**	**7 (9.6%)**	**25 (34.2%)**	**19 (26.0%)**	**10 (13.7%)**

In HTV, it's claimed that the ability of an industry to lead the stock market is related to its propensity to lead some economic indicators. To prove this relation HTV plot the coefficients of lagged industries in the vertical axis, where RM is endogenous and only lagged industries are exogenous. In the horizontal axis the coefficients of lagged industries, where IPG is endogenous and only lagged industries is exogenous. It's expected a positive relation between the coefficients. In their examples, only Japan does not exhibit a positive relationship, this can be due to high standard deviation in the time series, as explained by HTV. For the Brazilian case I can't find any relations between the coefficients, as Figure 1 demonstrates.

Fig. 1: Lambdas versus Gammas
No Clear relationship can be seen, this lack of relation it's possibly due to high standard deviation. Lambdas are the coefficients of industries when market return is endogenous. Gammas are the coefficients on industries when IPG is endogenous.



In Table 5, it's possible to analyze the causality between IPG and industries. There is no clear relation. Note that lagged IPG is better in forecasting industries than the reverse. The number of industries that lead IPG in the left side of Table 5 is smaller than the number of lagged IPG that forecast industries in the right side of the Table 5.

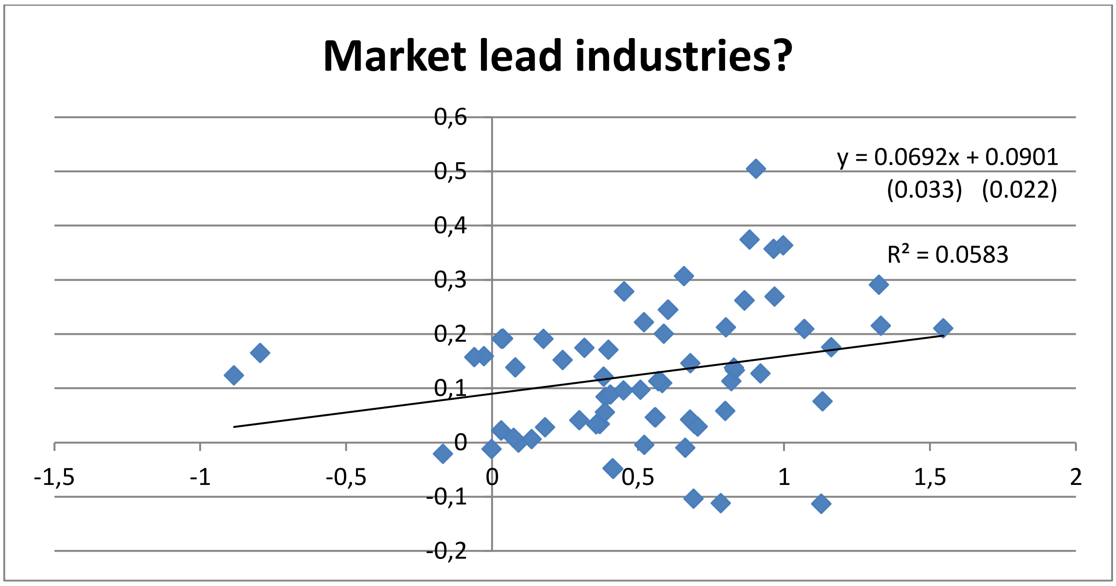
Table 5: Causality between IPG and industries
This Table presents on the left side, regressions (predictive regression 2) where Industrial Production Growth is an endogenous variable, as exogenous variables there are one month lagged industries (i=1) and a constant, I present the total of significant lagged industries coefficients (γ). On the right side on the Table, regressions where Industries is an endogenous variable and lagged one month industries (i=1) and a constant are exogenous variables. I present the total of significant lagged IPG coefficients (β2). I've used a 3 months Newey-West correction. Lagged IPG is better in forecasting industries returns than Lagged industries to forecast IPG. The same Industries portfolios are used. The percentages are related to the total of industries portfolios, 9.6% (7/73) of industries lead IPG with 1% significance level and in 26% (19/73) of the regressions, lagged IPG forecast industries portfolios.

	No of Significants γ(-i)			No of Significants β2(-i)		
	10%	5%	1%	10%	5%	1%
NW with 3 lags, IPG exogenus and end.						
Industry	4 (40.0%)	3 (30.0%)	1 (10.0%)	6 (60.0%)	5 (50.0%)	5 (50.0%)
Supersector	2 (25.0%)	1 (12.5%)	1 (12.5%)	4 (50.0%)	4 (50.0%)	1 (12.5%)
Sector	8 (32.0%)	4 (16.0%)	3 (12.0%)	13 (52.0%)	12 (48.0%)	6 (24.0%)
Subsector	15 (50.0%)	8 (26.7%)	2 (6.7%)	18 (60.0%)	14 (46.7%)	7 (23.3%)
Total	**29 (39.7%)**	**16 (21.9%)**	**7 (9.6%)**	**41 (56.2%)**	**35 (47.9%)**	**19 (26.0%)**

Comparing the ability of the Brazilian market to lead industries and the ability of IPG to lead industries, using a dispersion diagram, it can be observed a clear positive relationship in Figure 2, which, mistakenly, may lead to think that the market leads industries and there is an economic explanation. This forecasting power of the market is due to its relationship with economic activity. Correlation does not imply causality.

Fig. 2: Dispersion graph between the coefficients of lagged RM and lagged IPG with industries as endogenous variables.

Clear positive relation, but not enough to a causality conclusion.



The next step is to breakdown the causality relation by industries. The idea is to find leading and lagging industries in most of the models and industries which no conclusion is possible.

In Table 6, I investigate these relationships, in the white columns, the models that represent significant industries (λ) that lead the stock market, the linear regressions are the same as the ones described in the left part of Table 4. The idea is not to differentiate between the models but to try making conclusions that are robust to the chosen model. In Grey are the models that represent the significant industries that are led by the market (β), they are the same as the linear regressions represented in the right side of the Table 4. The significance level is 1% (***), 5% (**), 10% (*). In Table 6, it's possible to visually identify the leading and lagging industries, if there are asterisk in various white columns it's a leading industry, if there are various asterisk in the grey columns it's a lagging industry. When there are asterisks in both white and grey columns, no clear conclusion its possible .

In the analysis of Table 6 is evident that there are some industries that appear to be leading ones. This is the case of Consumer Goods, Oil & Gas, Utilities, Retail, Auto & Parts, Construction & materials, Electricity, Industrial Transport, Forestry & Paper and Drug Retailers.

By the same logic is possible to see that Electro & Electronic Equipments, Software & Computer Services, Commodity Chemicals and Speciality Finance, appear to be lagging industries.

Table 6: Causality per industries

In this table, I present the causality relations classified by industries portfolios. In white are the models where industries lead the stock market, in grey the models where the stock market lead industries. The Significance level is 1% (***), 5% (**), 10% (*). Take as example the AUTMBBR (Auto & Parts) Sector, is evident that all models lead to a conclusion that this is a leading sector, because in almost every model there is (***) on the white columns. M1 is the same as M1 in the left side of Table 4; M1B is the same as M1 in the right side of Table 4, and so on.

	M1	M1B	M2	M2B	M3	M3B	M4	M4B	M5	M5B	M6	M6B	M7	M7B	M8	M8B	M9	M9B	Description
INDUSTRIES																			
BMATRBR				*	***		***	**			**	*				**		***	Basic Mats
CNSMGBR					*		**				***								Consumer Gds
CNSMSBR									*		*								Consumer Svs
FINANBR									**	*	***	**			*				Financials
HLTHCBR					**		***		*		***					*		**	Health Care
INDUSBR		**		***							**	*		*		***		***	Industrials
OILGSBR	**		**										**			*	***	*	Oil & Gas
TECNOBR								***				***				*		***	Technology
TELCMBR			*												**				Telecom
UTILSBR					**		**		*		**	*			*				Utilities
SUPERSECTORS																			Description
BRESRBR				*	***		***	**			**					**		***	Basic Resource
FDBEVBR							*				**								Food & Bev
FINSVBR	**		*					**	*		***	**	**		*		**		Financial Svs(3)
INDGSBR																			Inds Gds & Svs
INSURBR			*	**														**	Insurance
PERHHBR									**		*								Pers & H/H Gds
RLESTBR	*								***	***	***	***	**		*				Real Estate
RTAILBR					***		***							*					Retail
SECTORS																			Description
AUTMBBR	**		**						**		**		***		***		***		Auto & Parts
BANKSBR									***		***	**							Banks
CHMCLBR																**		*	Chemicals
CNSTMBR	**		***	***			***		***		**		**				***		Con & Mat
AERSPBR																			Aero/Defence
BEVESBR				**							**			*				*	Beverages
ELECTBR					*		*		*		*				*				Electricity
ELTNCBR		***		**	*		*			*		***		**		***		***	Eltro/Elec Eq
FDRGRBR																			Fd & Drug Rtl
FNSVSBR	*		*			*		***	*		**	**	*		*		*		Financial Svs(4)
FOODSBR			*		**		*					*			*		*		Fd Producers
FSTPABR					***		***					*							Forestry & Pap
GNINDBR							*			*		**							General Inds
GNRETBR							**	*			**					*			Gen Retailers
GWMUTBR					*		**	*			***								Gs/Wt/Mul Util
HCEQSBR					**		***		*		***					*		**	H/C Eq & Svs
INDMTBR							**									***		**	Ind. Met & Mines
INDTRBR	**		***										*		**	*	***		Inds Transpt
NLINSBR				**														**	Nonlife Insur
OILGPBR	***		**										***			**	***	**	Oil & Gas Prod
PERSGBR						**		*											Personal Goods
RLISVBR									***	***	***	***	*		*				Real Est Inv,Svs
SFTCSBR								***				***				*		***	S/W & Comp Svs
TELFLBR															*				Fxd Line T/Cm
TELMBBR	*		**							***		***	**		**		**		Mobile T/Cm

Table 6 (Continued): Causality per industries

	SUBSECTORS																		
	M1	M1B	M2	M2B	M3	M3B	M4	M4B	M5	M5B	M6	M6B	M7	M7B	M8	M8B	M9	M9B	Description
AEROSBR																			Aerospace
ALTELBR					**		**	**	**		**								Alt. Electricity
APRETBR					*		**			*		***							Apparel Rtl
AUPRTBR	**		**						**		**		***		***		***		Auto Parts
BDRETBR							***				**					*			Broadline Rtl
BMATSBR				***			***		**		***	***					**		Build Mat/Fixt
BREWSBR				**							*			*				*	Brewers
CHEMSBR		***		***							*			***		***		***	Commodity Chem
CNFINBR					**		***												Consumer Fin
CNVELBR		*		**							*			*	*			***	Con. Electricity
CONPKBR							*			*		**							Cont & Pack
DGRETBR	*		**								*		*		*		**		Drug Retailers
ELEQPBR		**				*		*	*		**			**		**		*	Elec Compo/Eq
FDPRDBR			*		**		*					*			*		*		Food Products
FDRETBR																			Fd Rtl & W
FLINSBR				**														**	Full Lin Insur
FOOTWBR				**															Footwear
GASDSBR			*						**								*		Gas Dst
HCPROBR					**		***		*		***					*		**	H/C Providers
INVSVBR					*							*							Investment Svs
OILINBR	*															**		*	Int Oil & Gas
PAPERBR		*			***		***							*					Paper
PRSNLBR						**		*											Personal Prd
RLDEVBR									**	***	***	***	*		*				R/E Hld & Dvlp
SOFTWBR								***				***				*		***	Software
SPCSVBR										**									Spec Cons Svs
SPFINBR		*		**												**		***	Speciality Fin
STEELBR							**									***		**	Iron & Steel
TRNSVBR	**	*	***										**	**	**	*	***		Transpt Svs
WATERBR					*		**	*	*		***								Water

Using other econometrical techniques to investigate the forecasting power of industry portfolios I estimate, in a system, Model 7, M7 in Table 4. First I've used OLS as an estimator and then SUR (Table 7). As TSE and HTV noted[2], it's not clear how to form a system of equations with the same dependent variable.

Regarding the OLS estimated in a system, the coefficients and p-values are equivalent to the ones estimated individually. I tested the validity of industries portfolios coefficients using a Wald-test. Different than HTV, the Wald-test reject the relevance of the estimated coefficients, with 73 degrees of freedom the Chi-square value is 64.78, the p-value is 0.7428. I have tried many models, using different and less industries levels, they all have the same conclusion regarding the Wald-test.

In Table 7, a SUR is estimated, the SUR is useful when the error terms are in fact correlated between the equations, the standard errors, and consequently the p-values take in account these correlations. The results demonstrates that, all industries does not presents predictive power, the Wald-tests have big p-values. Various configurations and industry levels were estimated, all with the same conclusions. I've excluded the subsectors data because they are multicolinear with the other industry portfolios.

Table 7: Seemingly Unrelated Regressions (SUR) estimated and Wald-test
There is no significant industry portfolio, the Wald-test confirm this hypothesis. This is a modification of Predictive Regression 1.

SUR			SUR		
Ind. Port.	**Coefficient**	**p-value**	**Ind. Port.**	**Coefficient**	**p-value**
BMATRBR	-0.04	0.62	**AERSPBR**	0.00	0.81
CNSMGBR	0.05	0.65	**BEVESBR**	0.05	0.67
CNSMSBR	0.19	0.79	**ELECTBR**	0.05	0.72
FINANBR	0.13	0.75	**ELTNCBR**	0.51	0.73
HLTHCBR	-0.02	0.97	**FDRGRBR**	-0.05	0.98
INDUSBR	0.06	0.76	**FNSVSBR**	1.12	0.26
OILGSBR	-0.31	0.49	**FOODSBR**	0.05	0.65
TECNOBR	0.14	0.86	**FSTPABR**	-0.05	0.72
TELCMBR	0.06	0.73	**GNINDBR**	-0.19	0.82
UTILSBR	0.06	0.68	**GNRETBR**	0.00	0.68
BRESRBR	-0.03	0.65	**GWMUTBR**	0.04	0.98
FDBEVBR	0.04	0.62	**HCEQSBR**	-0.02	0.97
FINSVBR	1.11	0.25	**INDMTBR**	0.00	0.77
INDGSBR	0.01	0.80	**INDTRBR**	0.32	0.65
INSURBR	-0.26	0.67	**NLINSBR**	-0.25	0.68
PERHHBR	0.00	0.78	**OILGPBR**	-0.47	0.39
RLESTBR	0.35	0.57	**PERSGBR**	0.62	0.69
RTAILBR	0.12	0.44	**RLISVBR**	0.34	0.59
AUTMBBR	0.02	0.54	**SFTCSBR**	0.14	0.86
BANKSBR	0.09	0.78	**TELFLBR**	0.04	0.76
CHMCLBR	-0.09	0.61	**TELMBBR**	0.11	0.60
CNSTMBR	0.63	0.59			

Wald Test			
Test Statistic	**Value**	**df**	**Probability**
Chi-square	5.29	43	1
Null Hypothesis: All Industries are equal zero			

Note:
Example of the specification of the regressions in SUR
RM = C(1) + C(2)*DSPR(-1) + C(3)*INF(-1) + C(4)*MDY(-1) + C(5)*MVOL(-1) + C(6)*RM(-1) + C(7)*RM(-2) + C(8)*RM(-3) + C(10)*A_BMATRBR_267(-1)

[2] HTV Pg.379, TSE pg. 197.

Besides the SUR and the OLS, estimated in a system, in Table 8, I've estimated regressions using all industries simultaneously in the same regression. To do this, I've excluded some Industries portfolios, otherwise I would have multicolinearity problems. Like TSE, I experiment some model configurations, the models with the Newey-West corrections have p-values (Wald-test), on the industries portfolios coefficients, which are jointly significant. The models with no correction, with 3 lags of market returns does not presents significant industry portfolios coefficients. In all regression there are a small amount of individually significant industries, this can also be seen in Table 8.

Table 8: Forecasting the market using industries simultaneously

The regressions with the Newey-West correction have coefficients that are jointly significant. In all regression there are individually significant industries. There are two industries portfolios, with NW correction of 3 lags and using only sector level data, which are significant at a 10% level.

	No. Of Significant Industries				
	10%	5%	1%	p-value	obs
Newey-West 3 lags (except subsectors)	6	4	3	0.000	25
Newey-West 3 lags (only sectors)	2	2	1	0.003	13
3 lags of Dep Variables (except subsectors)	4	2	0	0.46	25
3 lags of Dep Variables (only sectors)	3	2	1	0.17	13

6. Robustness analysis

To check whether our results hold for two sub-samples I create diagrams that compare the whole sample results with the sub-sample ones. I arbitrary split the data in two parts. The data are from August of 1994 to September of 2005 in the first sample and from October of 2005 to September of 2016 in the last one. Since there are industries that have less observations in the first sample I've excluded the series with does not have all the months in both samples. Therefore, this analysis will include 37 industries portfolios.

In Table 9, I investigate the models and regressions that test if "industries lead the stock market" hold results for two sub-samples. The first white columns (All sample) in the left of the Table 9 are the same as the white columns in Table 6, they represent the regressions that test if industries portfolios are leading indicators for the market return. The next two class of columns (1994 – 2005 and 2005 – 2016) in grey and white are the results of the leading industries portfolios regressions for the two sub-samples. The Significance level is 1% (***), 5% (**), 10% (*). The columns in grey called "Both 3 Samples" check if the models for the both three samples exhibit significant industries portfolios. If there is a "1" it means that, in both three samples, the model have a significant industry portfolio.

Table 10 investigates the models and regressions that test if industries lag the stock market and if this results hold for two sub-samples. The logic of the table is equal to the logic in Table 9.

Table 9 and Table 10 demonstrates that the results are not, in general, robust. The columns called "Both 3 Samples" do not display a single industry portfolio that has more than one model with significant coefficients in both three samples. Of the "leading" industries portfolios in Table 6, only Utilities and Electricity have a single model that has significant coefficients in both three samples. Of the "lagging" industries portfolios in Table 6, only Commodity Chemicals have a single model that has significant coefficients in both three samples.

*Table 9: Robustness analysis for the **leading** models*

This Table investigates if the leading industries portfolios are robust to a sub-sample analysis. There is not an industry portfolio with more than a model with significant coefficients. Take as an example Auto & Parts, a leading industry portfolio if the whole sample is used, there is not a model where both three samples display significant coefficients. Only Utilities and Electricity have a model with significant coefficients in both three samples. If there is a "1" in the last nine columns "Both 3 Samples" it means that, in both three samples, the model have a significant industry portfolio.

	All Sample									1994 - 2005									2005 - 2016									Both 3 Samples									
INDUSTRIES	M1	M2	M3	M4	M5	M6	M7	M8	M9	M1	M2	M3	M4	M5	M6	M7	M8	M9	M1	M2	M3	M4	M5	M6	M7	M8	M9	M1	M2	M3	M4	M5	M6	M7	M8	M9	Description
BMATRBR			***	***		**							*	***							*	***		**							1						Basic Mats
CNSMGBR			*	**		***														*	**	***		***			*										Consumer Gds
FINANBR					**	***		*			*	*									***	***															Financials
INDUSBR						**						***										*															Industrials
OILGSBR	**	**					**		***		**				*			*																			Oil & Gas
TELCMBR		*						**								**		*																			Telecom
UTILSBR			**	**	*	**		*				*					*				**	**								1							Utilities
SUPERSECTORS																																					Description
BRESRBR			***	***		**							*	*							**	***		**							1						Basic Resource
FDBEVBR				*		**														***	**	***		***	*		***										Food & Bev
INDGSBR																	*																				Inds Gds & Svs
PERHHBR					**	*						*					**																				Pers & H/H Gds
RTAILBR			***	***						*																											Retail
SECTORS																																					Description
AUTMBBR	**	**			**	**	***	***	***	***	*	*				***		***			***	***															Auto & Parts
BANKSBR					***	***						**									**	*		**													Banks
CHMCLBR																			***	***					**		***										Chemicals
AERSPBR																				***																	Aero/Defence
BEVESBR						**								*								**															Beverages
ELECTBR			*	*	*	*		*				*					*				**									1							Electricity
FOODSBR		*	**	*				*	*										**	***	*	**			**	*	***										Fd Producers
FSTPABR			***	***															*	**	**	***			*		**										Forestry & Pap
GNRETBR				**		**						*					*																				Gen Retailers
INDMTBR				**																																	Ind. Met & Mines
OILGPBR	***	**					***		***	**	*				*			***																			Oil & Gas Prod
TELFLBR								*			*					**		**																			Fxd Line T/Cm
TELMBBR	*	**					**	**	**					***																							Mobile T/Cm
SUBSECTORS																																					Description
AEROSBR																				***																	Aerospace
ALTELBR			**	**	**	**						**									***	*								1							Alt. Electricity
AUPRTBR	**	**			**	**	***	***	***	***	*	*				***		***			***	***															Auto Parts
BDRETBR				***		**						*					*																				Broadline Rtl
BREWSBR						*								*								**															Brewers
CHEMSBR						*									***		**		**	***					**		***										Commodity Chem
CNVELBR						*		*						***	*		**			*							*										Con. Electricity
FDPRDBR		*	**	*				*	*										**	***	*	**			**	*	***										Food Products
OILINBR	*											**																									Int Oil & Gas
PAPERBR			***	***															*	**	**	***			*		**										Paper
SPFINBR																																					Speciality Fin
STEELBR				**																																	Iron & Steel

Note: The Significance level is 1% (***), 5% (**), 10% (*).

*Table 10: Robustness Analysis for the **lagging** models*

This Table investigates if the lagging industries portfolios are robust to a sub-sample analysis. There is not an industry portfolio with more than a model with significant coefficients. Take as an example Specialty Finance, a lagging industry portfolio if the whole sample is used, there is not a model where both three samples display significant coefficients. Only Commodity Chemicals have a model with significant coefficients in both three samples. If there is a "1" in the last nine columns "Both 3 Samples" it means that, in both three samples, the model have a significant industry portfolio.

	All Sample									1994 - 2005									2005 - 2016									Both 3 Samples									
INDUSTRIES																																					
	1B	2B	3B	4B	5B	6B	7B	8B	9B	1B	2B	3B	4B	5B	6B	7B	8B	9B	1B	2B	3B	4B	5B	6B	7B	8B	9B	1B	2B	3B	4B	5B	6B	7B	8B	9B	Description
BMATRBR		*		**		*		**	***							*	**		**	***		**															Basic Mats
CNSMGBR														**								**	***	***													Consumer Gds
FINANBR					*	**						*		*			*				*	*															Financials
INDUSBR	**	***				*	*	***	***	**	***	*	***	***			***	*						**		**									1		Industrials
OILGSBR								*	*		**	*	**																								Oil & Gas
TELCMBR													*	**				**									***										Telecom
UTILSBR						*								***					*	***							**										Utilities
SUPERSECTORS																																					Description
BRESRBR		*		**				**	***		*						**		**	***		**							1								Basic Resource
FDBEVBR														*								**	**	***	*												Food & Bev
INDGSBR																				*				*		**											Inds Gds & Svs
PERHHBR																								***													Pers & H/H Gds
RTAILBR							*								*		*																				Retail
SECTORS																																					Description
AUTMBBR														***	***																						Auto & Parts
BANKSBR						**							**	***																							Banks
CHMCLBR								**	*		**						*								**												Chemicals
AERSPBR																					***	***	*	*													Aero/Defence
BEVESBR		**					*		*		**			**									***	**		*											Beverages
ELECTBR														***					**	**																	Electricity
FOODSBR						*							***											*	**		***										Fd Producers
FSTPABR						*													***	**					*												Forestry & Pap
GNRETBR				*				*			*			***				***						*													Gen Retailers
INDMTBR								***	**		**					**	***																				Ind. Met & Mines
OILGPBR								**	**	*	***				**		*	**																			Oil & Gas Prod
TELFLBR																		*						*			**										Fxd Line T/Cm
TELMBBR					***	***						***	***	*			**																				Mobile T/Cm
SUBSECTORS																																					Description
AEROSBR																					***	***	*	*													Aerospace
ALTELBR				**										***					**																		Alt. Electricity
AUPRTBR														***	***																						Auto Parts
BDRETBR								*			*	*		***				***																			Broadline Rtl
BREWSBR		**					*		*		**			*									***	**		*											Brewers
CHEMSBR	***	***					***	***	***	***	***		*	*	**	*	***						**	**	**									1			Commodity Chem
CNVELBR	*	**					*		***	*	***			**				***									**									1	Con. Electricity
FDPRDBR						*							***											*	**		***										Food Products
OILINBR								**	*		*			***			*																				Int Oil & Gas
PAPERBR	*						*												***	**					*												Paper
SPFINBR	*	**						**	***		**						***	**																			Speciality Fin
STEELBR								***	**		*					**	***																				Iron & Steel

Note: The Significance level is 1% (***), 5% (**), 10% (*).

7. Conclusions

The main purpose of this thesis is to investigate the causality between the market and industries for Brazil. I conclude that is not possible to claim causality from industries to the market. It’s not clear that industries lead stock markets in Brazil, because using the same scope of tools that HTV, I have the same evidence from the market returns to industries portfolios returns, there is as much as the same amount of valid coefficients with market as exogenous variable, industries as endogenous and the opposite.

It’s possible to find some leading and lagging industries, when deepening the analysis, to check for robustness, the results are less significant. Utilities and Electricity appear to be leading industries and only Commodity Chemicals appear to be a lagging industry portfolio.

Lagged IPG explains industries returns more than the reverse, proportionally. IPG has a clear correlation with market return. This can be seen as an evidence of contemporaneous relationship between IPG and market return. In this case the economic information is absorbed by the market participants and, consequently, the market return, during the same month. However, is worth to note that if there is gradual diffusion of information from the market to industries this would be evident in this analysis, there is no strong evidence of significant lagging coefficients where, industries are endogenous and lagged market return is exogenous. This can be seen as evidence that IPG and market return are not leading indicators for industries portfolios in Brazil. Besides the evidence, a proper causality test between IPG and industries portfolios (individually) follow as a suggestion for a next study.

The SUR estimated in a system, does not present individually significant industry portfolios coefficients, a Wald-test confirm the insignificance of the coefficients jointly tested, this show that if the contemporaneous correlation between the industries portfolios is somehow taken into account, there is no gradual diffusion of information from the industries to the market. The Wald-test on the OLS estimated in a system reject the joint significance of the industries portfolios. Other possibilities are that the volatility of the Brazilian or the validity of the controls variables chosen, invalidate the models estimated in a system, although I’ve tested many model specifications. Since there is no consensus when estimating systems with the same endogenous variables, those are secondary results.

I have estimated a regression with exogenous lagged industries portfolios simultaneously. I found some models with jointly significant coefficients, but with a small amount on individually significant coefficients.

My overall results are consistent with the efficient market hypothesis.

References

Campbell, John Y. and Robert J. Shiller, 1988, "The dividend-price ratio and expectations of future dividends and discount factors," Review of Financial Studies 1, 195- 228.

Fama, Eugene and G. William Schwert, 1977, "Asset returns and inflation," Journal of Financial Economics 5, 115-146.

Fama, Eugene and Kenneth French, 1989, "Business conditions and expected returns on stocks and bonds," Journal of Financial Economics 25, 23-49.

Hong, Harris, Torous, Walter and Valkanok, Rossen, 2007 "Do Industries Lead Stock Markets?". Journal of Fiancial Economics 83, 367-396.

Hong, Harrison and Jeremy C. Stein, 1999, "A unified theory of underreaction, momentum trading and overreaction in asset markets," Journal of Finance 54, 2143-2184.

Industry Classification Benchmark, Industry Structure and Definitions Website.
< http://www.icbenchmark.com/ICBDocs/Structure_Defs_English.pdf >

Lamont, Owen, 2001, "Economic tracking portfolios," Journal of Econometrics 105, 161-184.

Merton, Robert C., 1987, "A simple model of capital market equilibrium with incomplete information," Journal of Finance 42, 483-510.

Shiller, Robert J., 2000, Irrational Exuberance, (Broadway Books: New York).

Sims, Christopher, 2001, "Rational Inattention," Princeton University Working Paper.

Torous, Walter, Ross Valkanov, and Shu Yan, 2004, "On predicting stock returns with nearly integrated explanatory variables," Journal of Business 77, 937-966.

Tse Yuman, 2015, "Do industries lead stock markets? A reexamination" Journal of Empirical Finance 34, 195-203.

Valkanov, Rossen, 2003, "Long-horizon regressions: Theoretical results and applications," Journal of Financial Economics 68, 201-232.