The 'Barn Door Closing Effect': application to the Portuguese mutual fund industry

Samuel Ferreira Correia

up201500299@fep.up.pt

Dissertation

Master in Finance

Supervisor:

Júlio Fernando Seara Sequeira da Mota Lobão, PhD

2018

Biographical Note

Samuel Ferreira Correia was born in Coimbra in 1993. He completed his undergraduate degree in Economics at University of Coimbra in January, 2015. Throughout his studies, he was involved in several activities. He did summer internships and was the founder and scientific coordinator of a project called *Olimpíadas da Economia*, developed at the University of Coimbra and created by a group of Economics' students. In September 2015, he started his Master course in Finance at School of Economics and Management, University of Porto, with the goal of studying the areas of corporate finance and financial markets. Here, he developed a major interest in the the area of behavioural finance. Along the second year of the master course he participated in a program called StartUp Voucher that assigns premiums to innovative ideas and products.

Acknowledgements

I would like to thank to everyone that directly or indirectly helped me to achieve this goal in my life. In particular to my family, mainly my grandparents, that were the core support in all my life. To my girlfriend and her family, for all the help and support during the last years and throughout the elaboration of the dissertation. Specially, a huge thanks to my supervisor, Professor Júlio Fernando Seara Sequeira da Mota Lobão, for his expertise, availability, advices, dedication and friendship along this work. Without them, this would not be possible. Once again, thank you all so much!

Abstract

Behavioural finance is nowadays one of the most interesting themes in finance, once the way investors think is affecting the world's transactions in financial markets. "Barn Door closing" is a phenomenon that consists in undertaking a behaviour in the future that would have been profitable in the past. There are several reasons underlying this phenomenon, as, for example, the uncertainty about the outcomes of financial markets. Some investors need to see a good performance of a mutual fund, measured by return, to invest. This will create a kind of "delay" in the investment because the investor will only earn money in the next period instead of earn it in the past.

Being the first time that this thematic is applied to European mutual funds, namely the Portuguese ones, we run a regression over a panel of 29 mutual funds from 2008 to 2012, to know if there is evidence of the phenomenon in the Portuguese market. Also, as we consider a period that comprises two different financial situations, one of pre-crisis and another of crisis, we run other regression with multiplicative dummies that interact with the explanatory variables, to know if the magnitude of the phenomenon is different in these two sub-periods.

We find out, in the overall period regression, evidence of the phenomenon in these type of mutual funds, despite of its low magnitude. This happens because the coefficient of the variable *return* is approximately zero. Concerning the impact of "barn door closing" in the two sub-periods, we conclude that there is no difference in its magnitude, because the variable that interacts with lagged *return* has no statistical significance.

Investors' performance is jeopardized. They can have higher profits by investing in the previous period instead of doing it in the following.

Key-words: Stocks; investment funds; "barn door closing"; return; financial crisis; flow of money.

JEL classification: G1, G11, G19

Resumo

Finanças comportamentais são, hoje em dia, um dos temas mais interessantes em finanças, uma vez que o pensamento dos investidores afeta as transações mundiais nos mercados financeiros. "*Barn Door Closing*" é um fenómeno que consiste em levar a cabo uma ação no futuro que teria sido lucrativa no passado. Há várias razões subjacentes a este fenómeno como, por exemplo, a incerteza acerca dos resultados dos mercados financeiros. Alguns investidores necessitam de observar um bom desempenho do fundo de investimento, medido pelo retorno, para investir. Cria-se, então, uma espécie de "atraso" no investimento, uma vez que o investidor irá ter lucro apenas no período seguinte em vez de o ter no anterior.

Sendo a primeira vez que esta temática é aplicada a fundos europeus, nomeadamente portugueses, fazemos uma regressão sobre um painel de 29 fundos com dados de 2008 a 2012, para verificar se há evidência do fenómeno no mercado nacional. Como estamos a considerar um período que envolve duas situações financeiras distintas, uma antes da crise e outra durante a crise, fazemos outra regressão com *dummies* multiplicativas que interagem com as variáveis independentes, para avaliar se a magnitude do fenómeno será diferente nestes dois subperíodos.

Verificamos, na regressão geral, que há evidência do fenómeno, apesar da reduzida magnitude. Isto acontece uma vez que o coeficiente da variável retorno é aproximadamente zero. Relativamente ao impacto do fenómeno nos dois subperíodos, concluímos que não há diferença na sua magnitude, uma vez que a *dummy* que interage com o retorno desfasado não tem significância estatística.

A *performance* dos investidores é, então, prejudicada, uma vez que podem ter lucros mais elevados investindo no período anterior em vez de o fazer no seguinte.

Palavras-chave: Ações; fundos de investimento; "*Barn Door Closing*"; retorno; crise financeira; fluxo de dinheiro.

Classificação JEL: G1, G11, G19

Contents

1. Introduction

Behavioural finance always has been one of the most interesting areas to study in finance and had gain more supporters because of the recent financial events that have occurred worldwide. Since the 2008 crisis that investors have been reluctant to invest. To do it, they need evidence that the market is going well, otherwise they will not invest because of their fear that a sort of crisis will happen again and make them run out of cash and, consequently, struggle.

Concerning the importance of this area, the theme that we propose to study is the "Barn door closing" phenomenon in Portuguese stocks mutual funds. It consists in undertaking a behaviour in the future that would have been profitable in the past.

The motivation to investigate this theme is related with personal and academic reasons, such as: it is an area of interest of the student and where the supervisor is specialized; there is a need to evaluate if the phenomenon happens in Portugal; know if the behaviour of investors can influence or not their actions and, consequently, financial markets; and there are few studies about this theme, what is a challenging opportunity for us to be pioneers in this area.

To develop this theme we use two main papers, among others, as basis, from Zeckhauser *et al.* (1991), and Patel *et al.* (1991).

The study is done for 29 Portuguese stocks mutual funds and the research sample comprises the period between 2008 and 2012. The data is collected from two main sources that are the following ones: *Comissão de Mercado de Valores Mobiliários* (CMVM) and *Associação Portuguesa de Fundos de Investimento, Pensões e Patrimónios* (APFIPP).

Since we focus on the behaviour of the investor, that our period of study includes two different sub-periods of financial and economic events, one before crisis (2008 to 2010) and other during crisis (2011 to 2012) and that in periods of crisis rational investors tend to become more reluctant to invest and it is expected that a higher increase in the return will lead to an higher increase in the flow of money to mutual funds in the following period, it is interesting to see if the magnitude of the phenomenon is different in these two periods.

Our main findings do not completely fulfil our expectations. We find out that the phenomenon exists in Portuguese stocks mutual funds between 2008 and 2012, but there is no difference in its magnitude in the two sub-periods previously described.

The structure of the study will proceed as follows. The second chapter presents the literature review that includes not only the analysis of papers written by other authors but also the main definitions to better understand the theme. Third chapter presents our hypothesis about the influence of the independent variables over the dependent one. Chapter four includes the data description while chapter five presents the empirical results and analysis. In chapter six we present the conclusions about the study where we discuss the main findings and some possible topics for further research.

2. Literature review

As previously said, "Barn Door Closing" is a phenomenon that "consists in undertaking a behaviour today that would have been profitable yesterday. Investors seek to reproduce actual or imagined past investment successes by investing today in the same way." (Zeckhauser *et al.*, 1991, p. 266). To better understand this phenomenon, try to imagine a stable in a farm. If a horse or another animal run away, the owner will construct a better stable that will prevent the horses from get away. So, the owner will try to "recover lost ground" by having better conditions for his horses. The same happens with investors that could have invested in the past and enjoy the profits instead of waiting for the market outcome to decide whether to invest or not.

"Barn door closing" phenomenon can also be compared to the behaviour of a person that did something wrong and then "seeks to contain postdecision regret and makes attempts to remove reminders of past errors through present choices" (Roy and Zeckhauser, 2015, p. 54).

Nowadays, the rationality of investors is becoming one of the major determinants of investment in financial markets. Zeckhauser *et al.* (1991) start to state that "some of the participants in financial markets do not act with perfect rationality: that is, they make mistakes. But more astute players stand ready to capitalize on their mistakes- for example, to buy when panic leads to speculative market crashes" (Zeckhauser *et al.*, 1991, p. 257). Here, there is an insight that some investors in financial markets do not act with rationality.

In the same paper, financial markets are seen "as an ecosystem, populated by rational and non-rational investors" (Zeckhauser *et al.*, 1991, p. 265). In consequence, they expose four different behavioural hypotheses of non-rational investors: barn door closing, expert and reliance effects, status quo bias and illusions, framing and data packaging. However, in this dissertation, we only explore the barn door closing, defined as in the beginning of this chapter. If market is said to be efficient, "prevailing asset prices capture all rational expectations about future prospects; it is impossible to identify profitable strategies" (Zeckhauser *et al.*, 1991, p. 266).

By using an econometric model to explain the annual flows to open-end no-load growth mutual funds, the authors, in the same paper, conclude that the barn door closing phenomenon, also called past performance, verifies: “a return 1% above the cross-sectional mean return in the previous period implies a $200,000 increased flow in this period” (Zeckhauser *et al.*, 1991, p. 270). So, when market goes up in the previous period, investors tend to invest more money in the following.

Patel *et al.* (1991) introduces two behavioural hypotheses instead of the previous four, to try to explain the financial phenomena. These two hypotheses are the barn door closing and the herd migration behaviour. Here, Barn door closing is defined as in the previous paper.

The authors argue that net flows of funds from individuals to the mutual fund sector can be explained by the barn door closing behaviour for two reasons. The first one is that mutual fund purchasers “rely on trends/patterns” (Patel *et al.*, 1991, p. 233), and the second one is that they “engage in personal window dressing (realigning their portfolio to a desirable composition for the sample period experienced) (Patel *et al.*, 1991, p. 233).

They also state that “if the barn door closing hypothesis is germane, individuals will buy more mutual funds after the stock market goes up, and sell after it plunges” (Patel *et al.*, 1991, p. 233), what is in line with what was stated earlier in this work. In fact, by doing an econometric regression and by running significance tests, the authors of the paper achieved the following result: “In a regression of f on four of its own lags, changes in Treasury bill interest rates, and returns on the equity market (proxied by the value-weighted NYSE index), we observe an economically large and statistically significant positive coefficient on equity returns that is consistent with barn door closing.” (Patel *et al.*, 1991, p. 234), where f is “the fraction of U.S. household sector's flow of financial purchases (composed of direct and intermediated net purchases of equities, bonds, and short-maturity or demand deposits) directed to mutual funds” (Patel *et al.*, 1991, p. 234). The existence of the phenomenon is also corroborated with their early findings in 1990: “In cross-sectional time-series regressions (for 96 funds over 1975-87), we explain 76 percent of the variance (R-squared) and highlight three interesting behavioural influences”, including “Past performance: A one percentage point return higher than the

average fund's return implies a $200,000 increased flow in the next year" (Patel *et al.*, 1991, p. 233).

Langevoort (1992) found out that "familiar concepts like "herd migration behavior", "barn door closing", and "Monday morning quarterbacking" have been observed, along with the presence of psychological barriers in daily trading activity" (Langevoort, 1992, p. 869).

Investors that follow the barn door closing phenomenon, i.e., that undertake a behaviour in the future that would have been profitable in the past are considered uninformed or non-rational, as it can be seen in Brennan and Cao (1996): "With multiple trading sessions, uninformed investors behave as rational trend followers, while more informed investors follow a contrarian strategy" (Brennan and Cao, 1996, p. 163). This conclusion was obtained by the authors with the help of the Hellwig's (1980) model, that was used to "analyse the value of improving trading opportunities by more frequent trading in the underlying asset, or by trading in a derivative asset" (Brennan and Cao, 1996, p. 163).

From the analysis of the papers, it can be seen that the "barn door closing" phenomenon is a constant in financial markets. Investors that are considered as irrational and trend-followers wait for the movements of the market to invest. So, there is a pattern that tells us that investors invest more money in the following period when the stock market had gone up in the previous, and vice-versa.

However, not all investors are irrational and trend-followers. For the more informed ones, there is an opportunity of poaching, because they "stand ready to capitalize on the mistakes of the naïve" (Patel *et al.*, 1991, p. 232).

In conclusion, it is possible to see that return is a key variable to explain the major movements of money in financial markets. Higher returns leads to higher flows and lower returns leads to lower flows. Investor's behaviour is in fact an important piece of the puzzle that are financial markets.

In the next chapter of the dissertation we develop some hypothesis used as basis for our econometric study.

3. Hypothesis development

In this chapter, based on the theories and empirical evidence of the previous studies presented in the literature review and on the expectations about the impacts of the independent variables over the dependent variable (positive or/and negative), we introduce our research hypotheses and the most important determining factors for the flow of money to stocks mutual funds.

Since we use a model from Zeckhauser *et al.* (1991) as basis, the hypotheses are formulated using the variables below.

i. Return

According to Zeckhauser *et al.* (1991), Patel *et al.* (1991), Langevoort (1992), Brennan and Cao (1996) and Roy and Zeckhauser (2015), the barn door closing phenomenon exists. This means that an increase in the return of a stock mutual fund in the previous period will produce an increase in the flow of money to the same fund in the next one. Following this, the expected relation between lagged return (return of the previous period) and flow of money to stock mutual funds in the actual period will be positive.

Therefore, we hypothesize:

Hypothesis 1 (H1): There is a positive/direct relation between lagged return and flow of money to stocks mutual funds.

ii. Flow of money to stocks mutual funds

According to Zeckhauser *et al.* (1991), financial markets are seen "as an ecosystem, populated by rational and non-rational investors" (Zeckhauser *et al.*, 1991, p. 265).

Knowing that fund purchasers "rely on trends/patterns" (Patel *et al.*, 1991, p. 233) what makes them non-rational investors, we suppose, without having an empirical study as basis, that if in the previous period the flow of money to stocks mutual funds have increased, the same will happen in the following period. The underlying reason is the fact exposed in **H1**: to invest in stocks mutual funds, investors need an increase in the return in the same fund in the previous year. So, if investors follow trends and patterns, the flow of money to stocks investment funds in one period will be influenced by the same variable in the previous period in a positive way, because there is a kind of herd migration

behaviour (the tendency of follow the actions, rational or non-rational, of other group of investors). This means that an increase in the lagged (previous period, in this case) variable will create an increase on the value of the same variable in the next period.

Therefore, we hypothesize:

Hypothesis 2 (H2): There is a positive relation between lagged flow of money to stocks mutual funds and the same variable in the next period.

iii. Size

The size of a fund is related with its total value, presented in euros once we are studying the Portuguese market. To be more precise, fund size is the month-end net assets of the mutual fund. If the value of net assets is high, this means that the fund has a good performance and therefore has more popularity among investors.

So, without having empirical studies to prove this relation and without running econometric tests, we suppose that investors, specifically the non-rational ones, think that it is a stable investment although its risk characteristics, what makes them want to invest. So, the trend is as follows: the higher the size of the mutual fund, the higher will be the flow of money to these type of funds.

Therefore, we hypothesize:

Hypothesis 3 (H3): There is a positive relation between the size of a mutual fund and the flow of money to the same fund.

The following table, based on the previous hypotheses, summarizes the expected impact of each variable over the flow of money to stocks mutual funds.

Variables (all variables are lagged by one period)	**Expected impact over flow of money to stocks mutual funds in *t***
Return	Positive
Flow of money to stocks mutual funds	Positive
Fund size	Positive

Table 1: Expected impact of the key variables over the flow of money to stocks mutual funds

However, as it will be seen later on in the dissertation, as we are considering two sub-periods of time, the expected relations between dependent and independent variables can change through time. The main reason can be the fact that during periods of crisis investors became more risk averse and instead of assuming some amount of risk and earn higher returns, they may prefer lower returns but no, or reduced risk. Also, the magnitude of the changes in the flow of money can diversify through time due to the same reason.

These considerations will be analysed when we present our econometric analysis in chapter 5.

4. Data description

In this dissertation, the sample is a panel of twenty nine Portuguese stocks mutual funds that ranges from 2008 to 2012. The ones that are used in the study are divided in four categories as follows:

- National funds, composed by *Alves Ribeiro Médias Empresas Portugal*, *Banif Ações Portugal*, *Barclays Premier Ações Portugal*, *BPI Portugal*, *Espírito Santo Portugal Ações*, *Millenium Ações Portugal* and *Santander Ações Portugal*;
- North American funds, composed by *BPI América*, *CaixaGest Ações EUA*, *Espírito Santo Ações América* and *Millenium Ações América*;
- European Union, Switzerland and Norwegian funds, composed by *Banif Euro Ações*, *BPN Ações Europa*, *CaixaGest Ações Europa*, *Crédito Agrícola – Raiz Europa*, *Espírito Santo Ações Europa*, *Millenium Eurocarteira*, *Montepio Ações* and *Montepio Ações Europa*;
- Other international funds, composed by *BPI Reestruturações*, *BPN Ações Global*, *CaixaGest Ações Emergentes*, *CaixaGest Ações Japão*, *CaixaGest Ações Oriente*, *Espírito Santo Ações Global*, *Espírito Santo Mercados Emergentes*, *Espírito Santo Momentum*, *Millenium Ações Japão* and *Millenium Mercados Emergentes*.

In the case of national mutual funds, we consider all funds for which we found values for the variables. However, in the other cases, we use a criteria to choose them. We only consider the funds of banks or other financial institutions that have been founded in Portugal.

We use monthly data for the three variables of the study, *return*, *flow* and *size*, that have been collected from APFIPP (*Associação Portuguesa de Fundos de Investimento, Pensões e Património*) and CMVM (*Comissão do Mercado de Valores Mobiliários*). It is used in two different ways, as follows: in chapter 5, we start by running a regression that let us know if the phenomenon happens during the entire period of the analysis; and, following this, we evaluate if the magnitude of the phenomenon varies between the pre-crisis and crisis periods by including a multiplicative dummy that interact with the explanatory variables.

In the next sub chapters, we present the dependent variable and independent variables, and the correspondent descriptive statistics. We also describe the methodology used in our study.

4.1 Dependent variable

Since the main purpose of our research study is to evaluate whether the "barn door closing" phenomenon exists or not, and knowing that it is done by studying the impact of some key variables over the flow of money to stocks mutual funds, the dependent variable must be the *flow* of money to these types of funds. This variable is obtained by subtracting the total redemptions to the total subscriptions.

To get the values attributable to the dependent variable, we hand collected data from APFIPP by consulting the monthly statistical report of securities mutual funds and looking to the column of the liquid subscriptions in each month.

4.2 Independent variables

Following the presentation of the hypothesis in chapter 3, we can see that the independent variables are the *flow* of money to stocks mutual funds, its *return* and its *size*, all lagged by one period.

i. Return

The returns of the funds have been taken from an excel file, collected from CMVM website, with data about the value of the participation units for each day. We used a logarithmic transformation to obtain the monthly returns of each fund.

ii. Size

As previously stated, the size of a fund is the total value of the fund, in this case in euros, or its month-end net assets.

The data for this variable has been taken from APFIPP website by consulting the monthly statistical report of securities mutual funds and looking to the column of the portfolio value in each month.

iii. Flow of money to stocks mutual funds

The flow of money to stocks mutual funds is both a dependent and an independent variable. This happens because, as stated in chapter 3, the lagged value of this variable can have an impact over the same variable in the next period.

As in the case of the dependent variable, the data for the independent one was taken from the APFIPP website.

4.3 Descriptive Statistics

The table below presents the descriptive statistics for the 29 stocks mutual funds over the period 2008 – 2012.

Descriptive Statistics, using observations 1:01 -29:52					
Variable	**Mean**	**Median**	**S.D.**	**Min.**	**Max.**
Return	-0.0114	-0.005	0.0613	-0.3	0.157
Flow	-4.43E+05	-6.57E+04	2.53E+06	-3.48E+07	1.36E+07
Size	3.24E+07	1.87E+07	3.75E+07	2.21E+06	2.72E+08

Table 2: Descriptive statistics for the 29 stocks mutual funds over the period 2008-2012

4.4 Methodology

As previously stated, in order to know if the "Barn Door Closing" phenomenon really happens in the Portuguese market, it is necessary to develop a model and run some tests over the variables that compose it.

As basis, we chose to use a model that have been developed in a paper from Zeckhauser *et al.* (1991), to explain the annual flows to open-end (type of mutual fund that does not have restrictions in the amount of shares that the fund can issue) no-load (type of mutual fund where shares are sold without a commission or sales charge) growth mutual funds.

By looking at the paper, we thought that the most suitable model (basic empirical model) to do our study would be as follows:

$$Flow_{it} = \alpha_t + \beta_1 Flow_{i,t-1} + \beta_2 Size_{i,t-1} + \beta_3 Return_{i,t-1} + \mu_{it}, \quad \textbf{(4.4.1)}$$

Where:

- Flow is the difference between the inflow and outflow of money to a specific mutual fund, that can be positive or negative;
- Size is the total value of the fund in each period;
- Return is the gain or loss over each period arising from investing in the fund;
- (i) refers to the mutual fund that is being considered;
- (t) refers to time.

In order to deal with panel data, the most common methods, suggested by the majority of the literature, are the pooled OLS, fixed effects and random effects model. To evaluate which one is the most suitable to do the study, we perform the appropriate tests in the next chapter of this work.

Taking into account that we are performing fixed and random effects models, the error term, μ_{it}, is composed by α_i, that captures all unobserved, time-constant factors that affect the dependent variable, and ε_i, which represents the unobserved factors that change over time and affect $Flow_{it}$ (Wooldridge, 2013).

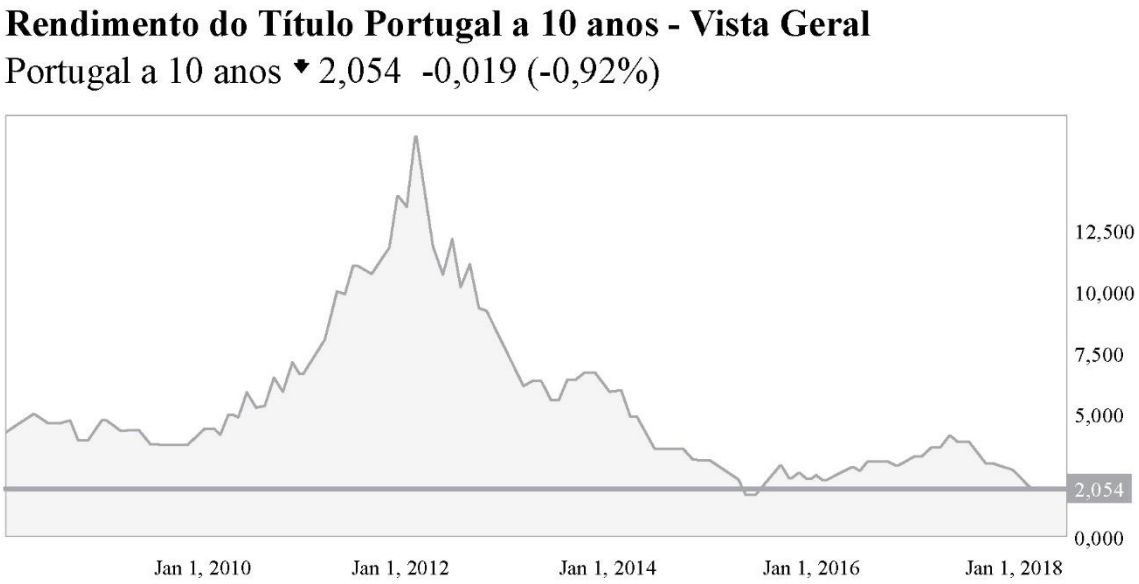
In this study, we also want to evaluate the impact of this phenomenon over two different periods, one before the financial crisis and another during it. In order to do it, we include multiplicative dummies in the model that are equal to 1 during crisis and 0 otherwise. So, the regression equation will be as follows:

$$Flow_{it} = \alpha_t + \beta_1 Flow_{i,t-1} + \beta_2 Flow_{i,t-1} * d + \beta_3 Size_{i,t-1} + \beta_4 Size_{i,t-1} * d + \\ + \beta_5 Return_{i,t-1} + \beta_6 Return_{i,t-1} * d + \beta_7 * d + \mu_{it}\,, \qquad \textbf{(4.4.2)}$$

Where d is the dummy variable and the others have the same meaning as in the case of the basic empirical model.

Since our data ranges from 2008 to 2012, we have to define the pre-crisis and crisis periods. From previous researches, we can see that there is not a consensus about the crisis period in Portugal. However, by looking at the yields of the Portuguese bonds, that can be seen in the graphic 1 below, the crisis, in Portugal, started in the second half of 2010, reaching the peak at the end of 2011, beginning of 2012, where the yields of the

bonds achieved a high value. In fact, there is a sharp increase in the yield between the second half of 2010 and the beginning of 2012.



Graphic 1: 10-year Portuguese Bond Yields (Source: Investing.com)

Since we have limited data, we consider 2011 and beggining of 2012 as the period of the Portuguese crisis, being 2008 – 2010 the pre-crisis period. Following this, the dummy variable will be equal to 1 during 2011 and beggining of 2012, and to 0 otherwise.

5. Empirical Results and Analysis

In this chapter we present the main results of the regressions analysis. It is divided into two sub chapters. In the first one, we approach the basic empirical model and its conclusions, while in the second one we do an analysis between the two different sub-periods that have been previously described.

5.1 Basic Empirical Model

Since we are dealing with panel data, we have three different models to estimate the regression. We have pooled OLS, random effects and fixed effects. In order to choose the model that suits best to our purposes, we have to do three main tests. F-test allows us to chose between pooled OLS and fixed effects, where a p-value below 0.05 make us reject the null hypothesis of using the first model. The Breusch-Pagan test let us choose between pooled OLS and random effects, where a p-value below 0.05 make us reject the null hypothesis of using the pooled OLS method. By its turn, the Hausman test let us choose between random and fixed effects. A p-value below 0.05 lead us to reject the null hypothesis and conclude that the fixed effects model is better.

So, as we are using GRETL, we start by running equation **(4.4.1)** with OLS method. Then we do the panel diagnosis through the Gretl's section test to obtain the results that let us compare the three previous models.

As we can see in the table below and in annex A, section 2 of the basic empirical model, the panel diagnosis of GRETL produces results for two of three tests.

Test	Models	Statistical value	P-value
F-test	Pooled OLS vs Fixed effects	2.53096	1.98186e-005 (<0.05)
Hausman test	Random effects vs Fixed effects	70.7875	2.89482e-015 (<0.05)

Table 3: Panel diagnosis of the basic empirical model

F-test, that allows us to choose between pooled OLS and fixed effects, shows a p-value below 0.05 leading us to reject the null hypotesis and concluding that the fixed effects model is better. By its turn, the Hausman test, that allows us to choose between random

and fixed effects, presents a p-value below 0.05 leading us to reject the null hypothesis of random effects and concluding that the fixed effects is better. If fixed effects is better than OLS and random effects, it will be the best method to estimate the model.

So, we run equation **(4.4.1)** through the fixed effects model. We consider robust standard errors to correct for any heteroskedasticity that may appear. So, the output of our basic empirical model estimated with fixed effects can be seen below:

	Coeficiente	Erro Padrão	Rácio-t	Valor p
Const	1.08174e+06	178430	6.063	1.55e-06 ***
Size_1	-0.0414694	0.00603702	-6.869	1.83e-07 ***
Return_1	2.13857e+06	654363	3.269	0.0029 ***
Flow_1	0.178010	0.0644900	2.760	0.0101 **

Media var. dependente	-370935.1	**D.P. var. dependente**	2360259
Soma res. quadrados	6.44e+15	**E.P. da regressão**	2109793
LSDV R-quadrado	0.217734	**Dentro R-quadrado**	0.153876
Log. da verosimilhança	-23619.78	**Critério de Akaike**	47303.56
Critério de Schwarz	47473.13	**Critério Hannan-Quinn**	47366.78
Rho	0.001887	**Durbin-Watson**	1.949424

Teste conjunto em regressores designados:
Estatística de teste: F(3;28) = 56.3515 com
valor p = P(F(3;28)>56.3515) = 5.52055e-012

Teste robusto para diferenciar grupos de intercepções no eixo x = 0:
Hipótese nula: Os grupos têm a mesma intercepção no eixo x=0
Estatística de teste: Welch F(28;515.5) = 5.42548 com
valor p = P(F(28;515.5)>5.42548) = 3.27085e-016

Table 4: Gretl output of the basic empirical model estimated with fixed effects

By looking at the model, we can see that despite of the low within R-squared of nearly 15 %, the F-statistic presents a really high value of 56.3515. In fact, with a p-value lower than 0.05 we reject the null hypothesis, what let us conclude that the explanatory variables are globally significant to the model. Allied to this, if we look to the inidividual significance of the variables, we can see that all of them are siginificant at 1 %, except

flow that is at 5 %. This means, respectively, that there is only 1% and 5% chance of rejecting the null hypothesis when it is true. Concerning collinearity, the usual way to see if it exists or not is by looking at the variance inflation factors. However, as we are using Gretl software and working with fixed effects, the collinearity problem does not exist because the software automatically eliminates the variables that have signs of multicollinearity. In our case all the initial variables are present in the model what indicates that there are not evidence of collinearity. Also, by looking at the value of the Durbin-Watson statistic, 1.949424, we can see that there is no evidence of autocorrelation between the residuals of the model, since the value is approximately 2.

In chapter 3 we hypothesized three different relations between the independent and dependent variables. Now, after having the model, it is necessary to evaluate if they still hold or not.

Hypothesis 1 tells that there is a positive relation between lagged *return* and *flow* of money to stocks mutual funds. By looking at the basic empirical model, we can see that this positive relation still holds.

By its turn, hypothesis 2 tells that there is also a positive relation between lagged *flow* of money to stocks mutual funds and the same variable in the next period. This expectation is also corroborated by the model, as we can see in annex B, section 1 of the basic empirical model.

Finally, hypothesis 3 that tells that there is a positive relation between the lagged fund *size* value and *flow* of money to stocks mutual funds do not hold. By looking at the model, we can see that the sign of the coefficient is negative what tells us that the bigger the fund size, the lower will be the flow of money to the mutual fund. This contradicts our early predictions of a positive relation between these two variables.

However, our focus is on the coefficient of lagged *return*. By looking at the output of the basic empirical model, we can see that there is evidence of the “barn door closing” phenomenon. An increase of 1 % on lagged *return* leads to an increase on the *flow* of money to stocks mutual funds. However, as the coefficient value is near 0, this increase will be really low.

5.2 Before and during crisis: is the impact of the phenomenon the same?

In this sub-chapter we focus on equation **(4.4.2)**. This is an equation that include a dummy variable used to evaluate if the magnitude of the effect in study is different in the pre-crisis and crisis periods.

To do this, we include a dummy variable named *crisis* that assumes value 1 during the period of crisis, 2011 and 2012, and 0 during the pre-crisis, 2008 to 2010. By looking at equation **(4.4.2)**, we can see that the dummy variable is multiplicative, meaning that it is multiplied by all the independent variables in the model. To obtain the data correspondent to these multiplication we multiplied the dummy by the values of the independent variables, obtaining the following ones: *Size_crisis*, *Return_Crisis* and *Flow_crisis*.

In this case, as we are also dealing with panel data, it is demanded to evaluate which model between pooled OLS, fixed effects and random effects is more appropriate. To do this, we start by running the equation using the OLS method. Then we do the panel diagnosis that let us decide between all the possible models.

By looking at the table below and to annex A, section 2 of the pre-crisis vs crisis model, we can see the panel diagnosis output.

Test	Models	Statistical value	P-value
F-test	Pooled OLS vs Fixed effects	3.12907	8.803e-008 (<0.05)
Hausman test	Random effects vs Fixed effects	88.4757	6.27963e-017 (<0.05)

Table 5: Panel diagnosis of pre-crisis vs crisis model

F-test, that let us choose between pooled OLS and fixed effects, presents a p-value lower than 0.05 leading us to reject the null hypothesis that the OLS is better and to non-reject the fixed effects model. Hausman test also presents a p-value below 0.05 what let us reject the hypothesis that the random effects model is better and to non-reject the fixed effects model. So, if the fixed effects model is better than random effects and pooled OLS, we will estimate equation **(4.4.2)** through the first method.

The estimation, done with robust standard errors in order to correct for any heteroskedasticity that may appear, can be seen in Annex B, section 1 of the pre-crisis vs

crisis model. Immediately, by looking at the output of the program we can see that there are three variables with no statistical significance: *Flow_crisis*, *Return_Crisis* and *Crisis*. So, in order to have a correct specification of the model we have to eliminate all the variables that are not significant.

First, we estimate a model without *Flow_crisis* and *Return_Crisis*. Then, we estimate another without the isolated dummy, *Crisis*.

	Coeficiente	Erro Padrão	Rácio-t	Valor p
Const	1.25271e+06	170871	7.331	< 0.0001 ***
Size_1	-0.0434207	0.00587181	-7.395	< 0.0001 ***
Return_1	2.13690e+06	648430	3.295	0.0027 ***
Size_crisis_1	-0.0127113	0.00437797	-2.903	0.0071 ***
Crisis	7132.01	106887	0.06672	0.9473
Flow_1	0.177133	0.0656666	2.697	0.0117 **

Media var. dependente	-370935.1	**D.P. var. dependente**	2360259
Soma res. quadrados	6.35e+15	**E.P. da regressão**	2096852
LSDV R-quadrado	0.228369	**Dentro R-quadrado**	0.165379
Log. da verosimilhança	-23609.66	**Critério de Akaike**	47287.32
Critério de Schwarz	47467.49	**Critério Hannan-Quinn**	47354.48
Rho	-0.001896	**Durbin-Watson**	1.960074

Teste conjunto em regressores designados:
Estatística de teste: F(5;28) = 110.484 com
valor p = P(F(5;28)>110.484) = 1.54493e-017

Teste robusto para diferenciar grupos de intercepções no eixo x = 0:
Hipótese nula: Os grupos têm a mesma intercepção no eixo x=0
Estatística de teste: Welch F(28;515.8) = 5.43072 com
valor p = P(F(28;515.8)>5.43072) = 3.10885e-016

Table 6: Gretl output of the pre-crisis vs crisis model without *Flow_crisis* and *Return_Crisis*

By looking at the output of the first model presented above, we can see that despite of the low within R-squared of nearly 17%, the F-statistic presents a value of 110.484 and a p-value lower than 0.05. This means that the regression is globally significant. Also, if we

look to the individual significance of the explanatory variables, we conclude that, despite of the dummy, all of them are statistically significant at 1% with exception of *flow* that is at 5%. By the same reasons presented in the case of the basic empirical model, this one also seems to not have any multicollinearity problems. The Durbin-Watson statistic presents a value near 2, 1.960074, meaning that there is no evidence of autocorrelation between the residuals of the model.

In chapter 3 of the study we hypothesized three different relations between the independent and dependent variables. As in the case of the basic empirical model, the hypothesis hold the same, meaning that there is a positive relation between lagged *return* and *flow*, lagged *flow* and the same variable in the following period and a negative relation between lagged *size* and *flow*.

In this model we also have multiplicative dummies. In this case there is only one, lagged Size_Crisis, because the others, as previously explained, were not relevant to the model. The coefficient of this variable assumes a negative value when the dummy, *Crisis*, is 1. This means that during crisis the lagged value of the fund (*size*) reinforces the negative impact over the *flow* of money to stocks mutual funds in the following period.

Again, to evaluate if the "barn door closing" phenomenon verifies, our focus is on the coefficient of the variable return. As it can be seen in the output of the model, its coefficient is positive, meaning that an increase of 1 % on lagged *return* leads to an increase on the *flow* of money to stocks mutual funds in the following period. However, as in the case of the basic empirical model, the value of the increase is really low since the value of the coefficient is approximately 0.

Despite the phenomenon verifies, since we eliminate the interaction between the dummy variable and return, *Return_Crisis*, because it was not significant to the model, there will not be any difference between the magnitude of the phenomenon in pre-crisis and crisis periods. So, the effects of the lagged *return* over the *flow* of money will hold over the entire period, 2008 to 2012.

	Coeficiente	Erro Padrão	Rácio-t	Valor p
Const	1.25430e+06	172556	7.269	< 0.0001 ***
Size_1	-0.0434304	0.00588464	-7.380	< 0.0001 ***
Return_1	2.13887e+06	645574	3.313	0.0026 ***
Size_crisis_1	-0.0126002	0.00353112	-3.568	0.0013 ***
Flow_1	0.177080	0.0658564	2.689	0.0119 **

Media var. dependente	-370935.1	**D.P. var. dependente**	2360259
Soma res. quadrados	6.35e+15	**E.P. da regressão**	2096129
LSDV R-quadrado	0.228368	**Dentro R-quadrado**	0.165378
Log. da verosimilhança	-23609.66	**Critério de Akaike**	47285.32
Critério de Schwarz	47460.19	**Critério Hannan-Quinn**	47350.51
Rho	-0.001881	**Durbin-Watson**	1.960051

Teste conjunto em regressores designados:
Estatística de teste: F(4;28) = 137.237 com
valor p = P(F(4;28)>137.237) = 5.75758e-018

Teste robusto para diferenciar grupos de intercepções no eixo x = 0:
Hipótese nula: Os grupos têm a mesma intercepção no eixo x=0
Estatística de teste: Welch F(28;515.5) = 6.0919 com
valor p = P(F(28;515.5)>6.0919) = 8.25829e-019

Table 7: Gretl output of the pre-crisis vs crisis model without *Flow_crisis, Return_Crisis and Crisis*

By looking at the output of the model without the dummy variable that is presented above, we can see that it is well specified according to the standards that we have been following in all of the previous models. All variables are globally and individually significant despite of the low within R-squared, and the Durbin-Watson statistic do not show any evidence of autocorrelation between the residuals.

Looking at the signs and values of the coefficients, we can see that all the hypothesis previously described are in line with what we have been writing.

Also in this model, there is evidence of the phenomenon but there is no difference in its magnitude between pre-crisis and crisis periods. Again, the effect of "barn door closing" will hold during the entire period.

6. Conclusions

As it was explained in the beginning of the study, the main goal is to evaluate if the "barn door closing" phenomenon occurs in Portuguese stocks mutual funds during the period between 2008 and 2012, by using a panel of 29 funds. As the data includes two different economic periods that range from 2008 to 2010 (pre-crisis) and 2011 to 2012 (crisis), we also felt the need to evaluate if the magnitude of the phenomenon is different in these two sub-periods.

Looking at the overall time range, 2008 to 2012, and to the sub-period analysis, we can see that, with exception of the variable size, all the others follow the expectations that we have described in the hypothesis development chapter. Concerning *size*, it was expected that a higher value of the fund, motivated by a higher value of the participation units or a higher level of investment in the mutual fund, would lead to an increase in its flow of money. Normally, these two situations leads the investor to believe that the fund is having a good performance. However, in this study, the expectation about the sign of the coefficient was not what we initially thought.

Focusing on the variable *return*, we can see that there is evidence of the phenomenon in the overall period of the study. However, despite the lagged return have a level of significance of 1%, the impact of an increase of 1% in *return* in the previous period will lead to a low increase in the *flow* of money to the mutual fund in the next period, because its coefficient is approximately 0.

Concerning the analysis of the pre-crisis and crisis periods, we can see that there is no difference in the magnitude of the phenomenon. This happens because when we run the model with dummies, the one that interacts with lagged return has no statistical significance, what obliges to take it from the model.

Globally, we can see that the phenomenon verifies in the Portuguese stocks mutual funds during the overall period of the study and that there is no difference in its magnitude in the pre-crisis and crisis periods.

Knowing that it is still not a widely explored theme, we suggest the study of this phenomenon in other type of funds and other financial securities as further research. It might also be interesting to do this study over a larger period of time in order to evaluate

if the phenomenon happened or not some years ago when there was not the availability of information of today.

References

Brennan, M. J. and Cao, H. Henry (1996), "Information, Trade and Derivative Securities", *The Review of Financial Studies*, Vol. 9, No. 1, pp. 163-208.

Devjani, R. and Zeckhauser, R. (2015), "Grappling with Ignorance: Frameworks from Decision Theory, Lessons from Literature", *Journal of Benefit-Cost Analysis*, Vol. 6, pp. 33-65.

Langevoort, Donald C. (1992), "Theories, assumptions and securities regulation: Market efficiency revisited", *University of Pennsylvania Law Review*, Vol. 140, pp. 851-920.

Patel, J., Zeckhauser, R., and Hendricks, D. (1991), "The rationality struggle: illustration from financial markets", *The American Economic Review*, Vol. 81, No. 2, pp. 232-236.

Wooldridge, J. M. (2013), *Introductory Econometrics: A Modern Approach*, (5th edition): South-Western.

Zeckhauser, R., Patel, J., and Hendricks, D. (1991), "Non rational actors and financial market behaviour", *Theory and Decision*, Vol. 31, pp. 257-287.

Annexes

Annex A – Statistic tests

Basic Empirical Model

1. Pooled OLS estimation

Modelo 1: Mínimos Quadrados de amostragem ("Pooled OLS"), usando 1479 observações

Incluídas 29 unidades de secção-cruzada
Comprimento da série temporal = 51
Variável dependente: Flow

	Coeficiente	*Erro Padrão*	*rácio-t*	*valor p*	
const	315990	74632,7	4,234	<0,0001	***
Size_1	−0,0175703	0,00156266	−11,24	<0,0001	***
Return_1	2,06635e+06	921387	2,243	0,0251	**
Flow_1	0,209081	0,0235463	8,880	<0,0001	***

Média var. dependente	−370935,1	D.P. var. dependente	2360259
Soma resíd. quadrados	6,76e+15	E.P. da regressão	2140231
R-quadrado	0,179422	R-quadrado ajustado	0,177754
F(3, 1475)	107,5049	valor P(F)	6,00e-63
Log. da verosimilhança	−23655,14	Critério de Akaike	47318,28
Critério de Schwarz	47339,47	Critério Hannan-Quinn	47326,18
rho	−0,007883	Durbin-Watson	1,984152

2. Panel diagnosis: Pooled OLS vs Fixed effects vs Random effects

Diagnósticos: usando n = 29 unidades de secção-cruzada

Estimador de efeitos fixos
permite diferenciar intercepções por unidade de secção-cruzada

	coeficiente	erro padrão	rácio-t	valor p	
const	1,08174e+06	128070	8,446	7,20e-017	***
Size_1	−0,0414694	0,00359786	−11,53	1,80e-029	***
Return_1	2,13857e+06	912334	2,344	0,0192	**
Flow_1	0,178010	0,0235687	7,553	7,51e-014	***

Variância dos resíduos: 6,44092e+015/ (1479 - 32) = 4,45122e+012

Significância conjunta da diferenciação das médias de grupo:
F (28, 1447) = 2,53096 com valor p 1,98186e-005
(Um valor p baixo contraria a hipótese nula de que o modelo Mínimos Quadrados (OLS) agrupado (pooled) é adequado, validando a hipótese alternativa da existência de efeitos fixos.)

Variance estimators:
between = 0
within = 4,45122e+012
theta used for quasi-demeaning = 0

Estimador de efeitos aleatórios permite uma componente de unidade-específica no termo do erro

	coeficiente	erro padrão	rácio-t	valor p	
Const	315990	74632,7	4,234	2,44e-05	***
Size_1	−0,0175703	0,00156266	−11,24	3,36e-028	***
Return_1	2,06635e+06	921387	2,243	0,0251	**
Flow_1	0,209081	0,0235463	8,880	1,91e-018	***

Estatística de teste de Hausman:
H = 70,7875 com valor p = prob (qui-quadrado (3)> 70,7875) = 2,89482e-015
(Um valor p baixo contraria a hipótese nula de que o modelo de efeitos aleatórios é consistente, validando a hipótese alternativa da existência do modelo de efeitos fixos.)

Pre-crisis vs Crisis Model

1. Pooled OLS estimation

Modelo 3: Mínimos Quadrados de amostragem ("Pooled OLS"), usando 1479 observações
Incluídas 29 unidades de secção-cruzada
Comprimento da série temporal = 51
Variável dependente: Flow

	Coeficiente	*Erro Padrão*	*rácio-t*	*valor p*	
const	383002	88494,0	4,328	<0,0001	***
Size_1	−0,0170376	0,00174287	−9,776	<0,0001	***
Return_1	2,86231e+06	1,01707e+06	2,814	0,0050	***
Flow_crisis_1	−0,0166533	0,0593430	−0,2806	0,7790	
Size_crisis_1	−0,00340656	0,00372440	−0,9147	0,3605	
Return_crisis_1	−4,39977e+06	2,43497e+06	−1,807	0,0710	*
Crisis	−197594	160329	−1,232	0,2180	
Flow_1	0,210160	0,0263282	7,982	<0,0001	***

Média var. dependente −370935,1 D.P. var. dependente 2360259

Soma resíd. quadrados	6,72e+15	E.P. da regressão	2137107
R-quadrado	0,184035	R-quadrado ajustado	0,180152
F (7, 1471)	47,39613	Valor P (F)	7,09e-61
Log. da verosimilhança	−23650,97	Critério de Akaike	47317,94
Critério de Schwarz	47360,33	Critério Hannan-Quinn	47333,75
rho	−0,008958	Durbin-Watson	1,987018

2. Panel diagnosis: Pooled OLS vs Fixed effects vs Random effects

Diagnósticos: usando n = 29 unidades de secção-cruzada

Estimador de efeitos fixos permite diferenciar intercepções por unidade de secção-cruzada

	coeficiente	erro padrão	rácio-t	valor p	
Const	1,28712e+06	139582	9,221	1,01e-019	***
Size_1	−0,0443859	0,00370467	−11,98	1,33e-031	***
Return_1	2,92951e+06	999883	2,930	0,0034	***
Flow_crisis_1	0,0824236	0,0593677	1,388	0,1652	
Size_crisis_1	−0,0115592	0,00380134	−3,041	0,0024	***
Return_crisis_1	−4,17196e+06	2,39403e+06	−1,743	0,0816	*
Crisis	−38116,8	158786	−0,2401	0,8103	
Flow_1	0,158472	0,0264892	5,983	2,77e-09	***

Variância dos resíduos: 6,33382e+015/ (1479 - 36) = 4,38934e+012

Significância conjunta da diferenciação das médias de grupo:
F (28, 1443) = 3,12907 com valor p 8,803e-008
(Um valor p baixo contraria a hipótese nula de que o modelo Mínimos Quadrados (OLS) agrupado (pooled) é adequado, validando a hipótese alternativa da existência de efeitos fixos.)

Variance estimators:
between = 0
within = 4,38934e+012
theta used for quasi-demeaning = 0

Estimador de efeitos aleatórios permite uma componente de unidade-específica no termo do erro

	Coeficiente	erro padrão	rácio-t	valor p	
Const	383002	88494,0	4,328	1,61e-05	***
Size_1	−0,0170376	0,00174287	−9,776	6,55e-022	***
Return_1	2,86231e+06	1,01707e+06	2,814	0,0050	***
Flow_crisis_1	−0,0166533	0,0593430	−0,2806	0,7790	

Size_crisis_1	−0,00340656	0,00372440	−0,9147	0,3605	
Return_crisis_1	−4,39977e+06	2,43497e+06	−1,807	0,0710	*
Crisis	−197594	160329	−1,232	0,2180	
Flow_1	0,210160	0,0263282	7,982	2,87e-015	***

Estatística de teste de Hausman:
H = 88,4757 com valor p = prob (qui-quadrado (6)> 88,4757) = 6,27963e-017
(Um valor p baixo contraria a hipótese nula de que o modelo de efeitos aleatórios é consistente, validando a hipótese alternativa da existência do modelo de efeitos fixos.)

Annex B – Models

Pre-Crisis vs Crisis Model

1. Fixed effects model

Modelo 4: Efeitos-fixos, usando 1479 observações
Incluídas 29 unidades de secção-cruzada
Comprimento da série temporal = 51
Variável dependente: Flow
Erros padrão robustos (HAC)

	Coeficiente	*Erro Padrão*	*rácio-t*	*valor p*	
const	1,28712e+06	184335	6,982	<0,0001	***
Size_1	-0,0443859	0,00641735	-6,917	<0,0001	***
Return_1	2,92951e+06	855257	3,425	0,0019	***
Flow_crisis_1	0,0824236	0,0796121	1,035	0,3094	
Size_crisis_1	-0,0115592	0,00497047	-2,326	0,0275	**
Return_crisis_1	-4,17196e+06	2,64257e+06	-1,579	0,1256	
Crisis	-38116,8	101383	-0,3760	0,7098	
Flow_1	0,158472	0,0662921	2,391	0,0238	**

Média var. dependente	-370935,1	D.P. var. dependente	2360259
Soma resíd. quadrados	6,33e+15	E.P. da regressão	2095076
LSDV R-quadrado	0,230742	Dentro R-quadrado	0,167946
Log. da verosimilhança	-23607,38	Critério de Akaike	47286,76
Critério de Schwarz	47477,53	Critério Hannan-Quinn	47357,88
Rho	-0,002381	Durbin-Watson	1,961121

Teste conjunto em regressores designados -
Estatística de teste: F (7, 28) = 76,7488
com valor p = P(F(7, 28) > 76,7488) = 1,40674e-016

Teste robusto para diferenciar grupos de intercepções no eixo x=0 -
Hipótese nula: Os grupos têm a mesma intercepção no eixo x=0

Estatística de teste: Welch F (28, 515,8) = 4,91108
com valor p = P(F(28, 515,8) > 4,91108) = 3,35202e-01