HAL
archives-ouvertes.fr

Essays in Empirical Financial Economics

Jean-Noël Barrot

▶ To cite this version:

Jean-Noël Barrot. Essays in Empirical Financial Economics. Business administration. HEC, 2012. English. <NNT : 2012EHEC0010>. <pastel-00829542>

HAL Id: pastel-00829542
https://pastel.archives-ouvertes.fr/pastel-00829542

Submitted on 3 Jun 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ECOLE DOCTORALE
Sciences du Management

Gestion – Organisation
Décision - Information

ECOLE DES HAUTES ETUDES COMMERCIALES DE PARIS

Ecole Doctorale « Sciences du Management/GODI » - ED 533

Gestion Organisation Décision Information

« ESSAYS IN EMPIRICAL FINANCIAL ECONOMICS »

THESE

présentée et soutenue publiquement le 25 octobre 2012

en vue de l'obtention du

DOCTORAT EN SCIENCES DE GESTION

Par

Jean-Noël BARROT

JURY

Président de Jury :	**Madame Antoinette SCHOAR** Professeur Massachusetts Institute of Technology - USA
Directeur de Recherche :	**Monsieur David THESMAR** Professeur, HDR Ecole des Hautes Etudes Commerciales Paris - France
Rapporteurs :	**Monsieur Denis GROMB** Professeur INSEAD Fontainebleau – France
	Monsieur Augustin LANDIER Professeur des Universités Université Toulouse 1 Capitole - France
Suffragants :	**Monsieur José ALLOUCHE** Professeur des Universités Université Paris 1 Panthéon Sorbonne - France
	Monsieur Ulrich HEGE Professeur Ecole des Hautes Etudes Commerciales Paris – France

Ecole des Hautes Etudes Commerciales

Le Groupe HEC Paris n’entend donner aucune approbation ni improbation aux opinions émises dans les thèses ; ces opinions doivent être considérées comme propres à leurs auteurs.

Essays in Empirical Financial Economics

Ph.D. thesis submitted by

Jean-Noël Barrot

Under the supervision of

David Thesmar

Members of the Committee

José Allouche

Denis Gromb

Ulrich Hege

Augustin Landier

Antoinette Schoar

Acknowledgements

Writing this thesis has been a fantastic experience. I've had the opportunity to interact and learn from many people who I admire and sincerely appreciate, and who I would like to thank here.

I am truly grateful to David Thesmar for his invaluable guidance and support. Our discussions have always been very inspiring for me, both inside and outside the academic sphere. Among many things, he taught me how to ask important questions. It's been both a pleasure and an honor to work under his supervision in these past few years.

I would like to thank Antoinette Schoar for giving me the opportunity to spend a fascinating and productive year at MIT Sloan. Her feedback has significantly sharpened my thoughts and greatly stimulated my research.

I am glad that José Allouche, Denis Gromb, Uli Hege and Augustin Landier kindly accepted to take part in my thesis committee. Their work has inspired mine throughout the Ph.D..

I am thankful to the members of the Finance department at HEC Paris for their feedback and advice. I acknowledge the support of the Fondation HEC, the Greghec and the Axa Research Fund for their financial support, which enabled me to work in the best possible conditions and present my research worldwide.

I had the opportunity to collaborate with several great researchers. I would like to thank Claire Lelarge for everything I learned from her and for her continuing support and encouragement. I'm also grateful to Ron Kaniel and David Sraer for having me on board our exciting research project.

I very much appreciated the support and friendship of my fellow Ph.D. candidates along the years. I would like to thank them and tell them how much I enjoyed sharing their office. On a similar note, I would like to mention the fruitful discussions I had with Ph.D. candidates at MIT. I sincerely hope we will continue to work together in the coming years.

Last, my sincere gratitude goes to my sisters and parents, to my friends and to Marina. They largely contributed to this thesis through their precious support and unconditional love. Thank you all for everything.

Executive summary

This dissertation is made of four distinct chapters. In the first chapter, I consider an exogenous restriction on the ability of French trucking firms to extend payment terms to their clients. I find that they provide trade credit at the cost of lower investment , lower return on assets, and higher default risk. I ask whether trucking firms trade off current profitability to lend to their constrained clients and extract a share in their future surpluses. Instead, I find that trucking firms provide more credit to their clients when the value of lending relationships is likely to be lower, i.e. when they face more competition. Altogether, these results point to a darker side of trade credit, whereby some firms extend trade credit at the cost of lower profitability.

In the second chapter, I consider investments made by private equity funds, which generally have a limited investment horizon contractually fixed *ex ante*. I use between and within-fund variations in investment horizon to show that funds with a longer horizon select younger companies at an earlier stage of their development, stage investment more and hold on to their investments for a longer period of time. Moreover, companies which receive funding from funds with a longer horizon increase their patent stock significantly more than companies which receive funding from investors with a shorter horizon. Altogether, these results provide new evidence on the behavior of private equity funds throughout their life cycle and suggest that investor horizon matters to an important extent for the funding of corporate innovation.

The third chapter presents a joint work with Ron Kaniel and David Sraer. We use detailed brokerage account data to provide a quantitative exploration of the behavior of retail investors during the financial crisis of 2008. We identify ex ante dimensions of heterogeneity based on trading frequency, amount invested or type of securities traded that are associated with different trading behavior ex post. We show that investors who appear more sophisticated on these dimensions in the pre-crisis period were, in the post-crisis period, less likely to flee to safety, more likely to engage in liquidity provisions and to earn higher returns. Our analysis thus extend our understanding of retail investors' behavior and provide a new light on one particular mechanics of the financial crisis.

In the fourth chapter, I develop the idea that households have an imprecise knowledge of their portfolio's exposure to systematic risk and that this leads them to make investment mistakes. This idea is tested in the context of the decision to actively trade rather than passively invest in the stock market. I show that the trading activity of individual investors increases (decreases) following high (low) performance. I carefully split individual investors performance into (i) the component of their performance related to the exposure of their trades to systematic risk factors and (ii) residual performance and show that their trading activity reacts to both. To account for these results, I contrast a story based on overconfidence against a simple model where individual investors have an imprecise knowledge of *both* their ability and systematic exposure ex ante, and learn about them as they trade.

Contents

List of Figures

List of Tables

Introduction

This dissertation tackles three important questions in corporate finance, some of which carry substantial policy implications.

Trade credit supports a large share of economic activity. Yet, there are diverging views on the decisions of firms to lend to their clients or borrow from their suppliers. On the one hand, trade credit is often seen as a last resort source of finance, and a substitute to bank credit in developing countries. On the other hand, *folk wisdom* suggests that small and credit constrained firms might be forced to extend credit to their powerful clients because of their weak competitive position. Interestingly both policies *supporting* and *restricting* trade credit have been undertaken recently. In the midst of the recent crisis, leaders from the G20 decided to pledge US$ 250bn in trade finance to curb the sharp drop in world trade. In contrast, the European Union recently adopted a directive reducing maximum payment terms, arguing that small and medium sized firms might finance a disproportionate share of trade credit at the cost of lower growth. The empirical evidence on trade credit has been relatively scarce, mainly because of the lack of appropriate data, and because the causal effect of trade credit on other corporate policies is hard to identify. Chapter 1 of this dissertation takes advantage of a quasi-natural experiment and of the richness of French firm-level data to contribute to this debate. The main findings are that some firms indeed provide trade credit to their clients at the cost of lower current investment and return on assets, and higher default risk. The rest of the chapter examines various explanations for this somewhat surprising result.

Frictionless financial markets should enable households to transfer wealth through time, making horizon irrelevant for investment opportunities to be funded. In practice, information asymmetries and agency problems drive agents to take decisions based on their investment horizon. Managers adopt myopic behaviors and forsake good investments so as to boost current earnings and mislead the market about their firms' worth (Stein, 1989) while arbitrageurs refrain from undertaking arbitrage opportunities to avoid facing redemptions from their investors before prices revert to fundamentals (Shleifer and Vishny, 1998). A related question is whether horizon matters for the funding of corporate innovation, an important driver of economic growth. The main econometric challenge associated with this question is the identification of a credible measure of horizon. Chapter 2 considers the case of private equity funds, which are major contributors to the funding of innovative startup companies. The fact that these funds have an investment horizon which is fixed ex-ante offers a unique opportunity to relate horizon to innovation. The results show that funds with a longer horizon select more innovative companies, and that they select less innovative companies as they get closer to the end of their investment life. These findings are interesting with respect to both our understanding of the role of investor horizon *and* the behavior of private equity funds throughout their life cycle.

The financial decisions of households have been the subject of a growing number of studies in the past few years. There are at least two reasons why this topic matters a great deal. First, as the world population ages, financial planning becomes a first order issue: investment mistakes can be very costly in the long run. Second, the allocation of households savings - or mis-allocation thereof - has an impact on the workings of financial markets themselves. The last two chapters of this dissertation investigate the decision by households to invest directly in the stock market. An extensive literature has shown that individual investors are prone to a number of behavioral biases which lead them to loose money and to potentially destabilize prices. The "noise trader approach to finance" (Shleifer and Summers, 1990) has generally seen individual investors broadly as a sentiment driven force pushing prices temporarily away from fundamentals. Chapter 3 analyzes the role of individual investors during the recent financial crisis and shows that while some individuals flee to safety, others provided liquidity to the bear markets. Chapter 4 tests whether some investment mistakes may be linked to the fact that households have a coarse understanding of their portfolio's exposure to risk factors. The results show that their trading activity depends on both their systematic exposure and their residual performance, and that the sensitivity of their trading activity to their systematic exposure decreases through time. This suggests that households may be learning about their systematic exposure through time, and make mistakes on the way. These findings add to our understanding of households financial management.

Chapter 1

A Darker Side of Trade Credit? Evidence from Trucking Firms

Abstract

I consider an exogenous restriction on the ability of French trucking firms to extend payment terms to their clients. I show that they provide trade credit at the cost of lower contemporaneous investment and return on assets, and higher default risk. I ask whether they might be trading off current profitability to offer credit to their constrained clients and extract a share in their future surpluses. Instead, I find that trucking firms provide more credit when the value of client relationships is likely to be lower, i.e. when they face more competition. Trade credit provision turns out to be higher in more competitive local markets, and to decrease following an increase in the barriers to entry in the trucking sector. Altogether, these results point to a darker side of trade credit, whereby some firms extend trade credit at the cost of lower profitability.

JEL classification: G32, G33, G34, D23

Keywords: Trade credit, investment, bankruptcy, corporate risk.

1.1 Introduction

Accounts payable represent the second largest liability on the aggregate balance sheet of non-financial businesses in the U.S., second only to bonds[1], and they finance a large share of global trade[2]. Yet, despite its significant economic importance, the dynamics of trade credit provision are not fully understood. Interestingly both policies *supporting* and *restricting* trade credit have been undertaken recently. In the midst of the recent crisis, leaders from the G20 decided to pledge US$ 250bn in trade finance to curb the sharp drop in world trade (Ahn et al., 2011). In contrast, the European Union recently adopted a directive reducing maximum payment terms, arguing that small and medium sized firms financed a disproportionate share of trade credit at the cost of lower growth[3].

I examine the decision of firms to extend trade credit to their clients. I use a quasi-natural experiment and a difference-in-differences approach and find that some firms offer payment terms to their clients at the cost of lower investment, increased default risk and lower return on assets. More specifically, I consider the response of French trucking firms to an exogenous restriction on their ability to extend payment terms to their clients beyond 30 days. This reform came into effect in 2006 and affected trucking activities exclusively. To build a relevant control group, I use the input-output matrix of the French economy and select the three-digit service sector closest to the trucking industry in terms of downstream client distribution. I then compare the change in various corporate policies of treated firms relative to control firms. I find that for one dollar less of receivables, the average trucking firm dedicates 5 cents to increase capex, 9 cents to reduce credit line use and 14 cents to reduce account payables. I replicate the analysis in the period prior to the reform to insure that these results are not driven by heterogeneous trends in the treatment and the control group. Moreover, I check whether these effects differ across firms exposure to credit constraints

[1] As of 2009, according to the U.S. Flow of Funds Account

[2] "Trade finance supports more than 80 per cent of global trade", in The Financial Times, April 8, 2012, "Banks test 'CDOs' for trade finance"

[3] See Financial Times, April 29, 2010, "EU late payment law moves a step closer".

based on size, payout status and age. While the effect on ROA seems to be prevalent across the board, the effect on capex is stronger among small, non dividend paying and young firms. Finally, the default probability of the average trucking firm drops sharply by a significant 50 percent relative to the default probability of control firms. The survival probability of newly created trucking firms also increases significantly following the reform.

A possible explanation for these somewhat surprising results is that firms lend to their constrained clients in the hope of sharing future surpluses. This is consistent with the common view in the literature, both theoretical and empirical, which sees trade credit provision as a way for suppliers to finance their temporarily constrained buyers. Petersen and Rajan (1997) use the National Survey of Small Business Finance (NSSBF) and document that firms with better access to credit from financial institutions offer more trade credit. They suggest that suppliers are willing to lend to constrained clients if they anticipate that growth in future business will compensate them for the risk they are taking. Why would a supplier accept to finance her client when other financial intermediaries would not? Biais and Gollier (1997) argue that suppliers have an informational advantage over other types of external investors (in particular, banks) and are therefore better at selecting solvent clients. Burkart and Ellingsen (2004) and Giannetti et al. (2008) suggest that it is typically less profitable for an opportunistic borrower to divert (specific) inputs than to divert cash, which increases the advantage of suppliers over banks in lending to their clients. In Cunat (2006) and Wilner (2000) the continuation value embedded in the relationship between the supplier and the client generates higher incentives for the client to repay the supplier than the bank. In turn, this increases the incentives for the supplier to provide liquidity insurance to a distressed client[4]. Hence, trucking companies might delay investment and be temporarily less profitable to offer credit to their constrained clients with investment opportunities, in the hope of sharing the future surplus.

[4]A fourth advantage of suppliers over bankers in supplying credit is related to the larger value they can extract from collateral (Petersen and Rajan, 1997; Burkart and Ellingsen, 2004; Fabbri and Menichini, 2010). This last advantage is probably not applicable to trucking firms since they might have as much difficulty as the banks to liquidate their clients' assets in case of default.

Unfortunately, I have no way to formally test whether the lower profitability of trucking firms prior to the reform was more than offset by the expected future cash-flow they were going to extract from their ongoing relationship with their clients. Instead, I ask whether trucking firms extend more credit when the value of relationships are likely to be higher. Petersen and Rajan (1995) show that credit market competition imposes constraints on the ability of a firm and its creditor to inter-temporally share surplus. In a competitive environment, the creditor faces a larger risk to be replaced and is reluctant to temporarily subsidize its client. A monopolistic creditor, on the other hand, expects to share the future surplus generated by the trade through the future rents she is able to extract. I hypothesize that the value of relationships with clients are higher when suppliers face less competition, in the spirit of Petersen and Rajan (1995). I use two different tests to check whether this is the case. I first compare the trade credit provision of trucking firms one year before the reform in areas where they face direct competition from other trucking firms and areas where they don't. I find that they extend more credit in more competitive areas. Since the value of relationships is likely to be lower in these areas, the results suggest that investing in client relationship is not the primary driver of trade credit provision. I go one step further and take advantage of an exogenous change to the barriers to entry in the trucking sector announced in 1999. Following this change, any trucking firm has to post a minimum amount of equity proportional to the number of trucks it operates. This change had a large impact on median book equity at creation and on the rate of entry in the trucking sector, which ensures that it indeed decreased the intensity of the competition in the trucking sector. In a difference-in-differences setting using the same control group as in the main analysis, I find that trucking firms decrease their provision of trade credit following the decrease in competition intensity, especially dominant and cash-rich ones. Again, this suggests that firms do not seem to sacrifice current ROA primarily to invest in client relationships.

These results seem to echo anecdotal evidence in Europe and the U.S.[5] as well recent

[5] See for instance: The Wall Street Journal, August 31, 2009, "Big Firms Are Quick to Collect, Slow to Pay" or The Wall Street Journal, June 7, 2012, "Big Customers are Taking Longer to Pay".

survey evidence (Robb and Reedy, 2012) according to which some firms in a weak competitive position are forced to extend long payment terms to their clients, especially in the recent financial crisis. There is evidence in the literature consistent with this view. Fabbri and Klapper (2008) show that Chinese SMEs with low bargaining power are more likely to extend trade credit and have a larger share of goods sold on credit. Giannetti et al. (2008) and Klapper et al. (2012) document respectively on the NSSBF and on a proprietary dataset of actual trade credit contracts that firms that have some buyer market power seem to receive more trade credit.

There is no evident *financial* reason why customers with more bargaining power should take more credit from their suppliers. If powerful customers intend to extract concessions from their suppliers, they would be better off by obtaining upfront price discounts from their supplier rather than obtaining financing which their supplier has probably little comparative advantage in providing. Trade credit provision by a powerless supplier to her powerful customer could make sense if the former is not credit constrained while the latter is. However, even in this case, suppliers should not turn out to be better off following a restriction to their ability to extend trade credit. Since financial reasons alone do not seem to account for this puzzle, the literature as came up with *operational* reasons why firms with more market power seem to receive more trade credit. Fisman and Raturi (2004) argue that a supplier with little bargaining power is willing to lend more to its customer because the customer makes more relationship specific investments when there is little risk she might be held up. Conversely, in Dass et al. (2011), a supplier with little bargaining power also lends more to its client in order to commit to make relationship specific investments, while a supplier with more bargaining power holds an implicit stake in the surplus of the trade and does not need to commit to do the relationship investment. Finally, Klapper et al. (2012) suggest that large buyers exercise their bargaining power by demanding longer ”trial” periods before they pay their suppliers. However, neither of these theories predict that the profitability of the supplier would increase following a restriction on trade credit provision. In fact, it is hard to think of a model with profit maximizing firms where a bargaining power differential between a supplier and a client drives the former to trade below her reservation profit.

The results highlighted above could be explained if there are private benefits of owning a trucking firm, as in Hurst and Pugsley (2011), and if these benefits offset the reduced profitability associated with extending long payment terms to clients. I have no way to formally test this assumption. However, the fact that ROA decreases for both small and large firms suggests that this might not be the only explanation. An alternative yet relatively unexplored hypothesis is suggested in Petersen and Rajan (1997): "trade credit may be a strategic tool for deep-pocket firms to increase the minimum scale of staying in the industry". Early contributions in the literature have already raised similar points. As emphasized by Caves and Porter (1977), incumbent firms may decide to expand as a strategic attempt to distort rivals' actions. In a similar spirit, Fudenberg and Tirole (1983) predict that incumbents may increase investment to deter the entry or expansion (or mobility) of potential rivals. If the long term costs of excess entry in the trucking sector outweigh the costs of additional credit provision, then incumbent firms may prevent the entry or expansion of new businesses by extending cheap credit to constrained clients and setting standard payment terms in the industry at a high level. The fact that most of the decrease in trade credit provision following the increase in barriers to entry in the French trucking sectors is due to dominant and cash-rich firms, although not definitive, is supportive of this hypothesis.

This paper contributes to several strands of the literature. It first relates to the literature on trade credit. There would be no need for trade credit in a frictionless world. More precisely, prices and payment terms would be perfect substitutes, so that a decrease in maximum payment terms would be fully offset by a decrease in prices and would have no real impact on firms corporate policy. However, in a world with imperfect information or moral hazard, trade credit and prices are not perfect substitutes. As discussed above, a first reason is the fact that some firms are credit rationed by external financiers, and that receiving trade credit from their supplier makes them better off. The second reason is that the supplier-client relationships are subject to agency and incentive problems which trade credit can help resolve. Kim and Shin (2012) show that trade credit can be used as a commitment device to mitigate moral hazard issues along the production chain. In Antràs and Foley (2011), late payment is a way for the client to incentivize the supplier to provide the right amount of effort in the production of the input. This paper is among the first to study the impact of trade credit provision on profitability and corporate policies. Closest to the results presented here is Murfin and Njoroge (2012) who use variation in large retailers' cash management policies to investigate the real impact of those cash management strategies on the suppliers of those firms. They show that firms are forced to cut back on investment in new plants and equipment when their buyers pay more slowly.

In addition, this paper relates to the literature on investment and liquidity constraints started by Fazzari et al. (1988) who find a strong positive relationship between cash flows and investment, suggesting that some informational frictions or agency problems constrain the amount of external financing that a firm may obtain. This result has been confirmed in studies of natural experiments such as lawsuits settlement and oil prices by Blanchard et al. (1994) and Lamont (1997), in a regression discontinuity design by Rauh (2006) and via a structural model by Whited (1992). Fazzari and Petersen (1993) consider the joint decision to invest in capex or working capital. They find that, if firms face financing constraints, working capital investment competes with fixed investment for the available pool of finance. I find very similar results: when trade credit is restricted, seemingly credit constrained firms invest more.

Finally, this paper adds to the growing literature on the interactions between product market competition and corporate finance. Amont the recent studies in this area, Leary and Roberts (2010) analyze how firms behavior is influenced by the decision of their peers, Hoberg and Phillips (2008) by the similarity of their competitors' products, Banerjee et al. (2008) by the relationships with their suppliers and customers.

The remainder of the paper is organized as follows. In section 1.2, I present the trade credit regulation reform which serves as the main source of identification in this paper. Section 1.3 presents the dataset. In section 1.4, I analyze the impact of the reform on trucking firms. In section 1.5, I analyze the behavior of transportation users. Section 1.6 presents two tests of the sensitivity of trade credit provision to competitive intensity. Section 1.7 concludes.

1.2 The 2006 trade credit regulation reform

The trade credit regulation reform which serves as the main source of identification in this paper came into effect in France in January 2006 with a law limiting payment terms to a maximum of 30 days for transactions involving a seller affiliated to one of six four-digit industries, listed in table 1.1. As appears in this table, 93 percent of the firms affected by the reform are trucking firms. I will refer to these firms as "treated firms" or "trucking firms" in what follows. The law stipulates that *both* suppliers and customers are eligible to a 75,000 euros fine should they breach the regulation. In case the statutory auditor of a firm observes a breach, she must give notice to the French Ministry of Finance, the equivalent to the U.S. Department of Commerce[6]. The reform had a huge impact on accounts receivables and working capital of trucking firms as evidenced in figure 1.1. Median receivables over sales drop by 4%, from approximately 80 to 65 days of outstanding sales between 2005 and 2007[7]

The regulation of payment terms has been on the European agenda for a number of years. On June 29, 2000, the European Commission adopted a directive (2000/35/EC) designed to fight late payment. This initiative was based on the belief that late payment terms were costly for businesses, especially small ones. The directive created a statutory right to interest

[6]The law also introduced an obligation to negotiate an updating process for prices of trucking contracts upon oil price changes. This obligation can however only be enforced if one of the parties brings it to the relevant court on its own initiative.

[7]Note that although this drop if very large, days of sales outstanding do not reach the theoretical 30 days set by the law. There are various reasons why this might be the case. Compliance with the law may be gradual. Moreover, sales outside France are not subject to the regulation.

payments after 30 days following the date of the invoice, unless another payment period was agreed upon in the contract. In July 2011, a new directive (2011/7/EU) was adopted to harmonize maximum payment terms at a maximum of 60 days, which Member States have to transpose in their respective national laws by March 16, 2013.

The reason why the reform targeted this subset of industries is somewhat related to the increased competition triggered by the enlargement of the EU to eight Eastern European countries in 2004, as well as the increase and volatility of oil prices, which in conjunction with unfavorable payment terms were thought to be excessively harmful to transport operators. It is likely that this legislation was lobbied for by trucking unions. However, it was extended to the remainder of the French economy in 2009, with a 45 days cap on payment terms applying to any transaction. This suggests that the regulation was addressing a problem not specific to trucking firms only.

The characteristics of the trucking industry have been thoroughly described in the work of Baker and Hubbard (2003, 2004) and Hubbard (2000, 2001, 2003) who show that the introduction of on board computers (OBCs) has two opposite effects on the structure of this industry: OBCs' incentive-improving capabilities leads to larger, more integrated firms, while OBCs' resource-allocation-improving capabilities leads to smaller, less integrated firms. Zingales (1998) finds that the initial size and leverage of trucking firms affected their survival following the deregulation of this industry in the U.S. in the early 80s.

The French trucking industry is arguably technologically similar to its U.S. counterpart. It is however less concentrated: while the U.S. trucking industry is constituted of large companies hiring many employee drivers, there are a more independent contractors owning and driving their own trucks in France (Arruñada et al., 2004).

Firms targeted by the trade credit regulation reform represent a decent share of the French economy. As of 2003, there were a total of 38,232 firms in this sector, employing 443,000 workers and generating combined sales of Euro 60bn and value added of Euro 1bn. They accounted for 46% of the sales and 40% of the value added of the total overall transportation industry (including water, rail and air transportation of goods and passengers). The main suppliers of trucking firms are oil producers and other transportation firms, which represent respectively 30 % and 22% of their input according to the 2005 input-output table of the French economy. Their main clients are other firms in the transportation industries and firms in the retail/wholesale sectors, which account for respectively 63% and 29% of their aggregate added value according to the 2005 input-ouput table. Not all firms in the economy resort to trucking firms for their road shipments. However, only a small share (about 15 percent) of road transportation is done internally as evidenced in table 1.3. Moreover, this proportion is stable through time, suggesting that the demand for road transportation services is pretty inelastic.

1.3 Data

Sample

I use accounting data extracted from the tax files used by the Ministry of Finance for corporate tax collection purposes, which is available up to 2007. This source has been used in various studies among which Bertrand et al. (2007). I obtain balance sheet and profit and loss accounts for the entire universe of French firms whose annual sales exceed 100,000 euros in the service sector and 200,000 in other sectors. Firms above these thresholds must fill in detailed balance sheet and profit statements on a yearly basis. Instead, smaller firms are subject to a simplified tax regime. Relative to the NSSBF, this data source has the advantage of being free of the misreporting concerns usually associated with survey-based data. Relative to Compustat, the dataset has the advantage of covering the entire universe of firms.

In the analysis of survival and bankruptcy, I use a file produced by the French statistical office which reports an exhaustive list of all bankruptcy filings each year in France, along with the unique identifying number of the corresponding bankrupt companies. To measure the probability of survival at creation, I need accounting information measured on the year of creation of new businesses. Since a large number of businesses do not reach 100,000 euros in sales in their first year, I use FICUS, another dataset constructed by the French statistical office. This dataset covers all incorporated firms in the economy, but provides less detailed information about their balance sheets and profit and loss accounts.

I track firms through time with the unique identifying number ascribed by the French statistical office. As emphasized in Bertrand et al. (2007) another feature of this data is that it includes, for each firm, a four-digit industry classification code that is very similar to the SIC coding system in the United States. This classification is prone to little measurement errors since codes are ascribed to each firm by the French statistical office itself.

To identify transport users, I take advantage of a survey conducted in 2005 by the French statistical office on a sample of 4,933 manufacturing firms with more than 5mln employees or total sales in excess of Euro 20mln. Surveyed firms are asked how much they spent in external purchases of a variety of services ranging from transportation to research and development in 2005. For the purpose of this study, the variable of interest is the 2005 amount spent in external transportation services.

Variables definition

The trade credit regulation reform of 2006 applies to trade credit contracts. Unfortunately, I do not have access to contract data. Hence, I proxy for the amount of trade credit given or received with the amount of accounts receivable and accounts payable on firms balance sheet as recorded at the end of their year of operations. I define "receivables" as accounts receivable (on the asset side of the balance sheet) minus cash received on orders (on the liability side) . I define "payables" as accounts payable (on the liability side) minus cash payed on orders (on the asset side).

While this might be a noisy measure of actual trade credit terms, I believe potential measurement errors are unlikely to drive the large observed change in receivables of treated (see figure 1.1). For measurement errors to bias the estimation of the impact of the reform on working capital upwards, it would have to be the case that firms in the treated group start giving more trade credit early and less trade credit towards the end of their year of operations. Suppose that prior to the reform, treated firms provided no trade credit in the first semester and large amounts of trade credit in the second semester of the year. Suppose that following the reform they adopted the opposite behavior. Then I might overestimate the impact of the reform. I doubt, however, that treated firms systematically behaved in such a way prior to the reform.

For virtually all analysis in this paper, I use a balanced panel, requiring that each firm included in the treatment and control groups reports information both on the year prior and following the reform. The only exception is the analysis of bankruptcies, where I also include firms which filed for bankruptcy in 2005 but did not report financial accounts in 2007, as well as firms created in 2006 and 2007, which did not report financial accounts in 2005.

Finally, the dependent variables of interest (payables, cash, credit lines, dividends, capex) are normalized by beginning of the year assets. ROA is measured as EBITDA over beginning of the year assets. All continuous variables are winzorised at the the 1% level.

Comparison with Compustat

One legitimate concern might be that firms in the sample are too different in terms of trade credit provision than firms in other samples used in the literature - such as Compustat. In table 1.2, I compare the 2005 level of accounts receivable and account receivables minus accounts payables of the 13,497 trucking firms in the sample with listed U.S. firms in the same sector for which accounting information is available in Compustat. As for the French trucking firms, the net trade credit position of U.S. listed firms is positive: they lend an average of 13.6 percent (net of the credit they receive from their suppliers) of their assets to their clients. The magnitude however differs, with French trucking firms lending twice as much as U.S. listed trucking firms.

1.4 Trade credit restriction: impact on trucking firms

Full-population difference-in-differences

Empirical strategy I analyze the response of trucking firms to the 2006 trade credit regulation in a difference-in-differences (DID) setting. Since the reform affects "relationships" between firms, the ideal empirical setting would exploit contract level information, as in Klapper et al. (2012). As detailed in section 1.3 however, I only observe the detailed balance sheets of the population of French firms. Firms affected by the reform are easy to identify using (four-digit) industry codes. I transpose the initial DID setting at the relation level into an approximate DID setting at the firm level.

In an ideal empirical setting, trucking firms would be randomized into a treatment and a control groups. Unfortunately, I have no way to construct such an ideal control group since the reform affected all trucking firms in the French economy. To build a relevant control group for the purpose of this study, I proceed as follow. I use the input-output tables of the French economy in 2005 (at the three-digit sector level), I compute the share of each industry's output in each of its downstream industries and then the distance to the trucking industry as the sum of the square difference of these shares. I consider only services industries (excluding financial services, hotels and restaurants), since the nature of the output has been shown to be a very important driver of the provision of trade credit in Burkart and Ellingsen (2004) and Giannetti et al. (2008). The "security, cleaning and other business services" industry has the smallest distance to the treated sector. It is not surprising that this industry turns out to be the closest to the trucking sector in terms of downstream markets since both sectors essentially include externalized business services in support of trade and production.

The treatment and controls groups might however differ along a number of dimensions, such as availability of external finance or investment opportunities. These differences could affect firms' corporate actions and explain part of the differential behavior observed following the reform. To mitigate this concern, I use a propensity score matching approach to construct an alternative control group. In the year prior to the reform (2005), I compute a propensity score for each firm using size (log assets), receivables as a fraction of sales and sales growth. Next, I match each treated firm to the control firm with the nearest predicted propensity score. The main advantage of this approach is to build an alternative control groups (the "matched control group") with firms that are as similar as possible to the treated firms. In particular, since I am interested in the response of trucking firms in terms of capex, it is important that both groups have comparable investment opportunities. Although there is no way to accurately measure the investment opportunities of private firms, sales growth has been used in the literature to roughly proxy for them (Petersen and Rajan, 1997; Asker et al., 2011).

A crucial assumption for the DID estimation to be valid is that the treatment and control groups should have parallel trends. In the absence of treatment, the observed DID estimates should be systematically zero. Because the reform occurs at the industry level, the parallel trends assumption might not hold if there are diverging latent trends. Although this assumption cannot be tested, I conduct a visual inspection of pre-reform trends in the variable of interest. Moreover, I conduct a placebo test by running the same DID estimates in the years prior to the 2006 reform. This placebo test also mitigates the concern that the reform might have been anticipated in any ways. Even after that, there remains some concerns that treated and control firms could differ along unobservable dimensions which may drive the results, in one direction or another.

With these concerns in mind, I turn to the main specification. I compare the behavior of firms one year prior to the reform (2005) and one year following the reform (2007). For each variable of interest, I estimate the difference between 2007 and 2005 in the treated and the control group and then take the difference of differences. I do so by using the following OLS regressions. I run this regression using alternatively the control group and the matched control group.

$$Y_{i,t} = \alpha_0 + \alpha_1.post + \alpha_2.treated_i + \alpha_3.post \times treated_i + \epsilon_{i,t}$$

where $Y_{i,t}$ is the outcome of interest measured in year t for firm i. *post* is a dummy equal to one in year 2007 and zero in year 2005, *treated* is a dummy indicating whether firm i belongs to the treatment or the control group. $\epsilon_{i,t}$ is an error term[8]. Standard errors are clustered at the firm level. In untabulated tests, I augment this specification with (i) lagged (two years) controls for size, tangibility (PPE over assets), receivables over sales and sales growth, and (ii) the same controls interacted with *post*. Results are similar or stronger in these complementary specifications.

Summary statistics Table 1.4 presents descriptive statistics that compare treated and control firms in 2005. The treated group includes 13,497 trucking firms. The control group includes firms from the security, cleaning and other business services industries (13,992 firms). The matched control group includes the nearest neighbor to each treated firm (13,497 firms) based on the propensity score matching procedure. T-tests present the t statistic of a test of the difference in means between the treated and the control and matched control groups. As can be observed, the treated and control firms are relatively similar on observables dimensions. The t-test of equality of means however rejects that they are equal, except for matching variables. The fact that the control and the matched control groups are close is reassuring with respect to the matching procedure.

Results Table 1.5 presents the effect of the trade credit regulation (2006) on a variety of corporate policies. Unsurprisingly, treated firms experience a very sharp drop of receivables (close to 8 percent of assets). Most of this effect is absorbed in cash holdings which increase

[8]As the law regulating trade credit also included a section facilitating the price indexation with respect to oil prices, I also include vector of controls P_t aiming at capturing this effect in unreported tests. P_t includes the (log) average annual price of diesel, as well as an indicator of its volatility computed as the log difference between the maximum and minimum monthly prices observed in the considered year. These two variables are interacted with the "treatment" dummy, and with the dummy variable indicating the post-reform period (*post*). It is unsure whether this part of the regulation has been actually enforced as results are very stable with or without this last set of variables. However, if anything, this insurance against the volatility of input prices should have a negative impact on *ex ante* prices.

by nearly 6 percent of assets. Credit lines decrease significantly, which is consistent with Petersen and Rajan (1997) who argue that lines of credit appear to be directly financing accounts receivables. With respect to control firms, capex in treated firms increases by 0.4%. ROA increases sharply by more than 3% between 2005 and 2007. Summing the effects on payables, credit lines, cash, dividend and capex, these results suggest that following the reform the average trucking firm dedicates 5 cents to increase capex, 9 cents to reduce credit line use and 14 cents to reduce account payables.

To insure that these results are not driven by unobserved trends in the control and treatment group, I conduct a placebo test by running the same experiment using the years 2003 and 2005 instead of 2005 and 2007. Results presented in table 1.6 show that if anything, the results obtained in the main specification are reversed or insignificant in the years prior to the reform: payables, cash and credit lines were increasing significantly prior to the reform, while there was no difference between the treatment and control group with respect to capex and ROA. I also make a visual inspection in figure 1.3 of trends in mean capex and ROA and find no diverging trend prior to the reform. Again, I cannot fully rule out that there may be some unobserved heterogeneity between treated and control firms which might drive the results in one direction or another.

I then check the extent to which these effects vary with the intensity of credit constraints. I split firms based on size, payout status and age, which have been used as proxies for financial constraints in Almeida et al. (2004) and Hadlock and Pierce (2010). Table 1.7 presents the DID estimates of the effect of the trade credit regulation (2006) on corporate policy variables between 2005 and 2007, conditional on firm size, payout status and age. Large firms are firms in the top half of the distribution of assets in 2005, Payout firms are firms that distributed dividends in 2005 and Old firms are firms in the top half of the distribution of age. An important caveat to this analysis is that unconstrained firms are likely to be the ones extending more trade credit prior to the reform, and therefore to be the most affected by the reform. Indeed, receivables and cash holdings decrease more for unconstrained firms.

Payables decrease more for unconstrained firms, consistent with the idea that constrained firms might still be willing to rely on their own suppliers to obtain finance. Although the difference is not significant, most of the increase in capex seems to occur in small, non dividend paying and young firms, while no effect is found on large, dividend paying and old firms. Finally, ROA seems to increase at all firms.

I then turn to the impact of the reform on survival. This analysis is conducted on the full sample of treated and control firms. Table 1.8 presents the results of a firm level DID estimation of the impact of the 2006 trade credit regulation reform on survival. Bankruptcy is a dummy equal to one on the year of the firm's bankruptcy, and zero in other years (until the firm stops filing tax forms). *Post* is a dummy equal to one in year 2007 and zero in 2005. Specifications in column 2 and 5 include firm level controls measured with a two year lag: log of assets, PPE over assets, receivables over sales, sales growth and debt over assets. Columns 3 and 6 include the same controls interacted with *Post*. Coefficients on control variables have the expected sign and are not reported for the clarity of exposition. Standard errors are corrected for clustering at the four digit sector level. As it turns out, treated firm experience a sharp and significant decrease of a 1 percent in their default probability, which amounts to a drop of the order of 50 percent with respect to their baseline default probability prior to the reform.

Cohort analysis

In a separate yet related test, I analyze the survival of newly created firms in the treatment and the control groups defined above around the 2006 reform. The previous DID setting is replicated at the cohort level using the following OLS specification:

$$B_{i,t_0+h} = \alpha + \sum_{t=2004}^{2007}(\beta_t.Cohort_{i,t}) + \sum_{t=2004}^{2007}(\gamma_t.treated_i.Cohort_{i,t}) + \delta.treated_i + \lambda.X_{i,t_0} + \epsilon_{i,t_0}$$

B_{i,t_0+h} denotes the outcome of interest, a dummy equal to 1 if firm i files for bankruptcy within h years following the year of creation t_0, and X_{i,t_0} denotes a set of controls measured in the year of creation, including the log number of employees, log assets per employee, log sales per employee, log debt per employee, ROA and log intangible assets per employee. The reference cohort is 2003. These regressions compare the 2004 cohort, which experienced the reform only two years after creation (so that, controlling for characteristics at creation, the probability of survival until 2 years was unaffected by the reform), with cohorts created in 2006 and 2007, which only experienced the new regime and its associated probabilities of default. The 2005 cohort has an intermediate status, as those firms experienced the change during their first (to second) year of existence.

Table 1.9 presents the regression of the default probability of firms in the second, third and fourth year following their creation. The reform has no effect on the survival of the treated firms in the 2004 cohort at a two year horizon but it has a an impact on their survival at the 3 year horizon. The treated firms of the 2006 cohort have the highest rate of survival at all horizons. It appears clearly that the reform had a strong impact on the survival probability of new trucking businesses.

Surprisingly, the coefficient for treated firms in the 2007 cohort is lower than the coefficient for the 2006 cohort, and not significantly different from zero. I interpret this as the result of adverse selection. As the reform somewhat decreased the cost in terms of payment terms to enter this industry, it might have drawn some lower quality entrepreneurs. As evidenced in figure 1.4, the treated sectors experienced a large number of entries in 2007. This flow of entrants probably included new businesses of lower quality in the treated sectors, which would explain the lower rate of survival of the 2007 cohort with respect to the 2006 cohort.

How should these results be interpreted so far? If suppliers have an advantage over other financial intermediaries in lending to their clients, then restricting their ability to lend should hurt them. However, all treated firms seem to be immediately better off following the reform, as measured by their ROA or default probability. For the results to be consistent with this hypothesis, it should be the case that lending to their clients is costly in the short run and profitable in the long run, which might be the case if they are able to extract a share in their clients future surpluses.

Before checking whether trucking firms lend more when the value of client relationships is likely to be higher, I briefly consider the impact of the reform on transport users.

1.5 Trade credit restriction: impact on transport users

To understand the behavior of trucking firms, it is useful to consider what happened to their clients around the trade credit reform. In particular, it is useful to identify which type of industries used to receive long payment terms from their trucking suppliers prior to the reform. Although I do not have information at the contract level, I take advantage of a survey described in section 1.3 which reports the total amount of external transportation services purchased by each firm in a sample of 4,933 manufacturing firms in 2005. I compute the intensity of external transportation as the ratio of external transportation services to total purchases in 2005. I then label the top two quintiles of the distribution as High Transport Users (HTU) and the bottom two quintiles as Low Transport Users (LTU). HTU have a mean intensity of external transportation use of 8.2 percent while LTU have a mean intensity of external transportation of less than 1 percent. The dynamics of the ratio of payables to total purchases is presented in figure 1.5. As expected, HTU experience a median drop of 1 percent in payables over total purchases with respect to LTU following the reform.

I then compare the change in payables over total purchases, for HTU and LTU in various industries one year before (2005) and one year following (2007) the reform. The idea of the test goes as follows: suppose that payables decrease more for HTU in growing industries than in non concentrated industries, relative to LTU. I can then infer that trucking companies where lending more to firms in growing industries prior to the reform.

This test rests on several assumptions. First, I do not observe the exact amount spent in external *road* transportation services. Instead, the amount of external transportation services reported in the data also includes water and air transportation. Since, as evidenced in table 1.3, trucking is the dominant transportation mode, I believe that using the share of external transportation services is reasonable proxy for the share of external trucking services. If anything, the measurement error should bias the results against finding any effect. A more serious concern might arise if transport users change their transportation mix after 2005. The change in payables that I observe would then be much less informative about the level of trade credit extended by trucking firms up to 2005. Although I can not test this, table 1.3 suggests that external road transportation demand is very inelastic through time. Third, within a given industry, HTU and LTU should not experience a different change in payables, with the exception of payables related to transportation services. Again, while I have no way to check whether this is the case, I believe this assumption is reasonable: according to figure 1.5, the median payables of HTU and LTU follow parallel trends up until 2005. As a robustness check, I do a back-of-the-envelop calculation of the drop in payables over total purchases which we should expect given the median intensity of external transportation of HTU (6.4 percent), their 2005 level of payables over total purchases (23 percent) and the median drop in trucking firms trade credit provision (19 percent, see figure 1.1). I find an expected drop of 1.2 percent of payables over total purchases, which is very close to the drop observed in figure 1.5.

I then run the following OLS firm-level regression:

$$\Delta.Pay_i = \alpha + \beta.HTU_i.X_j + \gamma.HTU_i.(1 - X_j) + \delta.X_j + \lambda_j + \epsilon_{i,j}$$

$\Delta.Pay_i$ is the change (difference) in payables over total purchases between 2005 and 2007 for firm i. HTU_i is a dummy equal to one if firm i is a High Transport User or zero otherwise. X_j is a dummy equal to one if industry j of firm i is of type X where $X \in (concentrated, profitable, growing, exporting)$. Concentrated (respectively: profitable, growing, exporting) industries are industries which Herfindahl index based on value added[9]

[9]The reason why I use value added rather than sales is that if there is contracting going on within a

(respectively: aggregate net income over aggregate sales, growth in aggregate value added and share of exports) lies above the 2005 sample median. λ_j is a sector fixed effect. Standard errors are clustered at the four-digit sector level.

Table 1.10 presents some summary statistics for HTU and LTU. HTU have an average ratio of external transportation purchases of 8.2 percent, vs. 0.9 percent for LTU. As it turns out, both set of firms are relatively similar in 2005 on observable dimensions such as size, tangibility or sales growth. Unsurprisingly given the scope of the survey, these firms are much larger on average than trucking companies. The results presented in table 1.11 indicate that High Transport Users experience a sharper decrease in payables over total purchases than Low Transport Users especially in concentrated, profitable and growing industries.

1.6 Competition, relationships and trade credit

If firms forgo current profitability to extend credit to their constrained clients and share their future surpluses, they should do so even more when the value of client relationships is higher. Although I have no way of measuring the value of such relationships, I hypothesize that it should be higher to the supplier when there is less competitive pressure. A lower level of competition implies that switching costs are higher for the client, and that relationships are more likely to be preserved. I use cross-sectional variation and time-series variations in the intensity of competition in the trucking sector to test this idea.

1.6.1 Local competition: impact on trade credit provision

I check whether trucking firms extend more credit to their customers in areas where there is indeed less competition. In columns 1 to column 3 of table 1.12, I replicate table 3 of Petersen and Rajan (1997) on the cross-section of 13,497 trucking firms in 2005. The dependent variable is receivables over sales, which I regress on the log of assets, the log of

sector, adding up sales results in double count. I check that results are similar when using a Herfindahl index based on 2005 sales.

age and a square term, the ratio of credit lines over sales, net profit (net income) over sales, a variable equal to sales growth if it's positive and zero otherwise, a variable equal to sales growth if it's negative and zero otherwise, gross margin (EBITDA margin) and a square term, and four-digit sector fixed effects. Standard errors are corrected for heteroskedasticity. I find very similar results as Petersen and Rajan (1997). Larger firms extend more trade credit. Older firms also extend more credit to their customers, although the relationship is non linear, and although the coefficient is non significant in my sample. Both firms with positive growth and negative growth have larger account receivables. Firms making losses extend more credit. Finally, the larger a firm's gross profit margin, the larger is the amount of receivables. This analysis suggests that the drivers of trade credit provision in the cross-section of firms are similar in my sample and in the NSSBF conducted in 1988-1989, which is reassuring in itself.

In columns 4 and 5, I add a variable measuring the intensity of local competition. As a first path, I identify the local market of the firm as the French "département"[10] (henceforth "district") where the trucking firm's headquarters are located. To measure the intensity of competitive pressure, I use alternatively the number of firms and the Herfindahl index of concentration computed at the four-digit sector × district level, based on value added in 2005[11]. In column 4, I add to the previous specification the log number of firms in the four-digit sector and district of the firm, as well as the log of the total number of firms in the district. In column 5, I add the Herfindahl index of concentration in the four-digit sector and district of the firm, as well as the median Herfindahl index of concentration across all sectors in the district of the firm. As it turns out, a higher intensity of competition measured by a larger number of firms or a lower Herfindahl index are associated with a larger amount of outstanding trade credit, controlling for all other determinants of trade credit provision.

This result is suggestive that trucking firm do not seem to extend more credit where the

[10]There are 101 French "département", with an average population of roughly 600,000 people

[11]The reason why I use value added rather than sales is that if there is contracting going on within a sector, adding up sales results in double count. I check that results are similar when using a Herfindahl index based on 2005 sales.

value of relationships are likely to be the highest. However, the result could be biased if an omitted variable drove both trade credit provision and concentration of trucking firms at the local level. Moreover, I cannot fully reject that reverse causality might be driving the effect, in the event that trucking firms concentrate more in areas where more trade credit is less needed from their clients.

1.6.2 Barriers to entry: impact on trade credit provision

To mitigate these concerns, I use an alternative source of variation in competition. I consider an exogenous legal change to barriers to entry to the trucking sector announced in 1999. Entry conditions were increased by a legislation which made it compulsory for firm owners to (i) hold a specific degree, (ii) be free of any past criminal record, (iii) conduct business in an incorporated firm listed on the "national transportation registry" and (iv) put forward minimum capital requirements related to the number of truck used, be them owned or rent: 1,800 (9,000) euros for the first truck carrying less (more) than 3.5 tons and 900 (5,000) for any additional one[12]. This reform was announced in September 1999 and came into effect in December 1999[13]. It had a significant impact on the median book equity at creation and consequently on entry in the trucking sector, as evidenced in figure 1.6.

This reform offers a nice setting to test how firms alter their trade credit provision when their competitive environment changes. To do so, I define a control group and a matched control group as in section 1.4. The trucking industry is the treated group (11,478 firms). The control group includes firms from the security, cleaning and other business services industries (11,761 firms). The matched control group includes the nearest neighbor to each treated firm (11,478 firms) based on a propensity score matching procedure based on size (log of assets), tangibility (PPE over assets) and sales growth, all measured in 1999. Matching

[12] Another reform was implemented in 1998 and gave the right to a trucking contractor to claim payment by the shipper if the outsourcing/contracting firm defaults. From an accounting point of view, this is likely to increase the value of receivables, since provisions for doubtful receivables will decrease. Moreover, this reform should drive trucking firms to increase receivables, since they face a lower risk of losses on their trade loans

[13] The actual administrative declaration file was published in December 2009

firms on size and tangibility ensures that the test compares firms with comparable access to outside financing. Using sales growth makes sure that firms in the matched control group have comparable investment opportunities - although, as discussed above, sales growth is a relatively rough proxy for investment opportunities.

For each firm, I compute the change (difference) in receivables over sales between 1999 and 2001. Moreover, I check whether, within the treatment group, dominant and cash-rich firms decrease change trade credit provision by more or less. I consider a firm as large (cash-rich) when its 1999 sales (cash over assets ratio) lie above the median of its four-digit sector in 1999. I then run the following firm-level OLS regression:

$$\Delta.Rec_i = \alpha + \beta.treated_i.C_i + \gamma.treated_i.(1 - C_i) + \delta.C_i + \epsilon_{i,j}$$

$\Delta.Rec_i$ is the change (difference) in receivables over sales for firm i between 1999 and 2001. C_i is a dummy equal to one if firm i is above the median of its four-digit sector in terms of total sales or the ratio of cash to total assets as of 1999. Standard errors are clustered at the firm level.

Table 1.13 presents descriptive statistics that compare treated and control firms in 1999. T-test present the t statistic of a test of the difference in means between the treatment and the control and matched control groups. Treated and control firms are relatively similar on observables dimensions. The t-test of equality of means however rejects that they are equal, except for sales growth and tangibility. The fact that the control and the matched control groups are close is reassuring with respect to the matching procedure. Table 1.14 presents the estimation of the effect of the change in barriers to entry to the trucking industry on the change (difference) in receivables (over sales). The results indicate that receivables of treated firms decreased by a significant one percent following the increase in barriers to entry, which amounts to a four percent decrease with respect to the average of 1999.

This is evidence that, if anything, trucking firms do not seem to lend more to their clients when the value of lending relationships supposedly increases following the change in barriers to entry in the industry. The fact that most of this increase is due to dominant and cash-rich firms is suggestive that some firms might have set receivables at a higher level prior to the reform, in an attempt to limit entry by new businesses. Of course, these results are subject to the same caveat as those obtained from the main DID specification of section 1.4. In particular, it could be that unobserved heterogeneity between control and treated firms drive the results in one direction or another.

1.7 Discussion and conclusion

I consider the response of French trucking firms to an exogenous restriction on their ability to extend payment terms to their clients beyond 30 days. I find that for one dollar less of receivables, the average trucking firm dedicates 5 cents to increase capex, 9 cents to reduce credit line use and 14 cents to reduce account payables. Moreover, ROA increases sharply, by more than 3 percent. While the effect on ROA seems to be prevalent across the board, the effect on capex is stronger among seemingly credit constrained firms. Finally, the default probability of the average trucking firm drops sharply by a significant 50 percent, and the survival probability of newly created trucking firms increases significantly.

A possible explanation for these somewhat surprising results is that trucking firms might invest in client relationships. They might extend long payment terms to their constrained clients in the hope of sharing future surpluses. However, I find that firms provide more credit when they face more competition, and hence when the value of relationships is likely to be lower.

Altogether, the results point to a somewhat darker side of trade credit, whereby some firms extend long payment terms to their clients at the cost of lower profitability. This could be explained if there are private benefits of owning a trucking firm and if these benefits offset the reduced profitability associated with extending more or cheaper credit to clients. The fact that ROA decreases for both small and large firms suggests that this might not be the only explanation. An alternative yet relatively unexplored hypothesis might be that if the long term costs of excess entry in the trucking sector outweigh the costs of additional or cheaper credit provision, then incumbent firms may prevent the entry or expansion of new businesses by extending cheap credit to constrained clients and setting standard payment terms in the industry at a high level. The fact that most of the decrease in trade credit provision following the increase in barriers to entry in the French trucking sectors is due to dominant and cash-rich firms is somewhat supportive of this hypothesis.

The evidence presented in this paper may be at play in other sectors and in other countries. They are likely to be stronger in countries with less developed financial markets, where clients may be more credit constrained and more reliant on their suppliers trade credit provision. I believe that the dynamics and welfare implications of inter-firm lending are important questions which deserve further research.

Bibliography

Ahn, JaeBin, Mary Amiti, and David E. Weinstein, "Trade Finance and the Great Trade Collapse," *American Economic Review*, May 2011, *101* (3), 298–302.

Almeida, Heitor, Murillo Campello, and Michael S Weisbach, "The Cash Flow Sensitivity of Cash," *The Journal of Finance*, 2004, *59* (4), 1777–1804.

Antràs, Pol and C. Fritz Foley, "Poultry in Motion: A Study of International Trade Finance Practices," *NBER working paper*, May 2011, (17091).

Arruñada, Benito, Manuel González-Díaz, and Alberto Fernández, "Determinants of organizational form: transaction costs and institutions in the European trucking industry," *Industrial and Corporate Change*, 2004, *13* (6), 867–882.

Asker, John, Joan Farre-Mensa, and Alexander Ljungqvist, "Comparing the Investment Behavior of Public and Private Firms," *NBER working paper*, September 2011, (17394).

Baker, George P and Thomas N Hubbard, "Make Versus Buy in Trucking:Asset Ownership, Job Design, and Information," *American Economic Review*, 2003, *93* (3), 551–572.

_ and _ , "Contractibility And Asset Ownership: On-Board Computers and Governance In US Trucking," *Quarterly Journal of Economics*, 2004, *119* (4), 1443–1479.

Banerjee, S, S Dasgupta, and Y Kim, "Buyer-supplier relationships and the stakeholder theory of capital structure.," *The Journal of Finance*, 2008, *63* (5), 2507–2552.

Bertrand, Marianne, Antoinette Schoar, and David Thesmar, "Banking Deregulation and Industry Structure: Evidence from the French Banking Reforms of 1985," *Journal of Finance*, 04 2007, *62* (2), 597–628.

Biais, Bruno and Christian Gollier, "Trade Credit and Credit Rationing," *Review of Financial Studies*, 1997, *10* (4), 903–37.

Blanchard, Olivier Jean, Florencio Lopez de Silanes, and Andrei Shleifer, "What do firms do with cash windfalls?," *Journal of Financial Economics*, 1994, *36* (3), 337–360.

Burkart, Mike and Tore Ellingsen, "In-Kind Finance: A Theory of Trade Credit," *American Economic Review*, 2004, *94* (3), 569–590.

Caves, R E and M E Porter, "From Entry Barriers to Mobility Barriers: Conjectural Decisions and Contrived Deterrence to New Competition," *The Quarterly Journal of Economics*, 1977, *91* (2), 241–262.

Cunat, Vicente, "Trade credit: suppliers as debt collectors and insurance providers," *Review of Financial Studies*, 2006, *20* (2), 491–527.

Dass, Nishant, Jayant Kale, and Vikram Nanda, "Trade Credit, Relationship-specific Investment, and Product-market Power," *Working paper*, 2011.

Fabbri, Daniela and Anna Maria C Menichini, "Trade credit, collateral liquidation, and borrowing constraints," *Journal of Financial Economics*, 2010, *96* (3), 413–432.

_ **and Leora Klapper**, "Market power and the matching of trade credit terms," *World Bank working paper*, October 2008, (4754).

Fazzari, Steven M and Bruce C Petersen, "Working Capital and Fixed Investment: New Evidence on Financing Constraints," *The RAND Journal of Economics*, 1993, *24* (3), 328–342.

_ **, R Glenn Hubbard, Bruce C Petersen, Alan S Blinder, and James M Poterba**, "Financing Constraints and Corporate Investment," *Brookings Papers on Economic Activity*, 1988, *1988* (1), 141–206.

Fisman, Raymond and Mayank Raturi, "Does Competition Encourage Credit Provision? Evidence from African Trade Credit Relationships," *The Review of Economics and Statistics*, 2004, *86* (1), 345–352.

Fudenberg, Drew and Jean Tirole, "Capital as a Commitment: Strategic Investment to Deter Mobility," *Journal of Economic Theory*, 1983, *31* (2), 227–250.

Giannetti, M, M Burkart, and T Ellingsen, "What You Sell Is What You Lend? Explaining Trade Credit Contracts," *Review of Financial Studies*, 2008, *24* (4), 1261–1298.

Hadlock, Charles J. and Joshua R. Pierce, "New Evidence on Measuring Financial Constraints: Moving Beyond the KZ Index," *Review of Financial Studies*, 2010, *23* (5), 1909–1940.

Hoberg, Gerard and Gordon M. Phillips, "Real and Financial Industry Booms and Busts," *NBER working paper*, August 2008, (14290).

Hubbard, Thomas N, "The Demand for Monitoring Technologies: The Case of Trucking," *Quarterly Journal of Economics*, 2000, *115* (2), 533–560.

_ , "Contractual Form and Market Thickness in Trucking," *The RAND Journal of Economics*, 2001, *32* (2), 369.

_ , "Information, Decisions, and Productivity: On-Board Computers and Capacity Utilization in Trucking," *American Economic Review*, September 2003, *93* (4), 1328–1353.

Hurst, Erik and Benjamin Wild Pugsley, "What Do Small Businesses Do?," *NBER working paper*, May 2011, (17041).

Kim, Se-Jik and Hyun Song Shin, "Sustaining Production Chains through Financial Linkages," *American Economic Review*, May 2012, *102* (3), 402–06.

Klapper, Leora, Luc Laeven, and Raghuram Rajan, "Trade Credit Contracts," *Review of Financial Studies*, 2012, *25* (3), 838–867.

Lamont, Owen, "Cash Flow and Investment: Evidence from Internal Capital Markets," *The Journal of Finance*, 1997, *52* (1), 83–109.

Leary, Mark T and Michael R Roberts, “Do Peer Firms Affect Corporate Financial Policy?,” *working paper*, 2010, (215).

Murfin, Justin and Kenneth Njoroge, “Small lending big: The real effect of trade credit demands on constrained suppliers,” *Working paper*, 2012.

Petersen, Mitchell A and Raghuram G Rajan, “The Effect of Credit Market Competition on Lending Relationships,” *The Quarterly Journal of Economics*, May 1995, *110* (2), 407–43.

_ **and** _ , “Trade credit: theories and evidence,” *Review of Financial Studies*, 1997, *10* (3), 661–691.

Rauh, Joshua D, “Investment and Financing Constraints: Evidence from the Funding of Corporate Pension Plans,” *The Journal of Finance*, 2006, *61* (1), 33–71.

Robb, Alicia and EJ Reedy, “An overview of the Kauffman Firm Survey: Results from 2010 Business Activities,” *Kauffman Foundation*, May 2012.

Whited, Toni M, “Debt, Liquidity Constraints, and Corporate Investment: Evidence from Panel Data,” *Journal of Finance*, 1992, *47* (4), 1425–1460.

Wilner, Benjamin S, “The Exploitation of Relationships in Financial Distress: The Case of Trade Credit,” *The Journal of Finance*, 2000, *55* (1), 153–178.

Zingales, Luigi, “Survival of the Fittest or the Fattest? Exit and Financing in the Trucking Industry,” *The Journal of Finance*, 1998, *53* (3), 905–938.

Appendix

Content of the 2006 law

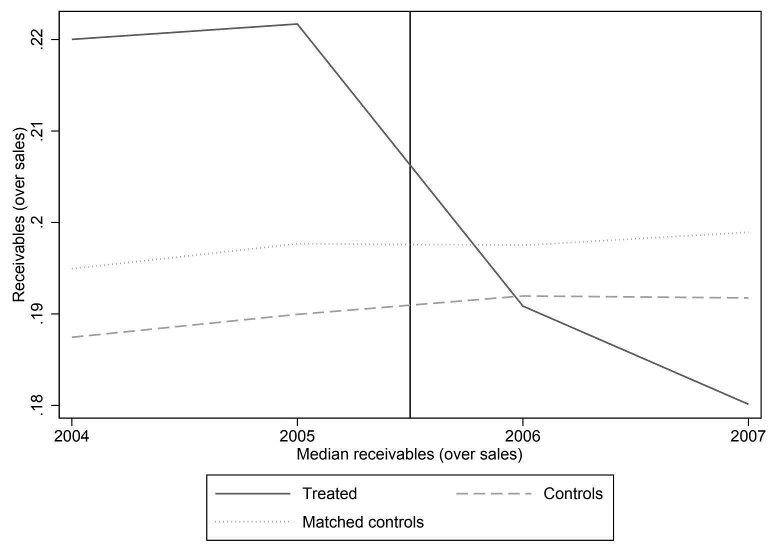
The following was introduced in the French "Code du Commerce": *Contrairement aux dispositions de l'alinéa précédent, pour le transport routier de marchandises, pour la location de véhicules avec ou sans conducteur, pour la commission de transport ainsi que pour les activités de transitaire, d'agent maritime et de fret aérien, de courtier de fret et de commissionnaire en douane, les délais de paiement convenus ne peuvent en aucun cas dépasser trente jours à compter de la date d'émission de la facture. Les conditions de règlement doivent obligatoirement préciser les conditions d'application et le taux d'intérêt des pénalités de retard exigibles le jour suivant la date de règlement figurant sur la facture dans le cas où les sommes dues sont réglées après cette date. Sauf disposition contraire qui ne peut toutefois fixer un taux inférieur à une fois et demie le taux d'intérêt légal, ce taux est égal au taux d'intérêt appliqué par la Banque centrale européenne à son opération de refinancement la plus récente majoré de 7 points de pourcentage. Les pénalités de retard sont exigibles sans qu'un rappel soit nécessaire. La communication prévue au premier alinéa s'effectue par tout moyen conforme aux usages de la profession. Toute infraction aux dispositions visées ci-dessus est punie d'une amende de 15000 euros. Les personnes morales peuvent être déclarées responsables pénalement, dans les conditions prévues par l'article 121-2 du code pénal. La peine encourue par les personnes morales est l'amende, suivant les modalités prévues par l'article 131-38 dudit code.*

Description of treated groups (in French, source INSEE)

- 60.2L: road transport in urban areas or close to shippers, which consists in removing or deliver goods packed or not during short trips, delivering concrete ready; collection of farm milk
- 60.2M: road transport of goods on long distance and international long-distance, including heavy transport in bulk containers
- 60.2P: rental of trucks and vans with driver
- 63.4A: collection of multiple shipments of less than 3 tons grouped on piers to load trucks for unbundling and delivery to the recipient's address; express freight
- 63.4B: land, sea and air chartering (or a combination thereof) which consists in handing in items without prior grouping to public carriers
- 63.4C: logistical organization of freight transport from or out of France or internationally by all modes of transport appropriate; land transit by sea or air; activities of customs agents

Figure 1.1: Impact of the trade credit reform (2006) on receivables and working capital
These figures show the impact of the trade credit regulation reform (2006) on median receivables over sales (top panel) and working capital over sales (bottom panel). The trucking industry is the treated group (13,497 firms). The control group includes firms from the security, cleaning and other business services industries (13,992 firms). The matched control group includes the nearest neighbor to each treated firm based on a propensity score matching procedure described in section 1.4 (13,497 firms).

A. Receivables over sales



B. Working capital over sales

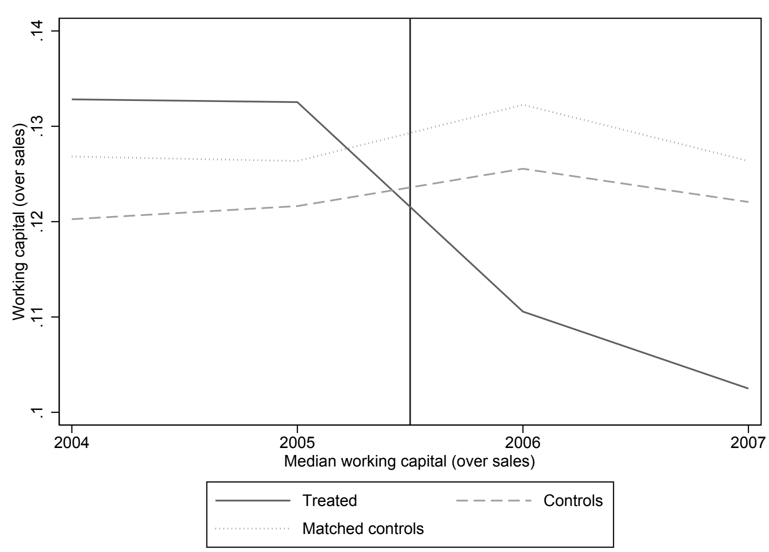
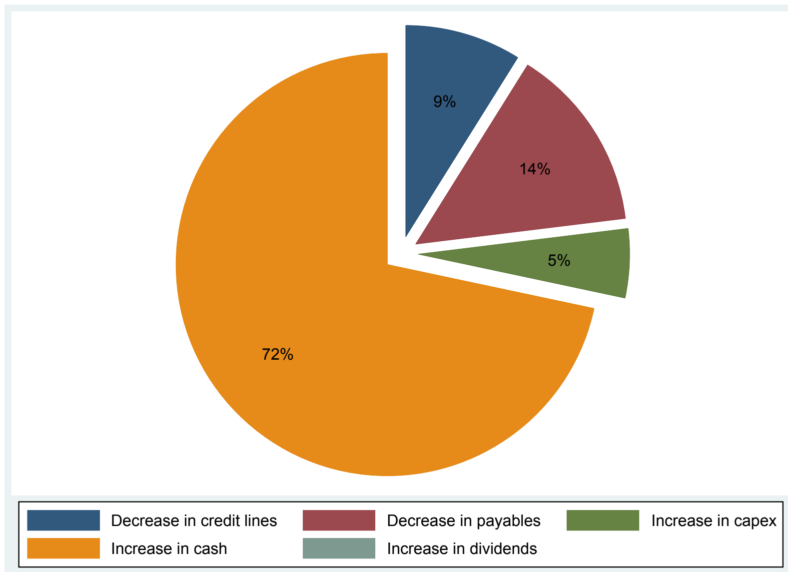


Figure 1.2: Change in policy variables following the trade credit reform (2006)

These figures shows the average response of policy variables to the drop in receivables induced by the trade credit regulation reform (2006). The trucking industry is the treated group (13,497 firms). The control group includes firms from the security, cleaning and other business services industries (13,992 firms). The matched control group includes the nearest neighbor to each treated firm based on a propensity score matching procedure described in section 1.4 (13,497 firms).

A. Treated vs. controls



B. Treated vs. matched controls

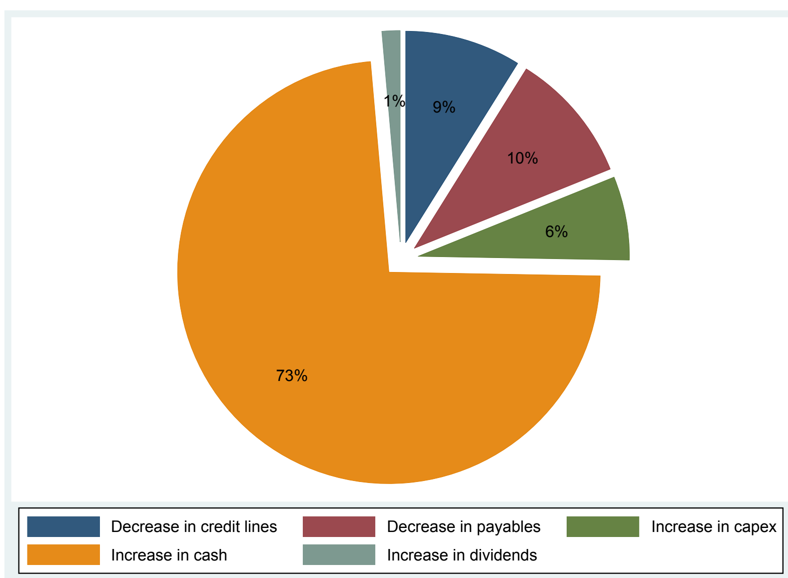
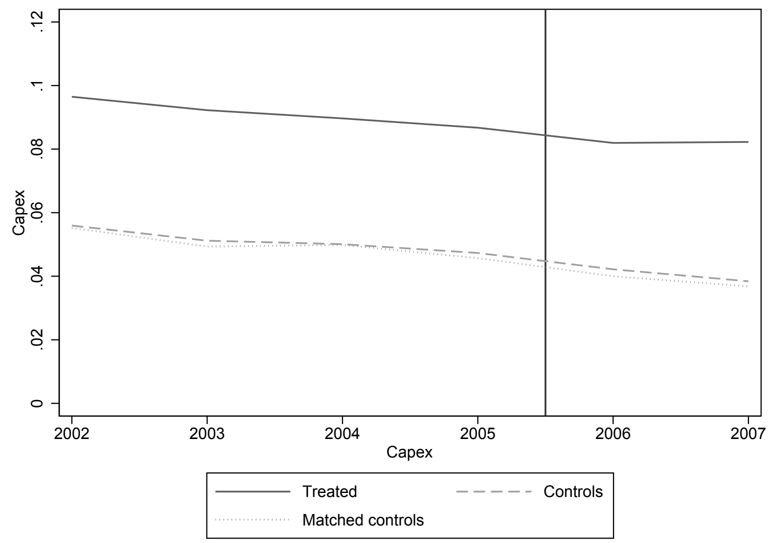


Figure 1.3: Parallel trends before the trade credit reform (2006)

These figures shows the trends in capex and ROA prior to the trade credit regulation reform (2006). The trucking industry is the treated group (13,497 firms). The control group includes firms from the security, cleaning and other business services industries (13,992 firms). The matched control group includes the nearest neighbor to each treated firm based on a propensity score matching procedure described in section 1.4 (13,497 firms).

A. Capex over assets (t-1)



B. ROA

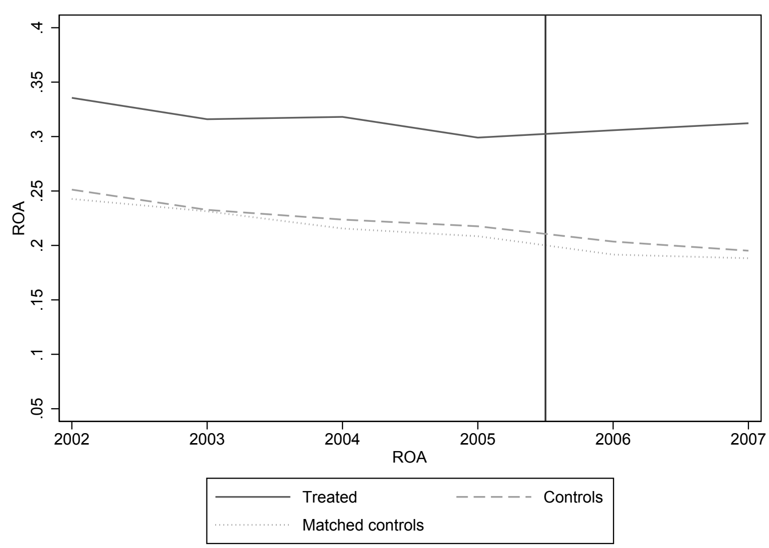


Figure 1.4: Entry rate around the trade credit reform (2006)

This figure shows the rate of entry, defined as the ratio of new business creations to existing businesses, for the trucking industry (the treated group), security, cleaning and other business services industry (the control group) and for the whole economy excluding the treated group.

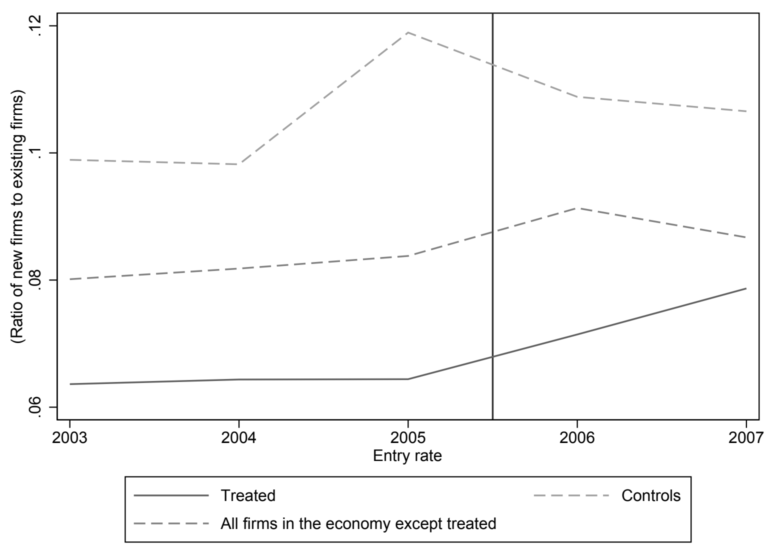
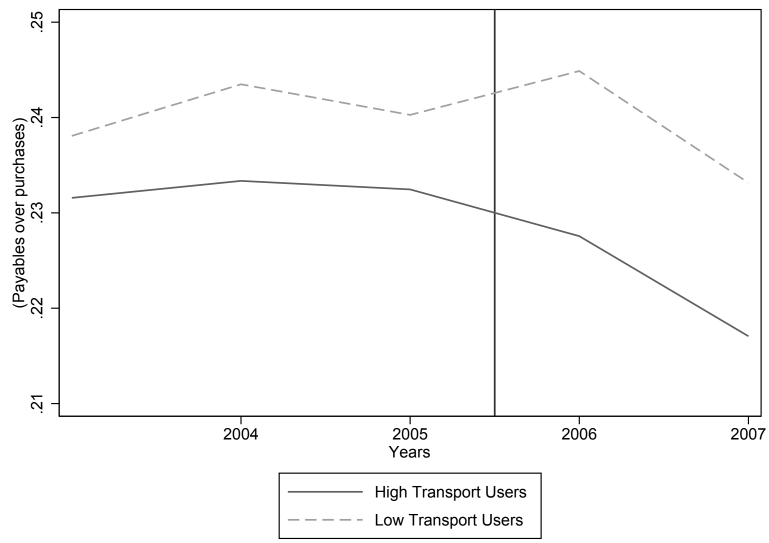


Figure 1.5: Change in payables of transport users around the trade credit reform (2006)

These figures show the change payables over total purchases around the trade credit regulation reform (2006) for high and low transportation users in a sample of 4,933 manufacturing firms. Firms are allocated to the High (Low) Transport Users group if their ratio of external purchases of transport services to total purchases in 2005 are in the top (bottom) 2 quintiles of the distribution.

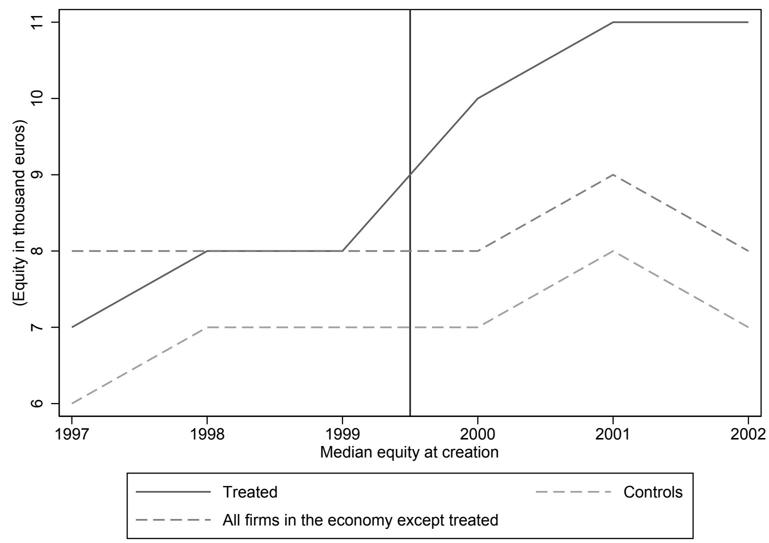
A.Median payables over total purchases, high vs. low transport users



B.Difference in payables over total purchases, high vs. low transport users

Figure 1.6: Start-up capital and rate of entry around the change in barriers to entry (1999)
These figures show the impact of a change in minimum start-up capital in the trucking industry in 1999. The top panel presents the median start up capital measured at creation for the trucking industry (the treated group), security, cleaning and other business services industry (the control group) and for the whole economy excluding the treated group. The bottom panel presents the rate of entry, measured as the as the ratio of new business creations to existing businesses.

A.Median start up capital



B. Rate of entry

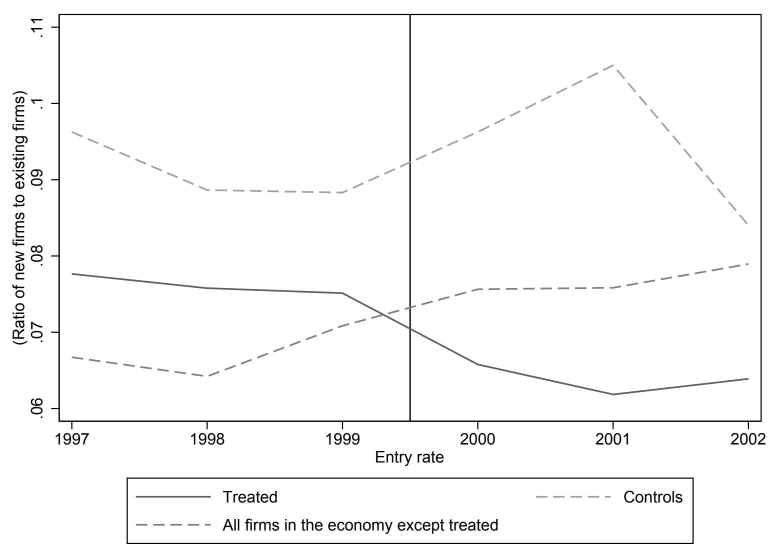


Table 1.1: Scope of the trade credit reform (2006)

Three-digit sector	Four-digit sector	Four-digit sector code	Number of firms in 2005 (in the main sample)
Freight transport by road	Proximity freight transport by road	60.2L	4,882
	Interurban freight transport by road	60.2M	6,228
	Rent of lorries with driver	60.2P	518
Transportation support activities	Freight services organization	63.4A	427
	International transportation organization	63.4C	1,113
	Chartering	63.4B	329

Table 1.2: Trucking industry (2005): Compustat vs. Treated firms

This table compares receivables and receivables minus payables over sales and over assets for the French trucking firms in the treated group and for equivalent firms in Compustat.

	Compustat				Treated firms (main sample)			
	Obs.	Mean	Median	Std. dev.	Obs.	Mean	Median	Std. dev.
Receivables over assets	63	0.276	0.229	0.203	13497	0.523	0.525	0.218
Receivables over sales	58	0.129	0.128	0.043	13497	0.252	0.227	0.195
Receivables - Payables over assets	63	0.136	0.134	0.159	13497	0.293	0.294	0.226
Receivables - Payables over sales	58	0.051	0.076	0.136	13497	0.144	0.138	0.126

Table 1.3: Share of road transportation in total transportation on the French territory
This table presents the share of road transportation in total transportation on the French territory, measured in km-tons as disclosed by the French Ministry of transportation. Internal transportation describes transportation activity by shippers using internal means.

	Share of million km-tons by transport mode				
	Year	Railroads	Waterways	Road (external)	Road (internal)
Domestic					
	2003	0.14	0.02	0.69	0.16
	2004	0.13	0.02	0.71	0.14
	2005	0.11	0.02	0.72	0.15
	2006	0.11	0.02	0.71	0.15
	2007	0.11	0.02	0.73	0.14
	2008	0.12	0.02	0.72	0.14
International (Exports)					
	2003	0.41	0.08	0.48	0.02
	2004	0.40	0.09	0.50	0.02
	2005	0.40	0.11	0.47	0.02
	2006	0.40	0.11	0.47	0.02
	2007	0.41	0.11	0.46	0.02
	2008	0.42	0.11	0.44	0.03
International (Imports)					
	2003	0.37	0.09	0.52	0.02
	2004	0.40	0.09	0.50	0.02
	2005	0.40	0.11	0.47	0.02
	2006	0.40	0.11	0.47	0.02
	2007	0.41	0.11	0.46	0.02
	2008	0.42	0.11	0.44	0.03

Table 1.4: Summary statistics: trucking vs. controls (2005)

This table presents descriptive statistics that compare treated and control firms in 2005. The trucking industry is the treated group (13,497 firms). The control group includes firms from the security, cleaning and other business services industries (13,992 firms). The matched control group includes the nearest neighbor to each treated firm (13,497 firms) based on a propensity score matching procedure based on size (log of assets), receivables over sales and sales growth, all measured one year prior to the trade credit regulation reform (2006). The null hypothesis is that the distribution functions are equal. T-test present the t statistic of a the difference in means between the treated and the control and matched control group.

	Treated firms			Control firms				Matched control firms			
	Obs.	Mean	Median	Obs.	Mean	Median	t-test	Obs.	Mean	Median	t-test
Matching variables											
Receivables over sales	13497	0.244	0.222	13992	0.238	0.190	0.033	13497	0.246	0.198	0.532
Log assets	13497	6.207	6.114	13992	5.806	5.654	0.000	13497	6.206	6.114	0.964
Sales growth	13497	0.056	0.045	13992	0.057	0.032	0.833	13497	0.053	0.035	0.442
Dependent variables											
Receivables over assets (t-1)	13497	0.539	0.516	13992	0.417	0.370	0.000	13497	0.420	0.377	0.000
Cash over assets (t-1)	13497	0.161	0.094	13992	0.259	0.161	0.000	13497	0.255	0.160	0.000
Credit lines over assets (t-1)	13497	0.032	0.000	13992	0.025	0.000	0.000	13497	0.023	0.000	0.000
Dividend over assets (t-1)	13497	0.017	0.000	13992	0.028	0.000	0.000	13497	0.031	0.000	0.000
Payables over assets (t-1)	13497	0.255	0.191	13992	0.213	0.120	0.000	13497	0.210	0.120	0.000
Capex over assets (t-1)	13497	0.087	0.031	13992	0.047	0.010	0.000	13497	0.047	0.012	0.000
ROA	13497	0.298	0.203	13992	0.213	0.149	0.000	13497	0.207	0.146	0.000

Table 1.5: Impact of the trade credit reform (2006): difference-in-differences

This table presents the effect of the trade credit regulation (2006) on a variety of corporate policies. The trucking industry is the treated group (13,497 firms). The control group includes firms from the security, cleaning and other business services industries (13,992 firms). The matched control group includes the nearest neighbor to each treated firm (13,497 firms) based on a propensity score matching procedure based on size (log of assets), receivables over sales and sales growth, all measured one year prior to the trade credit regulation reform (2006). Standard errors are corrected for clustering at the firm level and presented in parenthesis. ***,** and * indicate significance at the 1, 5 and 10% respectively.

		Control group		Matched control group	
	Mean treatment difference (2007 vs. 2005)	Mean control difference (2007 vs. 2005)	Difference in differences	Mean matched control difference (2007 vs. 2005)	Difference in differences
Receivables over assets (t-1)	-0.079***	-0.024***	-0.054***	-0.026***	-0.053***
	(0.002)	(0.003)	(0.004)	(0.004)	(0.005)
Payables over over assets (t-1)	-0.020***	-0.009***	-0.011***	-0.012***	-0.008**
	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
Cash over over assets (t-1)	0.056***	-0.000	0.056***	-0.003	0.058***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
Credit lines over over assets (t-1)	-0.008***	-0.001*	-0.007***	-0.001	-0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Dividends over assets (t-1)	0.003***	0.003***	0.000	0.002**	0.001
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Capex over assets (t-1)	-0.004***	-0.009***	0.004**	-0.010***	0.005**
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
ROA	0.013***	-0.023***	0.036***	-0.020***	0.033***
	(0.003)	(0.003)	(0.004)	(0.005)	(0.005)

Table 1.6: Impact of the trade credit reform (2006): placebo test

This table presents a placebo test of the effect of the trade credit regulation on a variety of corporate policies. The trucking industry is the treated group (13,403 firms). The control group includes firms from the security, cleaning and other business services industries (13,264 firms). The matched control group includes the nearest neighbor to each treated firm (13,403 firms) based on a propensity score matching procedure based on size (log of assets), receivables over sales and sales growth, all measured in 2003. Standard errors are corrected for clustering at the firm level and presented in parenthesis. ***,** and * indicate significance at the 1, 5 and 10% respectively.

		Control group		Matched control group	
	Mean treatment difference (2005 vs. 2003)	Mean control difference (2005 vs. 2003)	Difference in differences	Mean matched control difference (2005 vs. 2003)	Difference in differences
Receivables over assets (t-1)	-0.013***	-0.030***	0.018***	-0.038***	0.026***
	(0.002)	(0.003)	(0.004)	(0.004)	(0.005)
Payables over over assets (t-1)	0.002	-0.011***	0.013***	-0.016***	0.018***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
Cash over over assets (t-1)	-0.005***	-0.001	-0.004	-0.005	-0.001
	(0.001)	(0.002)	(0.003)	(0.003)	(0.004)
Credit lines over over assets (t-1)	0.001	-0.003***	0.003***	-0.004***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Dividends over assets (t-1)	0.000	0.002***	-0.001**	0.002**	-0.001
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Capex over assets (t-1)	-0.010***	-0.009***	-0.001	-0.011***	0.001
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
ROA	-0.029***	-0.033***	0.004	-0.036***	0.007
	(0.003)	(0.003)	(0.004)	(0.005)	(0.005)

Table 1.7: Heterogeneous response to the trade credit reform (2006): difference-in-differences

This table presents the DID estimation of the effect of the trade credit regulation (2006) on a variety of corporate policies between 2005 and 2007, conditional on firm size, payout status and age. The trucking industry is the treated group (13,497 firms). The control group includes firms from the security, cleaning and other business services industries (13,992 firms). The matched control group includes the nearest neighbor to each treated firm (13,497 firms) based on a propensity score matching procedure based on size (log of assets), receivables over sales and sales growth, all measured one year prior to the trade credit regulation reform (2006). Large firms are firms in the top half of the distribution of assets in 2005, Payout firms are firms that distributed dividends in 2005 and Old firms are firms in the top half of the distribution of age. Standard errors are corrected for clustering at the firm level and presented in parenthesis. ***,** and * indicate significance at the 1, 5 and 10% respectively.

Panel A: treated vs. controls

Differences-in-differences	Small	Large	Small minus Large	No payout	Payout	No payout minus Payout	Young	Old	Young minus Old
Receivables over assets (t-1)	-0.048***	-0.051***	0.003	-0.049***	-0.074***	0.026***	-0.048***	-0.065***	0.017**
	(0.005)	(0.004)	(0.007)	(0.004)	(0.006)	(0.007)	(0.006)	(0.004)	(0.007)
Payables over over assets (t-1)	-0.006	-0.008**	0.002	-0.008***	-0.018***	0.010**	-0.006	-0.019***	0.013**
	(0.004)	(0.004)	(0.005)	(0.003)	(0.004)	(0.005)	(0.004)	(0.003)	(0.005)
Cash over over assets (t-1)	0.054***	0.063***	-0.009*	0.054***	0.064***	-0.010*	0.051***	0.055***	-0.004
	(0.004)	(0.003)	(0.005)	(0.003)	(0.005)	(0.006)	(0.004)	(0.004)	(0.005)
Credit lines over over assets (t-1)	-0.007***	-0.006***	-0.001	-0.008***	-0.003***	-0.005***	-0.007***	-0.007***	0.000
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Dividends over assets (t-1)	-0.000	0.001	-0.002	-0.004***	0.017***	-0.021***	-0.000	0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Capex over assets (t-1)	0.009***	0.002	0.007*	0.005***	0.001	0.004	0.006**	0.003	0.002
	(0.003)	(0.002)	(0.004)	(0.002)	(0.003)	(0.004)	(0.003)	(0.002)	(0.004)
ROA	0.035***	0.037***	-0.002	0.034***	0.046***	-0.012	0.033***	0.035***	-0.002
	(0.007)	(0.004)	(0.008)	(0.005)	(0.006)	(0.008)	(0.006)	(0.005)	(0.008)

Panel B: treated vs. matched controls

Differences-in-differences	Small	Large	Small minus Large	No payout	Payout	No payout minus Payout	Young	Old	Young minus Old
Receivables over assets (t-1)	-0.047***	-0.058***	0.011	-0.046***	-0.072***	0.026***	-0.044***	-0.066***	0.022**
	(0.007)	(0.006)	(0.009)	(0.006)	(0.007)	(0.009)	(0.007)	(0.006)	(0.009)
Payables over over assets (t-1)	-0.005	-0.011**	0.006	-0.003	-0.021***	0.017**	-0.006	-0.014***	0.008
	(0.005)	(0.005)	(0.007)	(0.004)	(0.005)	(0.007)	(0.005)	(0.005)	(0.007)
Cash over over assets (t-1)	0.055***	0.062***	-0.007	0.055***	0.069***	-0.014*	0.050***	0.064***	-0.014*
	(0.006)	(0.005)	(0.007)	(0.004)	(0.006)	(0.008)	(0.005)	(0.005)	(0.007)
Credit lines over over assets (t-1)	-0.007***	-0.007***	0.000	-0.008***	-0.004***	-0.004**	-0.007***	-0.008***	0.002
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Dividends over assets (t-1)	0.000	0.002	-0.001	-0.005***	0.017***	-0.022***	0.000	0.002	-0.002
	(0.001)	(0.002)	(0.002)	(0.001)	(0.003)	(0.003)	(0.001)	(0.002)	(0.002)
Capex over assets (t-1)	0.008**	0.002	0.006	0.006**	0.002	0.004	0.006**	0.004	0.003
	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.005)	(0.003)	(0.003)	(0.004)
ROA	0.029***	0.037***	-0.009	0.026***	0.053***	-0.026**	0.029***	0.037***	-0.007
	(0.008)	(0.007)	(0.010)	(0.006)	(0.008)	(0.010)	(0.008)	(0.007)	(0.010)

Table 1.8: Impact of the trade credit reform on default, difference-in-differences

This table presents the results of a firm level DID estimation of the impact of the 2006 trade credit regulation reform on survival. Bankruptcy is a dummy equal to one on the year of the firm's bankruptcy, and zero in other years. Post is a dummy equal to one in year 2007 and zero in 2005. The trucking industry is the treated group. The control group includes firms from the security, cleaning and other business services industries. Specifications in column 2 and 5 include firm level controls measured with a two year lag: log of assets, PPE over assets, receivables over sales, sales growth and debt over assets. Columns 3 and 6 include the same controls interacted with post. Coefficients on control variables are not reported for the clarity of exposition. Standard errors are corrected for clustering at the four digit sector level. ***,** and * indicate significance at the 1, 5 and 10% respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Bankruptcy					
Post x treated	-0.012***	-0.010***	-0.011***	-0.012***	-0.009***	-0.011***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Post	-0.001	-0.001	-0.004	-0.001	-0.002	-0.005
	(0.002)	(0.002)	(0.005)	(0.001)	(0.002)	(0.005)
treated	0.011*	0.012**	0.011**			
	(0.006)	(0.005)	(0.004)			
Constant	0.021***	0.034***	0.049***	0.026***	0.039***	0.054***
	(0.005)	(0.007)	(0.011)	(0.001)	(0.005)	(0.009)
Firm level controls at year -2	No	Yes	Yes	No	Yes	Yes
Post x firm level controls at year -2	No	No	Yes	No	No	Yes
Four-digit sector FE	No	No	No	Yes	Yes	Yes
Observations	77,579	60,706	56,876	77,579	60,706	56,876
R-squared	0.001	0.019	0.004	0.006	0.022	0.008

Table 1.9: Cohort analysis of default probability at creation

This table presents a regression of the default probability of firms in the second, third and fourth year following their creation. Treated is the set of firms 2006 trade credit regulation reform.The control group includes firms from the security, cleaning and other business services industries. All regressions firm level controls measured at creation: log number of employees, log assets per employee, log sales per employee, log debt per employee, roa, log intangible assets per employee. All regressions include cohort and four-digit sector fixed effects. Coefficients on controls, post, year dummies and constant are not reported for the clarity of exposition. Standard errors are corrected for clustering at the four-digit sector. ***,** and * indicate significance at the 1, 5 and 10% respectively.

	(1) Bankruptcy before year 2	(2) Bankruptcy before year 3	(3) Bankruptcy before year 4
Treated x 2007 cohort	-0.007	-0.001	
	(0.006)	(0.007)	
Treated x 2006 cohort	-0.020***	-0.024***	-0.019**
	(0.006)	(0.008)	(0.008)
Treated x 2005 cohort	-0.011**	-0.021***	-0.019*
	(0.004)	(0.006)	(0.009)
Treated x 2004 cohort	-0.004	-0.011	-0.013*
	(0.005)	(0.007)	(0.007)
Constant	0.021***	0.042***	0.069***
	(0.004)	(0.007)	(0.009)
Firm level controls at creation	Yes	Yes	Yes
Four digit sector FE	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
Observations	36,655	36,655	28,361
R-squared	0.026	0.049	0.063

Table 1.10: Summary statistics: High vs. Low Transport Users (2005)

This table presents summary statistics for high and low transportation users in a sample of 4,933 manufacturing firms. Firms are allocated to the High (Low) Transport Users group if their ratio of external purchases of transport services to total purchases in 2005 are in the top (bottom) 2 quintiles of the distribution.

		Obs.	Mean	Median	Std. dev.
Low Transport Users	Share of external transport purchases	1614	0.009	0.009	0.007
	Sales growth	1614	0.027	0.019	0.115
	Delta payables	1614	-0.005	-0.005	0.100
	Fixed assets	1614	0.166	0.121	0.153
	Log assets	1614	9.798	9.681	1.808
High Transport Users	Share of external transport purchases	1607	0.082	0.064	0.062
	Sales growth	1607	0.018	0.016	0.097
	Delta payables	1607	-0.013	-0.012	0.094
	PPE over assets	1607	0.182	0.151	0.138
	Log assets	1607	9.522	9.401	1.481

Table 1.11: Change in payables of High vs. Low Transport users

This table presents the results of an estimation of the change (difference) in payables over total purchases around the trade credit regulation reform (between 2005 and 2007) for high and low transportation users in a sample of 4,933 manufacturing firms. Firms are allocated to the High (Low) Transport Users group if their ratio of external purchases of transport services to total purchases in 2005 are in the top (bottom) 2 quintiles of the distribution. Concentrated (profitable, growing, exporting) industries are industries which Herfindahl index based on value added (respectively aggregate net income over aggregate sales, growth in aggregate value added and share of exports) lies above the median. Standard errors are corrected for clustering at the three-digit sector level. ***,** and * indicate significance at the 1, 5 and 10% respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Difference in payables over total purchases, 2005-2007									
High Transport Users (HTU)	-0.010***	-0.009***								
	(0.003)	(0.003)								
HTU x high concentrated sectors			-0.016***	-0.015***						
			(0.005)	(0.004)						
HTU x low concentrated sectors			-0.004	-0.004						
			(0.004)	(0.004)						
HTU x high profitable sectors					-0.019***	-0.018***				
					(0.004)	(0.004)				
HTU x low profitable sectors					-0.002	-0.001				
					(0.004)	(0.004)				
HTU x high growing sectors							-0.013***	-0.012***		
							(0.004)	(0.004)		
HTU x low growing sectors							-0.008*	-0.007*		
							(0.004)	(0.004)		
HTU x high exporting sectors									-0.013***	-0.013***
									(0.004)	(0.004)
HTY x low exporting sectors									-0.007*	-0.006
									(0.004)	(0.004)
Constant	-0.004	-0.005	-0.004*	-0.004	-0.004*	-0.005	-0.004*	-0.006	-0.004*	-0.005
	(0.002)	(0.012)	(0.002)	(0.013)	(0.002)	(0.012)	(0.002)	(0.012)	(0.002)	(0.012)
Difference in coefficients			-0.012**	-0.012*	-0.017***	-0.017***	-0.005	-0.004	-0.006	-0.007
			(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Three digit sector FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm level controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	3,220	3,220	3,220	3,220	3,220	3,220	3,220	3,220	3,220	3,220
R-squared	0.042	0.050	0.043	0.050	0.044	0.051	0.042	0.050	0.043	0.050

Table 1.12: Local competition and trade credit provision

This table replicates table 3 in Petersen and Rajan (1997) on the cross-section of treated firms in 2005. The trucking industry is the treated group (13,497 firms). Log nb. firms, own sector, own district is the log of the total number of firms of the same (four-digit) sector in the same district as of 2005. Herfindahl index, own sector, own district is the Herfindahl index, computed based on 2005 value added, of the firm's (four-digit) sector in its district. Median Herfindahl index, own district is the median Herfindahl index across sectors in the firm's district, based on 2005 value added. Standard errors are robust to heteroskedasticity. ***,** and * indicate significance at the 1, 5 and 10% respectively.

	(1)	(2)	(3)	(4)	(5)
	Receivables over sales (2005)				
Log assets	0.019***	0.019***	0.018***	0.018***	0.018***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log age	0.011	0.011	0.009	0.009	0.009
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
(Log age)2	-0.001	-0.001	-0.000	-0.000	-0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Credit lines over sales	0.574***	0.575***	0.596***	0.589***	0.592***
	(0.096)	(0.095)	(0.095)	(0.095)	(0.095)
Net profit over sales	0.078*	0.084**			
	(0.040)	(0.038)			
Net profit over sales if positive			-0.022	-0.021	-0.021
			(0.014)	(0.014)	(0.014)
Net profit over sales if negative			-0.051***	-0.051***	-0.051***
			(0.014)	(0.014)	(0.014)
Sales growth if positive	0.038***	0.031***	0.028***	0.028***	0.028***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Sales growth if negative	-0.185***	-0.167***	-0.175***	-0.176***	-0.177***
	(0.027)	(0.027)	(0.026)	(0.027)	(0.026)
Gross margin	0.004	0.024	0.035**	0.033**	0.035**
	(0.022)	(0.017)	(0.015)	(0.015)	(0.015)
(Gross margin)2		0.078***	0.073***	0.072***	0.072***
		(0.023)	(0.022)	(0.022)	(0.022)
Log nb. firms, own sector, own district				0.010***	
				(0.003)	
Log nb. firms, own district				-0.008**	
				(0.003)	
Herfindahl index, own sector, own district					-0.051***
					(0.018)
Median herfindahl index, own district					-0.023
					(0.024)
Four-digit sector FE	Yes	Yes	Yes	Yes	Yes
Observations	13,411	13,411	13,411	13,408	13,408
R-squared	0.104	0.113	0.115	0.116	0.116

Table 1.13: Summary statistics: trucking vs. controls (1999)

This table presents descriptive statistics that compare treated and control firms in 1998. The trucking industry is the treated group (11,478 firms). The control group includes firms from the security, cleaning and other business services industries (11,761 firms). The matched control group includes the nearest neighbor to each treated firm (11,478 firms) based on a propensity score matching procedure based on size (log of assets), tangibility (PPE over assets), and sales growth, all measured in 1999. The null hypothesis is that the distribution functions are equal. t-test present the t statistic of a the difference in means between the treated and the control and matched control group.

	Treated firms			Control firms				Matched control firms			
	Obs.	Mean	Median	Obs.	Mean	Median	t-test	Obs.	Mean	Median	t-test
Matching variables											
PPE over assets	11478	0.210	0.171	11671	0.135	0.076	0.000	11478	0.211	0.174	0.536
Log assets	11478	7.619	7.575	11671	7.491	7.336	0.000	11478	7.549	7.402	0.000
Sales growth	11478	0.067	0.045	11671	0.075	0.042	0.056	11478	0.069	0.039	0.557
Dependent variables											
Delta receivables over sales	11478	0.002	-0.005	11671	0.012	0.000	0.000	11478	0.013	0.000	0.000
Receivables over sales	11478	0.230	0.219	11671	0.245	0.199	0.000	11478	0.217	0.184	0.000

Table 1.14: Change in trade credit provision following an increase in barriers to entry

This table presents the DID estimation of the effect of the change in barriers to entry to the trucking industry on the change (difference) in receivables over sales between 1999 and 2001. The trucking industry is the treated group (11,478 firms). The control group includes firms from the security, cleaning and other business services industries (11,671 firms). The matched control group includes the nearest neighbor to each treated firm (11,478 firms) based on a propensity score matching procedure based on size (log of assets), tangibility (PPE over assets), and sales growth, all measured one year prior to the change in barriers to entry (1999). Large (cash rich) firms are firms which sales (cash over assets) lie above the median of their four-digit sector in 1999. Standard errors are corrected for clustering at firm level. ***,** and * indicate significance at the 1, 5 and 10% respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Difference in receivables over sales, 1999-2001								
	Treated versus controls						Treated versus matched controls		
Treated	-0.009***	-0.010**					-0.011**		
	(0.003)	(0.003)					(0.004)		
Treated x large (sales)			-0.013***	-0.013**				-0.014**	
			(0.004)	(0.005)				(0.006)	
Treated x small (sales)			-0.006**	-0.006*				-0.007*	
			(0.003)	(0.003)				(0.003)	
Treated x cash rich					-0.015***	-0.016***			-0.019**
					(0.004)	(0.004)			(0.006)
Treated x cash poor					-0.004	-0.005			-0.003
					(0.003)	(0.004)			(0.004)
Constant	0.012***	0.031***	0.010***	0.040***	0.011**	0.028***	0.013***	0.011***	0.001
	(0.003)	(0.007)	(0.002)	(0.009)	(0.004)	(0.010)	(0.004)	(0.003)	(0.004)
Difference in coefficients			-0.007**	-0.007**	-0.011**	-0.011**		-0.007	-0.016**
			(0.003)	(0.003)	(0.004)	(0.004)		(0.005)	(0.006)
Firm level controls	No	Yes	No	Yes	No	Yes	No	No	No
Observations	23,149	23,149	23,149	23,149	23,149	23,149	22,956	22,956	22,956
R-squared	0.001	0.003	0.001	0.005	0.001	0.003	0.001	0.001	0.004

Chapter 2

Investor Horizon and Innovation: Evidence from Private Equity Funds

Abstract

Investments exploring new ideas typically take more time to payoff than investments exploiting existing ones. Hence, investors with a short horizon may be tilted towards the latter. I consider investments made by private equity funds, which generally have a limited investment horizon contractually fixed *ex ante*. I use between and within-fund variations in investment horizon to show that funds with a longer horizon select younger companies at an earlier stage of their development, stage investment more and hold on to their investments for a longer period of time. Moreover, companies which receive funding from funds with a longer horizon increase their patent stock significantly more than companies which receive funding from investors with a shorter horizon. Altogether, these results provide new evidence on the behavior of private equity funds throughout their life cycle and suggest that investor horizon matters to an important extent for the funding of corporate innovation.

JEL classification: M14, O31, G24, G34

Keywords: horizon, innovation, patents, venture capital, private equity

2.1 Introduction

Assets managed by private equity funds totaled $2,4 trillion in 2010 which amounts to approximately 5% of the world market capitalization[1]. These funds usually buy significant blocks of shares in private companies an divest after a few years through the initial public offering of the company or the sale to third party. However, little is known about how this class of investors select the companies they provide funding to. One important yet unexplored channel potentially affecting their decisions is their investment horizon. Contrary to most financial intermediaries, private equity funds generally have an investment lifetime of ten years fixed *ex ante*[2]. Funds are raised in year 0 from outside investors - "Limited Partners" (LPs) - and trusted to fund managers - "General Partners" (GPs), who invest and return funds and capital gains to LPs within ten years. This paper studies how private equity fund managers investment decisions shift throughout their fund's life cycle.

In frictionless capital markets, stock prices reflect firms fundamental value, assets are continuously traded and returns follow a random walk. Horizon does not affect asset allocation. Investors can meet their liquidity needs anytime by selling their shares before the firm's investments payoff, at no discount. Horizon may matter, however, when (i) there are asymmetries of information between sellers and buyers, and when (ii) investors are faced with projects exploiting existing ideas *and* projects exploring new ideas. Exploration arguably takes longer to produce observable outcomes and payoffs (Manso, 2011). Combined with asymmetries of information, this heterogeneity in the timing of payoffs implies that investor horizon matters. Two projects with the same net present value but involving different levels of exploration are of different interests for a short and a long-term investor.

I test the validity of this story in the context of private equity investments. I compare the behavior of funds with heterogeneous investment horizon as well as the evolution of a

[1]The World Federation of Exchanges estimates the worldwide market cap at $54.9 trillions; CityUK and Preqin estimate the worldwide assets under management of private equity funds to be $2.4 trillion

[2]In the case of Venture Capital funds, see Sahlman (1990); Gompers and Lerner (1996); Gompers (1996); Gompers and Lerner (1999); Lerner and Schoar (2004). In the case of Buyout funds, see Metrick and Yasuda (2009); or in the law literature: Masulis and Thomas (2009), Harris (2010), Cheffins and Armour (2007).

given fund's behavior as it moves closer to the end of its investment life. I expect private equity funds with a longer investment horizon to invest in more exploratory ventures, and to shift towards less exploratory investments as their horizon shrinks. I use a sample of private equity investments involving over 1,500 private equity funds from 1980 to 2010 to check whether this is indeed the case[3]. I show that funds further away from liquidation invest in younger companies at an earlier stage of their development. Moreover, funds with a longer horizon split their investment more through time and hold on to their investments for a longer period of time. Overall, with respect to a fund in its first year of activity, a fund in its fourth year of activity (i) invests in companies that are 1.2 to 2.9 months older, (ii) has a 4 to 5% larger probability to invest in companies that are beyond the seed or early stage of their development, (iii) selects companies which have already received 0.06 to 0.11 more rounds of financing in the past, (iv) stages its investments by 5 to 9% less and holds on to its investments for 1 to 6% longer. Although I cannot observe the novelty of companies projects ex ante, I analyze it ex post with standard patent-based metrics. Companies which receive funding from funds with a longer horizon increase their patent stock by 0.07 to 0.17 more patents following the investment than companies which receive funding from short horizon investors. Moreover, patents applied by companies which receive funding from funds with a longer horizon receive 0.06 to 0.18 more citations. Finally, I compare the sector exposure of funds with short and long horizon. I show that funds with a short horizon invest in industries which experience lower subsequent excess returns. A monthly rebalanced portfolio long in sectors invested by long horizon funds and short in sectors invested by short horizon funds yields positive and significant risk adjusted returns of the order of 96 basis points per year at the three and four year horizon. I interpret these results as evidence that long term investors

[3]VentureXpert collects information on both Venture Capital, Buyout and other funds. Although these types of funds differ in their investment style, they share the same contractual structure, with a usual finite lifespan of ten years. For the purpose of this study, I restrict the sample to funds which make more than half of their investments in companies which development stage is classified as "Startup/seed", "Early stage", "Expansion stage" or "Later stage" . Although the final sample is dominated by funds labeled by VentureXpert as Venture Capital, it also includes a small number of Buyout and Other funds. Results in untabulated regressions show that the results are identical when I keep all funds or when I restrict the analysis to funds labeled as Venture Capital only.

have a larger propensity to target sectors with mid-term rather than immediate prospects.

I carefully analyze and rule out alternative explanations for this shift in investment style. First, compensation and career concerns which GPs face might affect their risk aversion and their propensity to select exploratory projects throughout their fund's life. Suppose that compensation contracts are set in a manner that reduces the risk appetite of GPs through time, then this might explain that they fund less innovative companies as they move closer to liquidation. Compensation agreements in private equity funds typically give GPs a fixed management fee that is a percentage (around 2%) of the amount of capital committed to the fund, as well as a call option on a share (almost always 20%) of the fund's total cumulative profits which they receive at the fund's liquidation: the carried interest. The option-like nature of this compensation generates risk incentives when earlier performance has been low. If funds systematically perform well in their first years, their incentives to take risks might decrease with time and tilt their asset selection towards less risky targets. I address this concern by controlling for the fund's past performance. I find interesting evidence that funds with a lower track record of successful exits tend to select more innovative companies.

Moreover, it could be the case that funds focus on exploration to show skills to their investors in order to raise a follow-on fund and then shift towards more mature projects. Since private equity firms raise follow-on funds every three to five years on average, this is likely to influence their portfolio management to an important extent, especially in the case of first time funds (Gompers, 1996). I control for fundraisings in the regressions, to insure that this important pattern in a fund's life is not driving the results. I also use a "first time fund" dummy to insure that the results are not driven by the risk incentives of first time fund managers.

Altogether, the results presented in this paper suggest that horizon is a strong driver of private equity funds asset allocation. On a broader note, they stress the relevance of the horizon of corporate owners in the funding of new and innovative ideas.

This paper relates to several streams of the literature, the first of which focuses on the economic impact of private equity funds contractual structure. The structure of private equity funds is very similar across funds in North America and in Europe. Gompers and Lerner (2001) argue that the first (Venture Capital) limited partnership was formed in 1958 and followed the template of other limited partnerships common at the time such as those that had been formed to develop real estate projects and explore oil fields, and which had predetermined finite lifetimes of usually 10 years. Lerner and Schoar (2004) show that restrictions imposed by private equity funds on their investors are aimed at selecting liquid investors so as to invest in industries with longer cycles. Axelson et al. (2009) argue that giving incentives to managers in a series of deals rather than on a deal by deal basis prevents them from taking excessive risk. Kandel et al. (2011) model the behavior of Venture Capital funds and find that the age of the fund should have an important effect on the type of projects it takes and on the tendency to continue or stop projects. On the empirical side, the impact of fund age on performance metrics has been investigated in a few papers. Closer to liquidation, funds sell companies cheaper (Masulis and Nahata, 2009). Younger funds invest in riskier buyouts towards the end of their life, especially if they have underperformed earlier in their life, in an attempt to achieve superior performance (Ljungqvist et al., 2008). Leverage buyouts sponsored by private equity funds with more experience exit earlier, and funds that are publicly traded (and hence have an infinite investment horizon) take more time to exit their investments (Strömberg, 2008). Finally, Robinson and Sensoy (2011) show that cash flow variation within private equity funds is mostly idiosyncratic and that most predictable variation is explained by the age of the fund.

This paper also adds to a recent literature on the economic impact of investment horizon. Bushee (1998) finds that a large proportion of ownership by financial institutions that have a high portfolio turnover and engage in momentum trading significantly increases the probability that managers reduce R&D to reverse an earnings decline. Polk and Sapienza (2009) and Derrien et al. (2011) also use portfolio turnover rate as a proxy of shareholders time horizon. They show that when firms are mispriced, the horizon of their shareholders affects their corporate policy. Cella et al. (2010) find that institutional investors with shorter trading horizons where incited to sell their holdings to a larger extent than investors with longer horizons following Lehman Brother's bankruptcy, thereby contributing to the amplification of this market shock. Xu (2011) uses CEO employment contracts and finds that contract horizon is positively correlated with both capital expenditure and R&D expenses. This literature faces several challenges. The first of them is the identification of a credible proxy for investment horizon. To date, the literature has mainly used the turnover rate of institutional investors' portfolio. I use the time to liquidation as a measure of investor horizon. It is plausibly more exogenous than previously used measures since the evolution of a given private equity fund time horizon is deterministic and unlikely to be correlated with unobserved variables also correlated with investigated outcomes. Moreover, I am able to study *within investor* changes in investment horizon by using fund fixed effects. The second challenge that this literature faces is to identify causality from selection. The contribution of this paper is to emphasize that the horizon of investors has a strong impact on their asset selection.

Finally, the results presented here add to a growing literature on the relationship between corporate ownership and innovation. Lerner and Wulf (2007) examine the patenting activity in US corporations and show that long-term incentives (e.g. stock options and restricted stock) are associated with heavily cited patents. Ferreira et al. (2012) model the impact of public and private ownership structures on firms incentives to invest in innovative projects. They argue that publicly listed firms choose more conventional projects than private firms. Lerner et al. (2011) show that patent count does not decrease and that citation count increases when companies undergo a LBO. In the case of venture capital, Tian and Wang (2011) show that private equity firms that have shown a high propensity to finance failed ventures in the past (and thus have a high tolerance for early failure) invest in companies that are more innovative. Bernstein (2011) compares the patenting activity of firms going public and firms withdrawing their IPO and finds that going public reduces the quality of innovation. Chemmanur et al. (2011) show that companies held by corporate venture capitalists patent more actively than companies held by independent venture capitalists, an effect that they attribute to a difference in tolerance for early failure.

The rest of the paper is organized as follows. Section 2.2 develops the hypothesis to be tested, section 2.3 presents the sample and data. Section 2.4 presents the results and section 2.5 concludes.

2.2 Theoretical framework and empirical predictions

The theoretical framework used in this paper borrows from the representation of innovation in Manso (2011) and Ferreira et al. (2012). I provide a simple illustrative model to fix ideas.

Investors can provide funding to two company types which are operational for two periods. Companies of the first type exploit existing ideas while companies of the second type explore new ideas. Type 1 delivers cash flows of 1 with probability p and 0 otherwise. If 1 is obtained in the first period, then type 1 delivers 1 again with probability p in the second period and 0 otherwise. If 0 is obtained in the first period, then the company shuts down. Type 2 has a similar payoff structure and delivers 1 with probability δp and 0 otherwise in period 1. If 1 is obtained in the first period, then type 2 delivers 1 again with probability θp in the second period and 0 otherwise. If 0 is obtained in the first period, then the company shuts down.

I assume for simplicity that δ and θ are such that both types have the same net present value over the two periods:

$$\delta p + \delta\theta p^2 = p + p^2$$

However, type 2 is less profitable than type 1 in period 1 and the reverse is true in period 2:

$$\delta < 1 \text{ and } \theta > 1$$

Suppose there are short-term and long-term risk neutral investors deciding of a unique investment at the beginning of period 1. Short-term investors have to liquidate their investment by selling it to outside short-term investors at the end of period 1. Long-term investors can hold on to their investment for two periods and have the option to sell their investment at the end of period 1 to outside short-term investors.

Outside short-term investors can buy companies at the end of period 1 when the initial investors wish to sell them. They observe interim results (1 or 0) and whether the initial investor is short or long-term. But they do not observe the type of the company, i.e. whether the company is exploiting an existing idea or exploring a new one.

At the end of period 1, if investors wish to liquidate their investment, they need to agree on a price with outside buyers. These potential buyers will bid a price not less than their estimation of the residual project cash-flows, conditional on the information they observe at the end of period 1. Let us call this estimation E. Since they do not know and have no way to observe types, they will offer a single price E such that:

$$p \leq E \leq \theta p$$

Coming back to the beginning of period 1, long term investors are indifferent between the two types since they both have the same net present value. The only way for them to make more money would be to invest in type 1 and try to sell it at a price $E > p$ at the end of period 1. However, outside investors would be aware of that and offer a price $E = p$ to long-term investors, making them indifferent between selling or keeping type 1 company.

Consider now the decision of short-term investors. Since they live only one period, they have to sell their investment a the end of period 1. The payoff from funding type 1 is $p + pE$ while the payoff of funding type 2 is $\delta p + \delta pE$. Since $\delta < 1$, short-term investors will always select type 1.

This simple model shows that when there is information asymmetry between buyers and sellers and when there are projects exploiting existing ideas *and* projects exploring new ideas, then short-term investors have a preference for projects maturing more quickly.

I use this simple theoretical framework to generate a set of testable predictions.

Prediction 1: Funds with a longer horizon invest in less mature companies. Start-up companies are by essence the place of exploration of new ideas rather than the exploitation of existing ones. Of course, some established firms might decide to take some radical strategic orientation and seek private equity funding to finance exploration. However, I hypothesize that a larger share of each invested dollar finances exploration in a company in its first year than in a company in its fifth year of operations. In addition to their age, less mature companies are likely to be at an earlier stage of their development and to have received less external finance prior to the investment. Companies which are less mature along these three dimensions are likely to be more exploratory in the sense of the theoretical framework presented above. Hence, I expect funds with a longer investment horizon to select younger companies, companies at an earlier stage of their development or company which previously received less rounds of financing than funds with a shorter horizon. I also expect a given fund to shift its investments towards more mature companies as its investment horizon shrinks.

Prediction 2: Funds with a longer horizon stage investment more and hold on to their investments longer. Funds often "stage" their funding of young companies, especially Venture Capital funds. They split funding through time and sometimes also condition new funding to the achievement of certain operational milestones. Staging is a way to overcome agency costs related to low asset tangibility or high asset specificity (Gompers, 1995). Moreover, staging is a way for private equity funds to buy an option on a company's future equity (Bergemann and Hege, 1998). Staging is thus more likely to be prevalent in firms involved in exploration rather than exploitation. As a result, I expect funds with a longer investment horizon to stage investment more. Moreover, given that the exploration takes more time to yield observable payoffs, funds with a longer horizon should also hold on to their investments more than funds with a shorter period of time. Hence, I expect that as their horizon shrinks, private equity funds hold on to their investments for a shorter period of time.

Prediction 3: Funds with a longer horizon invest in more innovative companies. Consistent with the framework highlighted above, funds with a longer investment horizon should select companies which do more innovation and produce more or newer ideas. Although I do not observe the novelty of companies projects ex ante, I analyze it ex post. The most standard way to measure innovation is to compute the number of patents applied by a given company in a given year, and the number of citations that these patents receive. I expect companies which obtain funding from funds with a longer horizon to produce more patents and more cited patents than companies which obtain funding from funds with a shorter horizon. Again, funds should shift their portfolio of investments toward less innovative companies as their horizon shrinks.

Prediction 4: Funds with a longer horizon invest in sectors which experience higher subsequent risk adjusted returns. The framework presented above applies, to a certain extent, to the selection of sectors by investors with heterogeneous investment horizons. An investor with a longer investment horizon is more likely to select sectors with mid-term rather than intermediate prospects. In particular, she is more likely to pick undervalued sectors which experience subsequent risk adjusted returns. Sectors could be undervalued because of high uncertainty or because of investor sentiment, for instance. In either case, an investor with favorable private information is more likely to take a long position in these sectors when she has a long enough horizon to wait for uncertainty or sentiment to be resolved and prices to revert back to fundamentals. Hence I expect sectors selected by funds with a longer horizon to subsequently experience higher returns in excess of their exposure to systematic risk factors.

2.3 Data and sample

SDC Platinum VentureXpert (henceforth "Venture Xpert" or "SDC") is the main source used in this paper. SDC provides information on private equity investments between 1962 and 2010. For the purpose of this study, I restrict the sample to investments in the period running from 1980 to 2010. I then filter private equity limited partnerships as follows: I include an investment in the sample if the involved fund's name includes the acronym "LP", "L.P." or "Limited partnership" and excludes the term "Evergreen"[4], I exclude all funds for which the parent private equity firm is unknown, I exclude private equity firms which raised only one fund during the sample period, I restrict the sample to funds involved in at least two private equity rounds, I drop "real estate funds" and "fund of funds". I drop funds raised before 1971 and funds for which SDC does not provide either the "initial closing date" or the "fund year" which enable me to identify the starting point of the life of the fund. VentureXpert collects information on both Venture Capital, Buyout and other funds. Although these types of funds differ in their investment style, they share the same contractual structure, with a finite lifespan of ten years. For the purpose of this study, I restrict the sample to funds which make more than half of their investments in companies at the "seed", "early stage", "expansion stage" or "later stage" of their development[5].

To measure the investment horizon of any given fund in the sample at the time of any investment, I build a variable which I call fund age. Fund age is the difference in years between the month of an investment and the month when the fund was created. The creation date of a fund is a noisy concept: one could consider the date when the fund was launched, the date of its first closing or the date of its final closing. I identify the creation of the fund as the "initial closing date" provided by SDC. The "initial closing date" is unavailable for 25% of funds in the sample. In this case I use the "fund year" provided by SDC and set the creation of the fund in January of this year. Finally, when there are investments in the database prior to the fund creation date I computed, I reset the fund creation date at the time of the first investment for the 345 funds that have an investment within the 24 months prior to the computed creation date, and I drop 12 funds that have investments before that.

[4] "Evergreen" usually describes funds with an unbounded investment life.

[5] Although the final sample is dominated by funds labeled by VentureXpert as Venture Capital, it also includes a small number of Buyout and Other funds. Results in untabulated regressions show that the results are identical when I keep all funds or when I restrict the analysis to funds labeled as Venture Capital only.

I check that 50%, 70% and 99% of investments in the sample occur respectively within 2, 3 and 10 years following fund creation.

I am left with 1,527 limited partnerships, of which 1,328 labeled as "Venture Capital", 134 as "Buyout", and 64 others. I consider only the first investment of each fund in each company in the sample. In what follows, I call an "investment" or "deal" the initial investment of a distinct fund in a distinct company. Hence if there are two funds investing in the same company at the same date, this counts as two investments or two deals. When a funds makes several sequential investments in a given company, I only consider the first one.

I am left with 27,155 investments of distinct funds into 15,316 distinct companies. For some of the analysis, however, I only consider companies for which SDC provides a SIC code and only investments up to December 2006. This leaves me with 19,424 investments of 1,301 distinct funds in 10,077 distinct companies.

SDC's investment database has a companion database of private equity backed initial public offerings (IPOs) and merger acquisitions (M&As) which relates any such event to the names of the funds backing the company. I match my sample with this database on fund names to identify the timing of exits[6]. I am left with 8,609 exits of 1,258 distinct funds from 4,419 distinct companies, of which 2,721 IPOs and 5,888 M&As.

I collapse SIC codes in 30 industries using Fama-French 30 industry classification and correspondence tables with SIC codes. Kenneth French's website also provides industry monthly returns and Fama-French factors including momentum. In regressions using subsequent excess returns in the 12, 24, 36 and 48 months following any investment in the sample, the sample is restricted to investments up to December 2006 for which SDC provides a SIC code.

For a subset of companies in the main sample, I obtain patenting information from the NBER patent database and the HBS patent database (Lai et al., 2009) which together cover

[6] A little less than 20% of IPOs dates are anterior to the investment date in the main sample, I drop those. This occurs in less than 2% of cases in the M&A sub-sample. According to SDC, the funds associated with the companies in their IPO database could have invested in the company after the IPO and nonetheless appear in the sample.

US patents granted through December 2010. I merge it with my sample on company name and city. I then follow the procedure recommended by Hall et al. (2001) and applied in Lerner et al. (2011) to adjust patents and citations for the truncation bias. I restrict this sample to private equity investments occurring up to 2006. I only keep patents applied in the three years before and the five years after the investment of any given fund in any given firm. I am left with 6,755 investments by 1,104 distinct funds in 3,241 distinct companies, which file a total of 37,624 patents in the eight years around the investment year.

The great advantage of SDC over other private equity data providers is that it relates investments and companies to private equity funds rather than private equity firms. Although the representativeness of the main sample seems satisfactory, some data related concerns remain. I consider investment level data and do not use fund level performance data. Recently, Stucke (2011) and Harris et al. (2012) have established that Venture Economics performance data suffers from severe sample selection issues, with the coverage dropping sharply in the early 2000s. Since Venture Economics is a unit of SDC, one might be worried that the reporting bias also applies to the investment level data used in this study. To my knowledge, the answer to this question is still unclear[7]. However, two recent studies have assessed the ability of VentureXpert to accurately report deal level data. Kaplan et al. (2002) examine 143 financing rounds in 98 companies from 1986 to 1999. They argue that VentureXpert and VentureSource, another mainstream venture capital database, both exclude 15% of financing rounds and that the former oversamples larger rounds and California companies. Maats et al. (2011) examine investments made by a sample of 40 Venture Capital funds raised between 1992 and 2003 and compare the quality of the coverage of these investments by VentureXpert and VentureSource. They find that the consistency between both databases is low, but that the reliability of fund coverage is higher in VentureXpert, which should be the preferred source for collecting data at the fund level. They note however that fund coverage increases

[7]I address this potential concern by comparing the effect of horizon on funds behavior within a given vintage, by using vintage fixed effects. In addition, I am in the process of comparing the coverage of VentureXpert and VentureSource through time.

with the number of portfolio companies in a given fund. Overall, I am confident that the empirical strategy used in this paper addresses most of these concerns.

Measuring investor horizon Since virtually all private equity funds have a ten year finite horizon, I identify between-fund as well as within-fund variations in investment horizon by using the age of the fund at the time of the investment, measured as the log of the number of years between the creation of the fund and any given investment. Funds contractual agreements usually allow GPs to extend the fund's duration after ten years for up to three years in one year increments, with the consent of the LPs[8]. An extension of the fund's life enables GPs to liquidate stale investments at a profit instead of having to fire sell them. There is little room for LPs and GPs to extend the fund's duration *beyond* these contractual extensions. Gompers and Lerner (2001) note that *"unlike other agreements (for example, employment contracts or strategic alliances), these contracts are rarely renegotiated"*. Moreover, conversations with practitioners indicate that LPs are unlikely to agree to receive in-kind distributions of shares of unliquidated private companies. Instead, they will trust the private equity firm with a liquidation mandate and stop paying (or demand a cut in) management fees. This suggests that the contractual lifespan of the fund is indeed a binding constraint on GPs investment horizon.

Measuring innovation I follow Lerner et al. (2011) and Bernstein (2011) to measure innovation from the NBER and HBS patent databases. I first compute the number of patents per year which any company in the sample applies for in the three years before and the five years following any given private equity investment. Then for each patent, I count the number of times the patent has been cited by other patents in the calendar year of the patent grant and the three subsequent years. The innovation literature usually interprets the number of citations as a measures of the quality, or economic importance, of the patent.

[8] I analyzed a series of 24 hand collected private equity fund prospectus. They allow for an average of two extensions of one year. A majority of them require the consent of a majority of LPs for an extension to be granted.

The propensity to patent and to cite previously issued patents varies over time and across technologies. Moreover, towards the end of the sample, patent counts under report the actual patenting since many patents, although applied for, might not have been granted. I follow Hall et al. (2001) and compute scaled patents by dividing each patent by the average number of patents of all companies in the same year and technology class. Similarly, I compute scaled citations as the number of citations a patent receives divided by the average number of citations received by all patents granted in the same year and technology class.

Measuring sector excess returns I use industry returns computed from publicly listed companies to measure the excess returns of sectors invested by private equity funds. There are probably important differences between public and private firms within a given sector. However, I believe that using industry returns is reasonable for the purpose of this study since private equity funds rely on the stock market to exit their investments. As a matter of fact, the pricing of private companies in private equity transactions is significantly correlated with the pricing of listed companies in the same sector. In the largest hand collected sample of buyout deals with available pricing data, Axelson et al. (2010) find that *"there is a statistically significant, positive relationship between pricing and the prices of comparable public companies"*. More surprisingly, perhaps, Gompers and Lerner (2000) hand collect a sample of prices of Venture Capital deals and show that *"shifts in public market values appear to affect all transactions equally, regardless of stage or regions. This supports the suggestion that the industry public market index measure the expected future profitability of the industry and hence affect the prospects of all firms"*. Second, the sector exposure of a private equity investment impacts its overall performance. In the case of buyouts, Guo et al. (2011) and Acharya et al. (2010) find that it accounts for approximately one quarter to one third of investment level performance.

For each of the 30 Fama-French sectors, I first compute the average monthly sector return in excess of the market returns in the 12, 24, 36 and 48 subsequent months. Market returns are defined by Fama-French as the value-weighted returns on all NYSE, AMEX, and NASDAQ stocks (from CRSP). I compute the exposure of sector returns to each risk factor by running a five factor model (market, small-minus-large, high-minus-low, momentum, liquidity) model in the 60 months prior to any given month for each sector. I use the coefficients obtained in these regressions to compute the average monthly risk adjusted sector returns in the 12, 24, 36 and 48 subsequent months.

Summary statistics Table 2.1 to 2.5 present the summary statistics for the main analysis in the paper. I use the terms "investment" or "deal" to describe the initial investment of a distinct fund in a distinct company. If two funds invest in the same company at the same date, this counts as two investments or two deals. When a fund makes sequential investments in a given company, I keep only the first one.

Table 2.1 presents the distribution of fund creations, investments, club-deals[9] and exits through time. As expected, 1988-1990, 1998-2001 and 2005-2008 are the most active private equity periods. Fundraisings, investments and exits all increase sharply. Table 2.2 presents the distribution of investments across the 30 Fama-French sectors. Investments are more concentrated in the Business services, Health-care and Business equipments sectors, which account for respectively 38%, 18% and 14% of total investments. Panel A of table 2.3 presents fund level summary statistics. Funds invest on average in eighteen different companies in five different sectors. They are in their third year on average when they invest. Panel B presents statistics at the investment level, for dependent and explanatory variables used in the main regression analysis. Table 2.4 shows the distribution of private equity funds investments throughout their lives. Two thirds of investments occur within the three first years of the life of private equity funds. Half of the exits in the sample occur within the six first years, with IPOs occurring usually sooner than M&As. Table 2.5 presents the sub-sample for which patenting information is available. Companies apply for 1.2 patents per year on average in the three years before and the five years following any private equity investment. These

[9]Club-deals are simultaneous investments by different funds in the same company.

patents receive on average 8.9 citations in the year each patent was granted and in the three following years.

2.4 Results

Univariate tests Before turning to regression analysis, I compare the characteristics of companies receiving investments from funds close or far away from the end of their investment horizon.

Table 2.6 presents the mean and difference in means of characteristics of companies receiving an investment by funds within their first three years (18,677 investments) and funds beyond their third year (8,512 investments). Company age is the number of years between the creation of the company and the investment. The investment sequence number is the number of previous financing rounds (involving other funds) received by the company until the investment by the fund. The development stage of a company is measured with a dummy equal to zero for companies classified by VentureXperts as "Startup/Seed" or "Early Stage" and one for later stages. The number of rounds counts the number of financing rounds involving the fund following its first investment in the company. The holding period is the number of months between the investment and the exit through an IPO or a M&A deal. Standard errors are presented in parenthesis. *** indicates that the difference in means is significant at the 1% level.

It turns out that older funds invest in companies which are 3.5 year older on average and have received 0.5 more rounds of investments before that. The probability of younger funds to invest in "Startup/Seed" and "Early Stage" companies is smaller by 10%. Finally, younger funds do 0.3 more subsequent financing rounds and hold on to their investments for 5.4 additional months on average. Altogether, these results suggest that younger and older funds select projects with different characteristics. So far, however, these observations can not be tied down to the predictions of the theoretical framework above.

The way the sample is constructed could indeed generate these results mechanically. First, note that the sample is truncated towards the end of the sample (the last investments of the latest vintages are not observed). Suppose that for some reasons the general investment style of private equity funds changed through time and that the latest fund vintages specialized in very exploratory investments. I would find that younger funds invest on average in more exploratory projects. I address this issue in the OLS regressions of the following section by using fund and fund vintage fixed effects. Second, suppose that private equity firms differ systematically with respect to their investment styles, with some of them having higher skills at detecting and investing in exploratory projects early following their fundraising. The difference in means observed above might simply reflect the difference between skilled and unskilled private equity investors. I rule this channel out by adding private equity firm fixed effects in the regression analysis below. Finally, if new business creations are cyclical and if most fundraisings occur at the height of those cycles, then the univariate tests in table 2.6 might simply reflect this correlation. I shut this channel down by using year fixed effects in the regressions of the upcoming section.

Moreover, even if funds indeed systematically shift towards less mature projects as they get closer to liquidation, other features of private equity funds (rather than their limited lifespan only) may account for this fact . The first one is risk aversion. GPs are compensated based on an annual 2% of commitments and 20% of the overall performance of the fund above a hurdle rate (usually around 8%), the carried interest. Hence, at any given point in a given fund's investment life, if cumulative performance has been lower than the hurdle rate, the value of the carried interest is zero. Suppose that a fund manager can choose either a high risk and high return project and a low risk and low return project. It is clear that a lower level of past performance shifts the manager's preference towards the high risk and high returns project, since it is more likely to bring her carried interest back in the money in case of success and since it won't change the carried interest value in case of failure. I use the number of past exits as a proxy for past performance, a measure widely used in the literature.

Private equity firms raise new funds every three to five years. As has been evidenced by Gompers (1996) and more recently Chung et al. (2012) this is likely to dramatically influence manager's behavior, especially young ones that have not yet established reputation and that potentially face more difficulty in raising funds. I address this concern by including in all regressions a dummy for first time funds and the number of funds simultaneously run by the same private equity firm. If fundraisings have an influence on the change in investment behavior towards the end of funds investment life, then these variables should capture it.

Since funds have limited resources, they are likely to pick less complicated assets once they already have a number of other investments to manage, which is likely to happen towards the end of their investment life. Suppose that innovative projects are more costly to monitor. Then I might observe that as a fund gets closer to the end of its investment life, it invests in less innovative assets because it already devotes all its available resources to monitor its existing investments. I control for this with the number of investments that the fund has made since its creation.

Fund horizon and company's maturity I turn to a more formal test of *Prediction 1.* I use OLS regressions to show that as a fund gets closer to liquidation, it selects less mature companies. I analyze the maturity of a company along three dimensions: its age, its development stage at the time of the investment, and the number of financing rounds (involving other funds) it received prior to the investment.

I estimate the following OLS specification at the investment level:

$$V_{j,t} = \alpha + \lambda_1 Age_{i,t} + \lambda_2 X_{i,t} + \gamma_i + \mu_t + \epsilon_{i,j,t}$$

$V_{j,t}$ is the variable of interest at the time t of the investment of fund i in company j, $Age_{i,t}$ is the log of fund age, $X_{i,t}$ is a vector of fund level control and γ_i and μ_t are fund and time fixed effects.

I first consider how funds shift their investments towards younger companies as their horizon shrinks. To do so I run an investment level OLS regression of the log of company's age on the log of fund age, the log of the average age of companies in which the fund previously invested, and a variety of fund level controls including (i) the log number of past exits, (ii) the log number of past investments, (iii) the log number of funds operated by the same private equity firm, (iv) the log of fund size, (v) a dummy for first time funds and (vi) a dummy for "VC funds" and "Other funds". Several specifications are run with fund vintage, year, private equity firm and fund fixed effects. Standard errors are clustered by month. As evidenced in table 2.7, funds with shorter horizon invest in younger companies. A fund in its fourth year of activity invests in companies that are 1.2 to 2.9 months older than companies which receive funding from a fund in its first year of activity.

The age of the company might fail to account precisely for its development stage. Companies might have been founded for a few years and yet still be at a very early stage of their development. I measure the development stage of companies in the sample with a dummy equal to zero for companies classified by VentureXperts as "Startup/Seed" or "Early Stage" and one for later stages. I then run an investment level OLS regression of the development stage dummy on the log of fund age, the average of the development stage dummy in companies in which the fund invested prior to time t, and a variety of fund level controls including (i) the log number of past exits, (ii) the log number of past investments, (iii) the log number of funds operated by the same private equity firm, (iv) the log of fund size, (v) a dummy for first time funds and (vi) a dummy for "VC funds" and "Other funds". Several specifications are run with fund vintage, year, private equity firm and fund fixed effects. Standard errors are clustered by month. As expected, older funds invest in companies at a later stage of their development across specifications presented in table 2.8. A fund in its fourth year of activity has a 4 to 5% larger probability than a fund in its first year of activity to invest in companies that are beyond the startup or early stage of their development.

Finally, I approach the maturity of a company with its investment sequence number, i.e. the number of financing rounds (involving other funds) it already received until the investment of the fund. I run an investment level OLS regression of the log investment sequence number on the log of fund age, the log average investment sequence number of the fund's investments prior to time t, and a variety of fund level controls including (i) the log number of past exits, (ii) the log number of past investments, (iii) the log number of funds operated by the same private equity firm, (iv) the log of fund size, (v) a dummy for first time funds and (vi) a dummy for "VC funds" and "Other funds". Several specifications are run with fund vintage, year, private equity firm and fund fixed effects. Standard errors are clustered by month. As expected, older funds invest in companies at a later stage of their development across specifications presented in table 2.9. A company receiving an investment by fund in its fourth year of activity has had 0.06 to 0.11 more previous rounds of financing than a company receiving an investment by a fund in its first year of activity.

Interestingly, the level of past performance has a negative effect on the propensity of funds to select exploratory projects. In virtually all regressions, the coefficient on the number of past exits is statistically significant. This is consistent with the idea that funds make more exploratory projects when their risk incentives increase (when their carried interest is out of the money). Moreover, the coefficient on the interaction between the first time fund dummy and the number of past exits is of the opposite sign. This is consistent with Chung et al. (2012) and other related papers, which argue that most of the performance of first time funds should be related to future fund flows while the performance of established funds should be related to the carried interest on their current fund. Therefore, the behavior of established funds should be sensitive to the value of the carried interest: when performance has been low, their incentives to take on more risk increase, which might lead them to undertake more exploratory projects. Note however that these effects do not subsume the effect of fund horizon.

Fund horizon and investment style I now ask whether shorter horizon funds select companies in which they have to split funding into more investment rounds through time (*Prediction 2*). Staging investment is more relevant for the funding of startups involved in the exploration of new ideas rather than the exploitation of existing ones. I run an investment level OLS regression of the log number of rounds on the log of fund age, the size of the investment as a percentage of fund size, the log average number of rounds in past investments, and a variety of fund level controls including (i) the log number of past exits, (ii) the log number of past investments, (iii) the log number of funds operated by the same private equity firm, (iv) the log of fund size, (v) a dummy for first time funds and (vi) a dummy for "VC funds" and "Other funds". Several specifications are run with successively vintage, year, private equity firm and fund fixed effects. Standard errors are clustered by month. As evidenced in table 2.10, funds with a shorter investment horizon stage their investment less, controlling for the total amount they invest in the company across initial and subsequent rounds. A fund in its fourth year of operations stages its investments by 5 to 9% less than a fund in its first year.

I finally check whether shorter horizon funds hold on to their investments for a shorter period of time. This is what should be observed if funds getting closer to liquidation indeed select less exploratory projects which they can liquidate earlier. I run an investment level OLS regression of the log holding period (the number of months between the investment and the exit) on the log of fund age, the log average holding period of the fund's investments prior to time t, and a variety of fund level controls including (i) the log number of past exits, (ii) the log number of past investments, (iii) the log number of funds operated by the same private equity firm, (iv) the log of fund size, (v) a dummy for first time funds and (vi) a dummy for "VC funds" and "Other funds". Several specifications are run with successively vintage, year, private equity firm and fund fixed effects. Standard errors are clustered by month. Table 2.11 shows that on average funds with a longer investment horizon hold on to the companies they invest in for a longer period of time. Funds in their first year of operations hold on to their investments for 1 to 6% longer than funds in their fourth year of operations.

Fund horizon and innovation In this paragraph, I check whether companies which receive investments by younger funds are more innovative, as measured by the increase in the number of patents they issue each year and the number of citations their patents receive (*Prediction 3*).

I start by providing graphic evidence that funds further away from liquidation invest in companies that increase the number of patents they issue annually more than companies in which funds closer from liquidation invest. To do so, I keep any investments in the sample up to December 2006. I am left with 6,755 investments of 1,104 funds in 5,174 companies. I split the sample in two sub-samples of investments involving old and young funds. An investment is allocated to the young fund sample if it happens within the first 36 months of the life of the fund. It is allocated to the old fund sample otherwise. In each sub-sample, I compute the average number of patent applications in the three years prior and the four years following the investment. The results are presented in figure 2.1. Companies with long horizon investors issue 0.5 more patent and 0.2 more scaled patents per year following the investment.

I then consider the three years before and the five years after the year of the investment of any fund in any company in the sample up to December 2006 and run the following company$\times$year OLS regression:

$$\begin{aligned} PC_{j,t+k} &= \alpha_0 + \alpha_1 Age_{i,t} + \alpha_2 F_{i,t} + \alpha_3 C_{j,t} + \sum_{k=-3}^{5} \lambda_k Y_{t+k} + \sum_{k=-3}^{5} \beta_k Y_{t+k} \times Age_{i,t} \\ &+ \sum_{k=-3}^{5} \delta_k Y_{t+k} \times F_{i,t} + \sum_{k=-3}^{5} \theta_k Y_{t+k} \times C_{j,t} + \epsilon_{i,j,t} \end{aligned}$$

$PC_{j,t+k}$ is the log of one plus the number of patent applications by company j in year k around the investment year t. Y_{t+k} is a dummy equal to one in the k^{th} year around the investment of fund i in company j which occurs in year t. $Age_{i,t}$ is the log of the age of fund i at the time of the investment. $F_{i,t}$ is a vector of fund level controls including (i) the log number of past exits, (ii) the log number of past investments, (iii) the log number of funds operated by the same private equity firm, (iv) the log of fund size, (v) a dummy for first time funds and (vi) a dummy for "VC funds" and "Other funds". $C_{j,t}$ is a vector of company level controls, including the log of company age, state and sector dummies. Standard errors are clustered at the company level. Panel A of table 2.12 presents the results of the specifications using company fixed effects, while Panel B includes company level controls for age, sector and state of incorporation. As expected, the results presented in table 2.12 show that companies invested by private equity funds further away from liquidation increase their patenting activity more in the five years following the deal. Companies which receive funding from a fund in its first year of operations increase their patent count by 0.07 to 0.17 more patents than companies which receive funding from a fund in its fourth year of operations. When controlling for observable company characteristics in panel B, the difference in patenting *before* the investment is not significant. However, the difference remains significant following the investment.

I check whether the relatively stronger increase in patenting activity in companies with long term funding is not achieved at the cost of the quality of patents by studying the change in citation count per patent. I consider the three years before and the five years after the year of the investment of any fund in any company in the sample up to December 2006 and run the same regression as above, although at the patent level rather than the company×year level. Results in table 2.13 provide evidence that patent quality increases more in companies which have received investment from funds further away from liquidation. Following the investment, companies which receive funding from a fund in its first year of operations apply to patents that receive 0.06 to 0.18 more citations than patents applied by companies which receive funding from a fund in its fourth year of operations.

Horizon and sector exposure In this paragraph, I turn to the test of *Prediction 4* regarding the sector exposure of private equity funds. I check that funds with a longer investment horizon pick sectors that experience subsequent abnormal returns. I first provide graphic evidence that funds further away from liquidation (young funds) invest in sectors that experience larger excess returns than funds closer from liquidation (old funds). To do so, I split the main sample in two sub-samples of investments involving old funds and young funds. An investment is allocated to the young fund sample if it happens within the first 36 months of the life of the fund. It is allocated to the old fund sample otherwise. In each of eight years around the investment, I compute the annual excess returns of the sector to which the company belongs over the market returns. I then average the yearly excess returns in each sample, take the difference between excess returns in the sub-sample of young funds and the sub-sample of old funds, compute 95% confidence intervals of these differences and plot them in figure 2.2. On average, after (before) their investments, sectors targeted by young funds outperform (underperform) sectors picked by old funds.

The main competing explanation of the prediction that funds with a longer investment horizon invests in sectors that experience higher risk adjusted returns relies on private equity firms *expertise* and ability to strategically raise new funds when sectors they know the best are relatively worth investing in, in the spirit of Gompers et al. (2008) and Gompers et al. (2009). According to this explanation, sectors targeted early in a fund's life should experience larger subsequent returns than sectors targeted later on. One way to distinguish this explanation from the former is to test whether the difference between early and late investments for a given fund decreases when there are other funds investing at the same time. In the *expertise* story, the sensitivity of sector exposure to fund age should not be affected by the presence of other funds contemporaneously investing in the same sector. In the *exploration* story, an older fund is more likely to invest in an unexplored sector when there are more other private equity funds investing simultaneously in the same sector, as this signals that uncertainty about the sector is being resolved or that the market will have more appetite for the investment when it needs to be liquidated.

I now turn to an OLS regression setting and provide evidence that investment horizon acts as a limit on the propensity of private equity funds to invest in sectors which experience subsequent excess returns. I show that the relation between fund age and sector excess returns is weaker when there are more private equity investments around. I compute the industry subsequent excess returns in the 12, 24, 36 and 48 months following any investment in the sample, hence the sample is restricted to deals up to December 2006.

The following regression is run at the investment level:

$$ARet_{j,t+k} = \eta Age_{i,t} + \lambda Age_{i,t} \times SHARE_{j,t} + \beta SHARE_{j,t} + \delta X_{i,t} + \alpha_i + \epsilon_{i,j,t}$$

$ARet_{j,t+k}$ is the average five factor adjusted monthly return of the sector of company j in the k months following the investment by fund i in month t. $Age_{i,t}$ is the log of the age of fund i. $SHARE_{j,t}$ is the share of total investments by private equity funds in sector j in month t. $X_{i,t}$ is a vector of time varying characteristics of fund i in month t including (i) the log number of exits, (ii) the number of investments involving fund i up to month t, (iii) the number of funds of the same class as fund i contemporaneously managed by the private equity firm managing fund i. α_i is a fund fixed effect. $\epsilon_{i,j,t}$ are residuals. Standard errors are corrected for month clustering. Results are presented in table 2.14. As expected, the coefficient on the interaction term is positive and significant at all horizons and the coefficient on the log of fund age holds at the 2, 3 and 4 year horizon after controlling for time varying fund characteristics.

I complement this analysis by building a monthly calendar time portfolio long in sectors invested by funds far away from liquidation and short in sectors invested by funds closer to liquidation. The test borrows from Mitchell and Stafford (2000) and uses sector returns provided by Fama-French. I build a series of monthly calendar time portfolios as follows. Each month I split investments in tertiles of age of the funds involved. I define young fund investments as investments of funds in the first tertile and old fund investments as investments of funds in the third tertile of age. I exclude sectors which are both invested by young and old funds. I first build a monthly calendar time portfolio of sectors picked by young private equity funds. Each month, I average the returns of sectors that have received investments by young private equity funds in the previous 12 months. I do the same at the 24, 36 and 48 months frequency. I build similar monthly calendar time portfolios of sectors which have received investments by old funds. I then compute the monthly raw returns, CAPM alpha and five factor alpha generated by each of these portfolios. Finally, I build a portfolio long in sectors invested by young funds and short in sectors invested by old funds and compute monthly excess returns of this portfolio using the CAPM and five factor model. Results in table 2.15 show that the monthly rebalanced portfolio long in the exposure of young funds and short in the exposure of old funds generates yearly excess returns of the order of 96 basis points.

2.5 Conclusion

Investments exploring new ideas typically take more time to payoff than investments exploiting existing ones. I check that, consistent with this idea, investors with a longer horizon have a larger propensity to fund innovation than investors with a shorter horizon. I consider the case of private equity funds, which investment horizon is fixed ex ante.

I show that funds further away from liquidation invest in younger companies at an earlier stage of their development. Moreover, funds with a longer horizon split their investment more through time and hold on to their investments for a longer period of time. Overall, with respect to a fund in its first year of activity, a fund in its fourth year of activity (i) invests in companies that are 1.2 to 2.9 months older, (ii) has a 4 to 5% larger probability to invest in companies that are beyond the seed or early stage of their development, (iii) selects companies which have already received 0.06 to 0.11 more rounds of financing in the past, (iv) stages its investments by 5 to 9% less and holds on to its investments for 1 to 6% longer. Companies which receive funding from funds with a longer horizon increase their patent stock by 0.07 to 0.17 more patents following the investment than companies which receive funding from short horizon investors. Moreover, patents applied by companies which receive funding from funds with a longer horizon receive 0.06 to 0.18 more citations. Finally, I compare the sector exposure of funds with short and long horizon. I show that funds with a short horizon invest in industries which experience lower subsequent excess returns. A monthly rebalanced portfolio long in sectors invested by long horizon funds and short in sectors invested by short horizon funds yields positive and significant risk adjusted returns of the order of 96 basis points per year at the three and four year horizon. I interpret these results as evidence that long term investors have a larger propensity to target sectors with mid-term rather than immediate prospects.

Altogether, these results suggest that horizon is a strong driver of private equity funds asset allocation. Private equity funds collect a substantial share of global savings, which they channel to the real economy. Among other things, they play an important role in supporting and fueling corporate innovation. Improving our understanding of the behavior of private equity fund managers in response to their contractual incentives is thus a promising avenue for future research.

Bibliography

Acharya, Viral V., Moritz Hahn, and Conor Kehoe, "Corporate Governance and Value Creation: Evidence from Private Equity," *CEPR Discussion Papers*, March 2010, (7242).

Axelson, Ulf, Per Strömberg, and Michael S. Weisbach, "Why Are Buyouts Levered? The Financial Structure of Private Equity Funds," *The Journal of Finance*, 2009, *64* (4), 1549–1582.

_ , Tim Jenkinson, Per Strömberg, and Michael S Weisbach, "Borrow Cheap, Buy High? The Determinants of Leverage and Pricing in Buyouts," *National Bureau of Economic Research Working Paper Series*, 2010, *No. 15952.*

Bergemann, Dirk and Ulrich Hege, "Venture capital financing, moral hazard, and learning," *Journal of Banking & Finance*, August 1998, *22* (6-8), 703–735.

Bernstein, Shai, "Does Going Public Affect Innovation?," *Working Paper*, 2011.

Bushee, Brian J, "The influence of institutional investors on myopic R&D investment behavior," *Accounting Review*, 1998, *73* (3), 305–333.

Cella, Cristina, Andrew Ellul, and Mariassunta Giannetti, "Investors' horizons and the Amplification of Market Shocks," *CEPR Discussion Papers*, 2010, (8083).

Cheffins, Brian and John Armour, "The Eclipse of Private Equity," *ESRC Centre for Business Research*, March 2007, (wp339).

Chemmanur, Thomas J, Elena Loutskina, and Xuan Tian, "Corporate Venture Capital, Value Creation, and Innovation," *Working Paper*, 2011.

Chung, Ji-Woong, Berk A. Sensoy, Léa Stern, and Michael S. Weisbach, "Pay for Performance from Future Fund Flows: The Case of Private Equity," *Review of Financial Studies*, 2012.

Derrien, François, Andrus Kecskés, and David Thesmar, "Investor Horizons and Corporate Policies," *Working Paper*, 2011.

Ferreira, Daniel, Gustavo Manso, and André C. Silva, "Incentives to Innovate and the Decision to Go Public or Private," *Review of Financial Studies*, 2012.

Gompers, Paul, "Optimal Investment, Monitoring, and the Staging of Venture Capital," *Journal of Finance*, 1995, *50* (5), 1461–1489.

_ , "Grandstanding in the venture capital industry," *Journal of Financial Economics*, 1996, *42* (1), 133–156.

_ **and Josh Lerner**, "The Use of Covenants: An Empirical Analysis of Venture Partnership Agreements," *Journal of Law and Economics*, 1996, *39* (2), 463–98.

_ **and** _ , "An analysis of compensation in the US venture capital partnership," *Journal of Financial Economics*, 1999, *51* (1), 3–44.

_ **and** _ , "Money chasing deals? The impact of fund inflows on private equity valuation," *Journal of Financial Economics*, 2000, *55* (2), 281–325.

_ **and** _ , "The Venture Capital Revolution," *Journal of Economic Perspectives*, 2001, *15* (2), 145–168.

_ **, Anna Kovner, and Josh Lerner**, "Specialization and Success: Evidence from Venture Capital," *Journal of Economics Management Strategy*, 2009, *18* (3), 817–844.

_ **,** _ **,** _ **, and David Scharfstein**, "Venture capital investment cycles: The impact of public markets," *Journal of Financial Economics*, 2008, *87* (1), 1–23.

Guo, Shourun, Edith S Hotchkiss, and Weihong Song, "Do Buyouts (Still) Create Value?," *Journal of Finance*, 2011, *66* (2), 479–517.

Hall, Bronwyn H, Adam B Jaffe, and Manuel Trajtenberg, "The NBER patent citation data file: Lessons, insights and methodological tools," *NBER Working Paper*, 2001, *2001* (8498), 1–77.

Harris, Lee, "A Critical Theory of Private Equity," *Delaware Journal of Corporate Law*, 2010, *35* (1), 1–35.

Harris, Robert S, Tim Jenkinson, and Steven N Kaplan, "Private Equity Performance: What Do We Know?," *Working Paper*, 2012.

Kandel, Eugene, Dima Leshchinskii, and Harry Yuklea, "VC Funds: Aging Brings Myopia," *Journal of Financial and Quantitative Analysis*, April 2011, *46* (02), 431–457.

Kaplan, Steven N., Berk A. Sensoy, and Per Strömberg, "How well do venture capital databases reflect actual investments?," *Working Paper*, 2002.

Lai, Ronald, Alexander D'Amour, and Lee Fleming, "The careers and co-authorship networks of U.S. patent-holders, since 1975," *Working Paper*, 2009.

Lerner, Josh and Antoinette Schoar, "The illiquidity puzzle: theory and evidence from private equity," *Journal of Financial Economics*, 2004, *72* (1), 3 – 40.

_ **and Julie Wulf**, "Innovation and Incentives: Evidence from Corporate R&D," *The Review of Economics and Statistics*, 2007, *89* (4), 634–644.

_ **, Morten Sorensen, and Per Strömberg**, "Private Equity and Long-Run Investment: The Case of Innovation," *The Journal of Finance*, 2011, *66* (2), 445–477.

Ljungqvist, Alexander, Matthew Richardson, and Daniel Wolfenzon, "The Investment Behavior of Buyout Funds: Theory and Evidence," *NBER Working Paper*, 2008, *No. 14180.*

Maats, Frederike, Andrew Metrick, Ayako Yasuda, Brian Hinkes, and Sofia Vershovski, "On the Consistency and Reliability of Venture Capital Databases," *Working Paper*, 2011.

Manso, Gustavo, "Motivating Innovation," *Journal of Finance*, 2011, *66* (5), 1823–1860.

Masulis, Ronald and Rajarishi Nahata, "Venture Capital Conflicts of Interest: Evidence from Acquisitions of Venture Backed Firms," *Journal of Financial and Quantitative Analysis*, 2009.

Masulis, Ronald W and Randall S Thomas, "Does Private Equity Create Wealth? The Effects of Private Equity and Derivatives on Corporate Governance," *University of Chicago Law Review*, 2009, *76* (1), 219–259.

Metrick, Andrew and Ayako Yasuda, "The Economics of Private Equity Funds," *Review of Financial Studies*, 2009, *23* (6), 1–57.

Mitchell, Mark L and Erik Stafford, "Managerial Decisions and Long-Term Stock Price Performance," *The Journal of Business*, 2000, *73* (3), 287–329.

Polk, Christopher and Paola Sapienza, "The stock market and corporate investment: a test of catering theory," *Review of Financial Studies*, 2009, *22* (1), 187–217.

Robinson, David T. and Berk A. Sensoy, "Cyclicality, Performance Measurement, and Cash Flow Liquidity in Private Equity," *NBER Working Paper*, September 2011, (17428).

Sahlman, William, "The structure and governance of venture capital organizations," *Journal of Financial Economics*, 1990, *27*, 473–521.

Strömberg, Per, "The new demography of private equity," *Globalization of Alternative Investments Working Papers*, 2008, *1*, 3–26.

Stucke, Rüdiger, "Updating History," *Working Paper*, 2011.

Tian, Xuan and Tracy Yue Wang, "Tolerance for Failure and Corporate Innovation," *Review of Financial Studies*, 2011.

Xu, Moqi, "Ceo Contract Horizon and Investment," *Working Paper*, 2011.

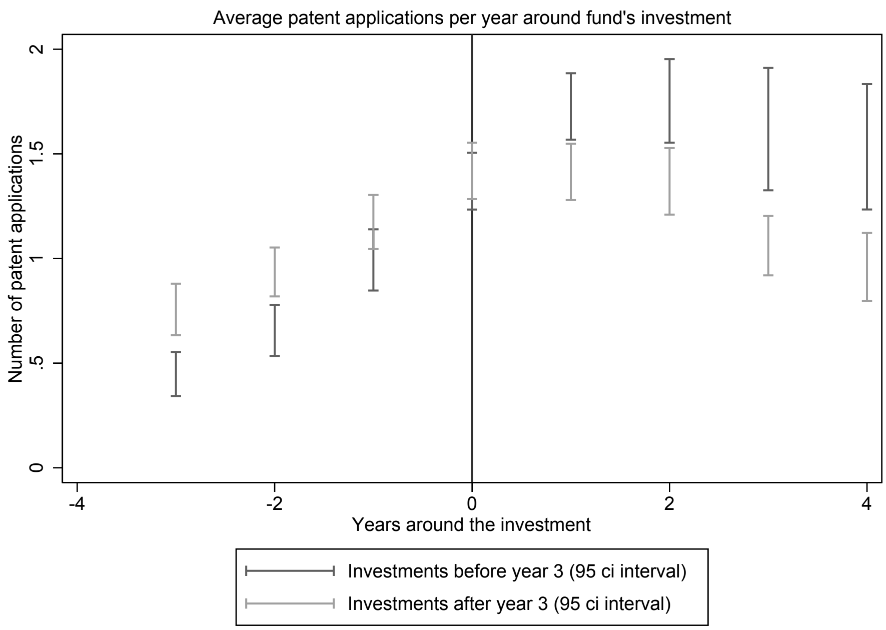
Appendix

Definition of the main variables

Log fund age	Log of one plus the number of years between the months of the investment and the month when the fund was created.
Log number of rounds	Log of one plus the number of investment rounds of the fund in the company following the initial investment of a fund in this company.
Investment sequence number	Number of financing rounds (involving other funds) received by the company until the investment of the fund.
Development stage dummy	Dummy equal to zero for companies classified by VentureXperts as "Startup/Seed" or "Early Stage" and one for later stages.
Log company age	Log of one plus the number of years between the month when the company was founded and the month of the initial investment of a fund in this company.
Log holding period	Log of the number of months between the month of the initial investment of a fund in the company and the initial public offering of the company or its sale to a third party (M&A).
Log number of past exits	Log of one plus the number of IPOs or M&As of companies which previously received an investment from the fund.
Log number of past investments	Log of one plus the number of previous investments made by the fund.
Log nb. of funds, same firm	Log of the number of funds of the same type (Venture capital, Buyout or Other funds) managed by the private equity firm at the time of the investment.
Size of investment	Ratio of the total dollar amount invested by the fund in the company to the size of the fund.
Log fund size	Log of the size of the fund measured in million dollars.
First time fund dummy	Dummy equal to 1 if the fund is the first of its type (Venture capital, Buyout or Other funds) raised by the private equity firm.
"VC fund" dummy	Dummy equal to 1 if the fund is labeled by VentureXpert as a Venture Capital fund.
"Other fund" dummy	Dummy equal to 1 if the fund is labeled by VentureXpert as an other fund.
Log patents	Log of one plus the number of patents applied by a company in a given year around a private equity investment.
Log scaled patents	Following Hall et al. (2001): log of one plus the number of patents applied by a company in a given year scaled by the average number of patents applied by all companies in the same year and technology class.
Log citations	Log of one plus the number of citations received by a patent in the year it was granted and in the three following calendar years.
Log scaled citations	Following Hall et al. (2001): log of one plus the number of citations a patent receives divided by the average number of citations received by all patents granted in the same year and technology class.
Share of sector in total private equity investments	Number of investments by funds of the same type (Venture Capital, Buyout, Other) in the same Fama-French 30 industry divided by the total number of investments by funds of the same type in a given month
Mean 48 (36,24,12) month five factor adjusted return	Mean monthly return of the unweighted Fama-French 30 industry index in the 48 (36,24,12) months following an investment, adjusted for the exposure to five risk factors (computed in the sixty month prior to the investment).

Figure 2.1: Fund horizon and patenting around private equity investments
I keep any investments in the sample up to December 2006. I am left with 6,755 investments of 1,104 funds in 3,241 companies. I split the sample in two sub-samples of investments involving old funds and young funds. An investment is allocated to the young fund sample if it happens within the first 36 months of the life of the fund. It is allocated to the old fund sample otherwise. In each sub-sample, I compute the average number of patent applications in the three years prior and the four years following the investment.

A. Patent count



B. Scaled patent count

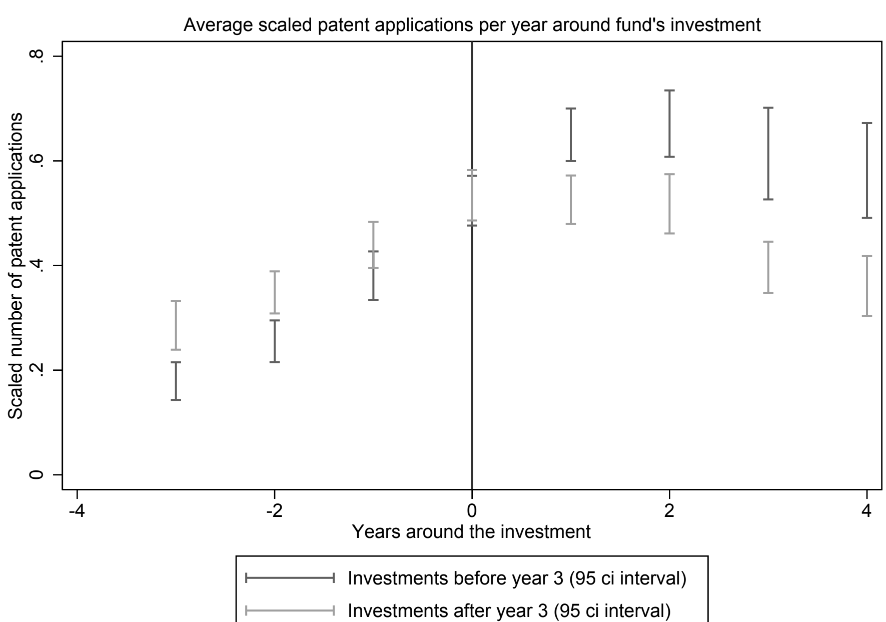


Figure 2.2: Fund age and sector excess returns around private equity investments
I use a sample of 19,424 investments of 1,301 distinct funds in 10,077 distinct companies up to December 2006 and split it in two sub-samples of investments involving old funds and young funds. An investment is allocated to the young fund sample if it happens within the first 36 months of the life of the fund. It is allocated to the old fund sample otherwise. Then, in each of eight years around the investment, I compute the past 12 month excess returns of the sector to which the company belongs over the market returns. I then average the yearly excess returns in each sample, take the difference between excess returns in the sub-sample of young funds and the sub-sample of old funds, compute the 95% confidence intervals of these differences and plot them.

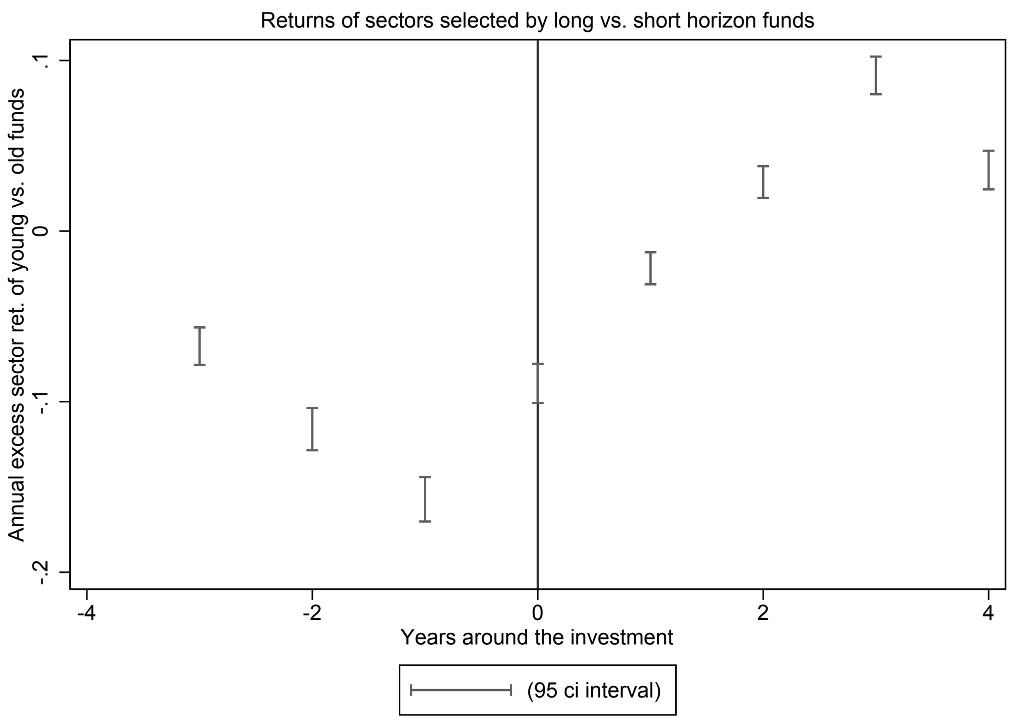


Table 2.1: Distribution of fund creations, investments, exits and and club-deals over time

This table presents the distribution of fund creations, investments, club-deals and exits across years. Club-deals are simultaneous investments by different funds in the same company.

Year	Nb. of new funds	Nb. of investments	Nb. of club-deals	Nb. of exits
1980	3	17	-	1
1981	5	44	-	2
1982	6	96	8	2
1983	14	221	30	20
1984	7	200	32	13
1985	12	208	35	10
1986	17	321	68	43
1987	8	299	51	32
1988	14	271	35	24
1989	20	274	35	23
1990	12	230	22	26
1991	4	208	21	53
1992	17	323	32	90
1993	29	319	51	99
1994	43	380	57	94
1995	38	498	66	137
1996	55	782	110	177
1997	97	1007	125	140
1998	128	1299	184	155
1999	149	2384	384	294
2000	212	3655	604	337
2001	107	2113	365	200
2002	52	1232	209	199
2003	54	1330	215	190
2004	78	1557	293	306
2005	87	1561	262	270
2006	93	1619	286	305
2007	80	1672	261	318
2008	56	1540	243	231
2009	17	781	105	192
2010	13	714	75	326

Table 2.2: Distribution of investments across sectors

This table presents the distribution of investments in the sample across Fama-French 30 sectors (classification obtained from Kenneth French's website).

Sector	Number of investments	% of deals
Food Products	134	0.49%
Beer & Liquor	5	0.02%
Recreation	205	0.75%
Printing and Publishing	167	0.61%
Consumer Goods	130	0.48%
Apparel	31	0.11%
Healthcare	4785	17.62%
Chemicals	134	0.49%
Textiles	15	0.06%
Construction	133	0.49%
Steel Works	87	0.32%
Fabricated Products	196	0.72%
Electrical Equipment	225	0.83%
Automobiles and Trucks	47	0.17%
Aircraft, ships	19	0.07%
Mines, Precious Metals	10	0.04%
Coal	10	0.04%
Oil	119	0.44%
Utilities	125	0.46%
Communication	1335	4.92%
Business Services	10271	37.82%
Business Equipment	3822	14.07%
Business Supplies	35	0.13%
Transportation	136	0.50%
Wholesale	360	1.33%
Retail	527	1.94%
Restaraunts, Hotels	80	0.29%
Finance	883	3.25%
Other	3130	11.53%

Table 2.3: Summary statistics

This table presents summary statistics for the sample used in most of the analysis. Panel A shows fund level statistics while panel B shows investment level statistics. In regression using subsequent excess returns in the 12, 24, 36 and 48 months following any investment in the sample, the sample is restricted to deals up to December 2006 for which SDC provides a SIC code. Returns are shown in percentages.

Variables	Obs.	Mean	Median	Std. dev.
Fund level summary statistics				
Number of investments (in distinct companies)	1527	17.81	14.00	14.53
Number of sectors	1527	4.90	4.00	4.90
Fund age	1527	3.11	2.82	1.50
Investment level summary statistics				
Dependent variables				
Log company age (in years)	20542	1.320	1.386	0.877
Log investment sequence number	25219	0.642	0.693	0.681
Stage dummy (=0 for seed and early stage)	25219	0.599	1.000	0.490
Log number of rounds	24222	0.596	0.693	0.630
Log holding period (in months)	7245	3.579	3.714	0.840
Mean 48 month five factor adjusted return	19351	0.502	0.430	0.718
Mean 36 month five factor adjusted return	19351	0.519	0.487	0.815
Mean 24 month five factor adjusted return	19351	0.462	0.434	0.987
Mean 12 month five factor adjusted return	19351	0.391	0.260	1.341
Explanatory variables				
Log fund age (in years)	25219	0.943	1.099	0.627
Log number of past exits	25219	0.457	0.000	0.679
Log number of past investments	25219	2.353	2.398	0.855
Log nb. of funds, same firm	25219	1.064	1.099	0.395
Size of investment (% of fund size)	24222	0.030	0.019	0.042
Log fund size	25219	5.243	5.220	1.153
First time fund dummy	25219	0.345	0.000	0.476
Share of sector in total private equity investments	19351	0.243	0.237	0.152

Table 2.4: Distribution of investments and exits throughout fund life

This table presents the distribution of investment and exits through initial public offerings of merger and acquisitions by fund age (in years).

Fund age	Investments			Exits (IPOs + M&As)			IPOs		
	Nb.	Percent.	Cum . percent.	Nb.	Percent.	Cum. percent.	Nb.	Percent.	Cum. percent.
1	6498	24%	24%	108	1%	1%	75	3%	3%
2	6948	26%	49%	406	5%	6%	212	8%	11%
3	5231	19%	69%	760	9%	15%	319	12%	22%
4	3284	12%	81%	981	11%	26%	378	14%	36%
5	2135	8%	89%	1009	12%	38%	311	11%	48%
6	1298	5%	93%	1090	13%	51%	325	12%	60%
7	660	2%	96%	1015	12%	62%	314	12%	71%
8	420	2%	97%	817	9%	72%	245	9%	80%
9	267	1%	98%	704	8%	80%	140	5%	85%
10	149	1%	99%	541	6%	86%	105	4%	89%
11	100	0%	99%	427	5%	91%	90	3%	92%
12	74	0%	100%	253	3%	94%	64	2%	95%

Table 2.5: Summary statistics: patenting firms

This table presents summary statistics for the sub-sample of firms matched with the NBER and HBS patent databases. Panel A displays the distribution of investments, patent applications and grants per year. Panel B show the distribution of patents (per year around the investment) and citations (per patent).

PANEL A: Distribution of investments, patent applications and patent grants per year

Year	Investments	Patent applications	Patent grants
1978	-	3	-
1979	-	3	-
1980	-	11	2
1981	7	31	5
1982	23	44	6
1983	31	54	14
1984	40	97	23
1985	53	133	38
1986	73	186	69
1987	66	200	104
1988	66	233	148
1989	46	299	215
1990	35	345	204
1991	58	363	233
1992	84	447	272
1993	93	484	307
1994	96	731	351
1995	131	1092	370
1996	190	968	510
1997	266	1318	714
1998	376	1829	865
1999	594	2453	954
2000	862	3381	1242
2001	718	3963	1585
2002	529	4103	1838
2003	589	3690	2513
2004	619	3963	2622
2005	536	2909	2487
2006	483	2049	3290
2007	-	1308	2929
2008	-	483	2658
2009	-	88	2910
2010	-	10	88

PANEL B: Mean patent and citation count

Variables	Obs.	Mean	Std. dev.
Patent applications per year	59976	1.21	6.69
Scaled patent applications per year	59976	0.46	2.07
Citations per patent	37271	8.90	15.47
Scaled citations per patent	37215	1.80	4.56

Table 2.6: Univariate tests

This table presents the mean and difference in means of characteristics of companies receiving an investment by funds within their first three years (18,677 investments) and funds beyond their third year (8,512 investments). Company age is the number of years between the creation of the company and the investment. Investment sequence number is the number of previous financing rounds (involving other funds) received by the company until the investment by a given fund. The development stage of a company is measured with a dummy equal to zero for companies classified by VentureXperts as "Startup/Seed" or "Early Stage" and one for later stages. The number of rounds counts the number of financing rounds involving the fund following its first investment in the company. The holding period is the number of months between the investment and the exit through an IPO or a M&A deal. Standard errors are presented in parenthesis. *** indicates that the difference in means is significant at the 1% level.

		Investments until year 3	Investments beyond year 3	Difference
Company age at investment (years)	Mean	1.47	4.95	-3.48***
	Std. dev.	(0.81)	(1.78)	
Investment sequence number	Mean	2.29	2.82	-0.53***
	Std. dev.	(1.91)	(2.30)	
Development stage dummy	Mean	0.57	0.67	-0.10***
	Std. dev.	(0.50)	(0.47)	
Number of rounds	Mean	2.29	2.01	0.29***
	Std. dev.	(1.75)	(1.54)	
Holding period (months)	Mean	48.84	43.49	5.35***
	Std. dev.	(34.50)	(29.83)	

Table 2.7: Fund horizon and company's age

This table presents the results of an investment level OLS regression of the log of company's age on the log of fund age, the log of the average age of companies in which the fund previously invested, a variety of fund level controls including (i) the log number of past exits, (ii) the log number of past investments, (iii) the log number of funds operated by the same private equity firm, (iv) the log of fund size, (v) a dummy for first time funds and (vi) a dummy for "VC funds" and "Other funds". Several specifications are run with fund vintage, year, private equity firm and fund fixed effects.Standard errors are clustered by month and presented in parenthesis. ***, ** and * indicate significance at the 1, 5 and 10 percent level.

Dependent variable: log company age (in years)						
Log fund age	0.25***	0.17***	0.28***	0.27***	0.16***	0.36***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Log nb. of exits	0.07***	0.09***	0.03**	0.05***	0.06***	0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
First time fund x Log nb. of exits	-0.05***	-0.06***	-0.04**	-0.05**	-0.05**	
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	
First time fund	0.02	0.03	-0.10***	-0.02	0.01	
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	
Log mean company age in past inv.	0.25***	0.25***	0.05***	0.04***	0.03*	-0.23***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Log nb. of past investments	-0.11***	-0.11***	-0.09***	-0.08***	-0.07***	-0.10***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Log nb. of funds, same PE firm	-0.03	-0.04*	-0.05*	-0.04*	-0.03	-0.07**
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
Log fund size	0.01	0.01	0.01	-0.03**	-0.03***	
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	
"VC fund" dummy	-0.17***	-0.18***	-0.06	-0.07	-0.07	
	(0.04)	(0.04)	(0.09)	(0.09)	(0.09)	
"Other fund" dummy	0.20***	0.17***	0.18	0.13	0.04	
	(0.06)	(0.06)	(0.13)	(0.13)	(0.13)	
Constant	1.14***	1.24***	1.29***	1.21***	1.47***	1.57***
	(0.07)	(0.07)	(0.11)	(0.15)	(0.30)	(0.05)
Vintage fixed effects	Yes	No	No	Yes	No	No
Inv. year fixed effects	No	Yes	No	No	Yes	No
PE firm fixed effects	No	No	Yes	Yes	Yes	No
Fund fixed effects	No	No	No	No	No	Yes
Observations	20,542	20,542	20,542	20,542	20,542	20,542
R-squared	0.10	0.10	0.16	0.17	0.17	0.23

Table 2.8: Fund horizon and company's development stage

This table presents the results of an investment level OLS regression of a development stage dummy on the log of fund age, the average of the development stage dummy in companies in which the fund invested prior to time t, a variety of fund level controls including (i) the log number of past exits, (ii) the log number of past investments, (iii) the log number of funds operated by the same private equity firm, (iv) the log of fund size, (v) a dummy for first time funds and (vi) a dummy for "VC funds" and "Other funds". The development stage dummy is equal to zero for companies classified by VentureXperts as "Startup/Seed" or "Early Stage" and one for later stages. Several specifications are run with fund vintage, year, private equity firm and fund fixed effects.Standard errors are clustered by month and presented in parenthesis. ***, ** and * indicate significance at the 1, 5 and 10 percent level.

	Dependent variable: Development stage dummy					
Log fund age	0.07***	0.07***	0.09***	0.09***	0.08***	0.09***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Log nb. of exits	0.04***	0.05***	0.04***	0.03***	0.04***	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
First time fund x Log nb. of exits	-0.02**	-0.03***	-0.04***	-0.03***	-0.04***	
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	
First time fund	0.01	0.01	0.00	0.00	0.01	
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	
Mean dev. stage of past inv.	0.40***	0.41***	0.13***	0.11***	0.12***	-0.34***
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Log nb. of past investments	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***	-0.02**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Log nb. of funds, same PE firm	0.02**	0.01	0.08***	0.05***	0.04***	0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Log fund size	0.02***	0.01***	0.00	-0.00	-0.01	
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	
"VC fund" dummy	-0.11***	-0.12***	-0.02	-0.03	-0.01	
	(0.01)	(0.01)	(0.03)	(0.03)	(0.03)	
"Other fund" dummy	0.01	-0.00	0.08	0.08	0.08	
	(0.02)	(0.02)	(0.05)	(0.05)	(0.05)	
Constant	0.36***	0.37***	0.41***	0.38***	0.57***	0.70***
	(0.03)	(0.03)	(0.04)	(0.06)	(0.12)	(0.03)
Vintage fixed effects	Yes	No	No	Yes	No	No
Inv. year fixed effects	No	Yes	No	No	Yes	No
PE firm fixed effects	No	No	Yes	Yes	Yes	No
Fund fixed effects	No	No	No	No	No	Yes
Observations	25,219	25,219	25,219	25,219	25,219	25,219
R-squared	0.09	0.09	0.14	0.14	0.14	0.20

Table 2.9: Fund horizon and company's investment sequence number

This table presents the results of an investment level OLS regression of the log investment sequence number on the log of fund age, the log average investment sequence number of companies in which the fund invested prior to time t and a variety of fund level controls including (i) the log number of past exits, (ii) the log number of past investments, (iii) the log number of funds operated by the same private equity firm, (iv) the log of fund size, (v) a dummy for first time funds and (vi) a dummy for "VC funds" and "Other funds". The investment sequence number is the number of previous financing rounds (involving other funds) received by the company until the investment of the fund. Several specifications are run with fund vintage, year, private equity firm and fund fixed effects.Standard errors are clustered by month and presented in parenthesis. ***, ** and * indicate significance at the 1, 5 and 10 percent level.

	Dependent variable: Log investment sequence number					
Log fund age	0.13***	0.09***	0.15***	0.16***	0.12***	0.17***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Log nb. of exits	0.05***	0.07***	0.05***	0.04***	0.06***	0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
First time fund x Log nb. of exits	0.02	0.00	-0.02	-0.02	-0.03**	
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	
First time fund	-0.01	-0.00	-0.04***	-0.02	-0.00	
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	
Log mean inv. sequence nb. of past inv.	0.13***	0.13***	0.05***	0.05***	0.04***	-0.09***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Log nb. of past investments	-0.04***	-0.04***	-0.05***	-0.05***	-0.05***	-0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Log nb. of funds, same PE firm	0.04***	0.02	0.11***	0.08***	0.05***	0.07***
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Log fund size	-0.01	-0.01	-0.01	-0.02***	-0.03***	
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	
"VC fund" dummy	0.11***	0.10***	0.06*	0.07*	0.05	
	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)	
"Other fund" dummy	0.07**	0.04	0.11*	0.12*	0.06	
	(0.03)	(0.03)	(0.06)	(0.07)	(0.06)	
Constant	0.17***	0.24***	0.36***	0.38***	0.62***	0.69***
	(0.04)	(0.04)	(0.06)	(0.08)	(0.21)	(0.04)
Vintage fixed effects	Yes	No	No	Yes	No	No
Inv. year fixed effects	No	Yes	No	No	Yes	No
PE firm fixed effects	No	No	Yes	Yes	Yes	No
Fund fixed effects	No	No	No	No	No	Yes
Observations	25,219	25,219	25,219	25,219	25,219	25,219
R-squared	0.08	0.08	0.13	0.14	0.14	0.20

Table 2.10: Fund horizon and investment staging

This table presents the results of an investment level OLS regression of the log number of rounds on the log of fund age, the size of the investment as a percentage of fund size, the log average number of rounds in past investments, and a variety of fund level controls including (i) the log number of past exits, (ii) the log number of past investments, (iii) the log number of funds operated by the same private equity firm, (iv) the log of fund size, (v) a dummy for first time funds and (vi) a dummy for "VC funds" and "Other funds". Several specifications are run with successively vintage, year, private equity firm and fund fixed effects. Standard errors are clustered by month and presented in parenthesis. ***, ** and * indicate significance at the 1, 5 and 10 percent level.

	Dependent variable: Log number of rounds					
Log fund age	-0.09***	-0.09***	-0.07***	-0.09***	-0.06***	-0.05**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Log nb. of exits	-0.05***	-0.05***	-0.03***	-0.04***	-0.03***	-0.06***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
First time fund x Log nb. of exits	-0.01	0.00	-0.03**	-0.01	-0.00	
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	
First time fund	-0.04***	-0.05***	0.04***	-0.05***	-0.07***	
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	
Log mean nb of rounds in past inv.	0.65***	0.60***	0.46***	0.32***	0.24***	-0.27***
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)
Size of the investment (% of fund size)	4.00***	3.94***	4.18***	4.28***	4.21***	4.53***
	(0.22)	(0.22)	(0.25)	(0.25)	(0.24)	(0.22)
Log nb. of past investments	0.00	0.01*	-0.01	-0.01	-0.00	-0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Log nb. of funds, same PE firm	-0.03**	-0.04***	-0.00	-0.02	-0.03*	0.00
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Log fund size	0.07***	0.07***	0.08***	0.12***	0.13***	
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	
"VC fund" dummy	0.22***	0.21***	0.10***	0.09**	0.06	
	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)	
"Other fund" dummy	0.09***	0.06**	0.01	-0.01	-0.02	
	(0.03)	(0.03)	(0.05)	(0.06)	(0.05)	
Constant	-0.74***	-0.68***	-0.47***	-0.26***	-0.17	0.94***
	(0.05)	(0.05)	(0.07)	(0.09)	(0.17)	(0.05)
Vintage fixed effects	Yes	No	No	Yes	No	No
Inv. year fixed effects	No	Yes	No	No	Yes	No
PE firm fixed effects	No	No	Yes	Yes	Yes	No
Fund fixed effects	No	No	No	No	No	Yes
Observations	24,222	24,222	24,222	24,222	24,222	24,222
R-squared	0.19	0.21	0.22	0.24	0.26	0.30

Table 2.11: Fund horizon and investment holding period, conditional on exit

This table presents the results of an investment level OLS regression of the log holding period (number of months between the investment and the exit) on the log of fund age, the log average holding period of the fund's past investments, and a variety of fund level controls including (i) the log number of past exits, (ii) the log number of past investments, (iii) the log number of funds operated by the same private equity firm, (iv) the log of fund size, (v) a dummy for first time funds and (vi) a dummy for "VC funds" and "Other funds". Several specifications are run with successively vintage, year, private equity firm and fund fixed effects. Standard errors are clustered by month and presented in parenthesis. ***, ** and * indicate significance at the 1, 5 and 10 percent level.

	Dependent variable: Log holding period (in months)					
Log fund age	-0.06**	-0.01	-0.13***	-0.15***	-0.03	-0.34***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)
Log nb. of exits	-0.06**	-0.05**	0.01	0.02	0.02	0.19***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
First time fund x Log nb. of exits	0.06**	0.03	0.09***	0.09***	0.07**	
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	
First time fund	-0.06**	-0.08**	-0.02	-0.11**	-0.13***	
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	
Log mean holding period in past inv.	0.13***	0.09***	0.08***	-0.03	-0.07***	-0.33***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.04)
Log nb. of past investments	-0.01	-0.00	0.01	0.00	-0.02	0.06*
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.04)
Log nb. of funds, same PE firm	0.06**	-0.05	0.19***	0.18***	-0.03	0.24***
	(0.03)	(0.03)	(0.04)	(0.05)	(0.04)	(0.05)
Log fund size	-0.01	-0.01	-0.03*	0.02	0.02	
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	
"VC fund" dummy	0.07	0.09	0.08	-0.04	-0.02	
	(0.06)	(0.05)	(0.10)	(0.10)	(0.10)	
"Other fund" dummy	0.01	-0.01	-0.08	-0.16	-0.07	
	(0.09)	(0.09)	(0.19)	(0.21)	(0.20)	
Constant	3.14***	3.31***	3.21***	3.93***	1.88***	4.68***
	(0.14)	(0.12)	(0.18)	(0.23)	(0.25)	(0.15)
Vintage fixed effects	Yes	No	No	Yes	No	No
Inv. year fixed effects	No	Yes	No	No	Yes	No
PE firm fixed effects	No	No	Yes	Yes	Yes	No
Fund fixed effects	No	No	No	No	No	Yes
Observations	7,245	7,245	7,245	7,245	7,245	7,245
R-squared	0.05	0.09	0.10	0.12	0.15	0.21

Table 2.12: Fund horizon and increase in patent count

This table presents the results of the following company × year regression (investment made up to December 2006 are included):

$$
\begin{aligned}
PC_{j,t+k} \quad = \quad & \alpha_0 + \alpha_1 Age_{i,t} + \alpha_2 F_{i,t} + \alpha_3 C_{j,t} + \sum_{k=-3}^{5} \lambda_k Y_{t+k} + \sum_{k=-3}^{5} \beta_k Y_{t+k} \times Age_{i,t} \\
& + \sum_{k=-3}^{5} \delta_k Y_{t+k} \times F_{i,t} + \sum_{k=-3}^{5} \theta_k Y_{t+k} \times C_{j,t} + \epsilon_{i,j,t}
\end{aligned}
$$

$PC_{j,t+k}$ is successively the log of one plus the number of patent applications and the log one plus the number of scaled patent applications by company j in year k around the investment year t. Y_{t+k} is a dummy equal to one in the k^{th} year around the investment of fund i in company j which occurs in year t. $Age_{i,t}$ is the log of the age of fund i at the time of the investment. $F_{i,t}$ is a vector of fund level controls including (i) the log number of past exits, (ii) the log number of past investments, (iii) the log number of funds operated by the same private equity firm, (iv) the log of fund size, (v) a dummy for first time funds and (vi) a dummy for "VC funds" and "Other funds". $C_{j,t}$ is a vector of company level controls, including the log of company age, state and sector dummies. Standard errors are clustered at the company level. Panel A the results of the specifications using company fixed effects, while Panel B includes company level controls (age, sector and state of incorporation). Standard errors are clustered by company and presented in parenthesis. ***, ** and * indicate significance at the 1, 5 and 10 percent level.

PANEL A: Within company

	Log patents + 1			Log scaled patents + 1		
Log fund age	0.12***	0.09***	0.11***	0.07***	0.05***	0.07***
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Inv. year -3 × Log fund age	0.06***	0.06***	0.07***	0.04***	0.04***	0.04***
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Inv. year -2 × Log fund age	0.08***	0.08***	0.08***	0.04***	0.04***	0.04***
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Inv. year -1 × Log fund age	0.04**	0.04**	0.05***	0.02*	0.02*	0.03**
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Inv. year +1 × Log fund age	-0.12***	-0.12***	-0.11***	-0.07***	-0.07***	-0.07***
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Inv. year +2 × Log fund age	-0.17***	-0.17***	-0.15***	-0.10***	-0.10***	-0.09***
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Inv. year +3 × Log fund age	-0.21***	-0.21***	-0.17***	-0.12***	-0.12***	-0.10***
	(0.03)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)
Inv. year +4 × Log fund age	-0.24***	-0.24***	-0.20***	-0.14***	-0.14***	-0.12***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
Inv. year +5 × Log fund age	-0.26***	-0.26***	-0.21***	-0.14***	-0.14***	-0.12***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
Constant	-0.34**	-0.88***	0.30***	-0.25***	-0.58***	0.16***
	(0.14)	(0.33)	(0.05)	(0.10)	(0.21)	(0.03)
Inv. year dummies × Fund controls	Yes	Yes	Yes	Yes	Yes	Yes
Inv. year dummies × Company controls	No	No	No	No	No	No
Company fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Inv. year fixed effects	No	Yes	No	No	Yes	No
Vintage fixed effects	Yes	No	No	Yes	No	No
PE firm fixed effects	Yes	Yes	No	Yes	Yes	No
Fund fixed effects	No	No	Yes	No	No	Yes
Observations	59,976	59,976	59,976	59,976	59,976	59,976
R-squared	0.42	0.42	0.41	0.44	0.44	0.44

PANEL B: Controlling for company's observable characteristics						
	Log patents + 1			Log scaled patents + 1		
Log fund age	0.05**	0.10***	0.08***	0.03*	0.06***	0.05**
	(0.02)	(0.02)	(0.03)	(0.01)	(0.02)	(0.02)
Inv. year -3 × Log fund age	0.03	0.03	0.04*	0.02	0.02	0.02
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Inv. year -2 × Log fund age	0.04	0.04	0.04*	0.02	0.02	0.02*
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Inv. year -1 × Log fund age	0.01	0.01	0.02	0.00	0.00	0.01
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Inv. year +1 × Log fund age	-0.11***	-0.11***	-0.10***	-0.07***	-0.07***	-0.06***
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Inv. year +2 × Log fund age	-0.15***	-0.15***	-0.13***	-0.09***	-0.09***	-0.08***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
Inv. year +3 × Log fund age	-0.18***	-0.18***	-0.15***	-0.10***	-0.10***	-0.09***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
Inv. year +4 × Log fund age	-0.21***	-0.21***	-0.16***	-0.12***	-0.12***	-0.10***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
Inv. year +5 × Log fund age	-0.23***	-0.23***	-0.19***	-0.13***	-0.13***	-0.10***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
Constant	-0.14	-0.32	0.17	-0.02	-0.12	0.15
	(0.39)	(0.40)	(0.31)	(0.28)	(0.29)	(0.22)
Inv. year dummies × Fund controls	Yes	Yes	Yes	Yes	Yes	Yes
Inv. year dummies × Company controls	Yes	Yes	Yes	Yes	Yes	Yes
Company fixed effects	No	No	No	No	No	No
Inv. year fixed effects	No	Yes	No	No	Yes	No
Vintage fixed effects	Yes	No	No	Yes	No	No
PE firm fixed effects	Yes	Yes	No	Yes	Yes	No
Fund fixed effects	No	No	Yes	No	No	Yes
Observations	51,426	51,426	51,426	51,426	51,426	51,426
R-squared	0.15	0.15	0.17	0.14	0.15	0.17

Table 2.13: Fund horizon and increase in citation count

This table presents the results of the following patent-level regression (investment made up to December 2006 are included):

$$CC_{j,t+k} = \alpha_0 + \alpha_1 Age_{i,t} + \alpha_2 F_{i,t} + \alpha_3 C_{j,t} + \sum_{k=-3}^{5} \lambda_k Y_{t+k} + \sum_{k=-3}^{5} \beta_k Y_{t+k} \times Age_{i,t}$$
$$+ \sum_{k=-3}^{5} \delta_k Y_{t+k} \times F_{i,t} + \sum_{k=-3}^{5} \theta_k Y_{t+k} \times C_{j,t} + \epsilon_{i,j,t}$$

$CC_{j,t+k}$ is successively the log of one plus the number of citations and the log of one plus the number of scaled citations received by a patent applied by company j in year k around the investment year t. Y_{t+k} is a dummy equal to one in the k^{th} year around the investment of fund i in company j which occurs in year t. $Age_{i,t}$ is the log of the age of fund i at the time of the investment. $F_{i,t}$ is a vector of fund level controls including (i) the log number of past exits, (ii) the log number of past investments, (iii) the log number of funds operated by the same private equity firm, (iv) the log of fund size, (v) a dummy for first time funds and (vi) a dummy for "VC funds" and "Other funds". $C_{j,t}$ is a vector of company level controls, including the log of company age, state and sector dummies. Standard errors are clustered at the company level. Panel A the results of the specifications using company fixed effects, while Panel B includes company level controls (age, sector and state of incorporation). Standard errors are clustered by company and presented in parenthesis. ***, ** and * indicate significance at the 1, 5 and 10 percent level.

PANEL A: Within company

	Log citations + 1			Log scaled citations + 1		
Log fund age	-0.18***	0.09**	-0.19***	-0.08***	0.05**	-0.08***
	(0.04)	(0.04)	(0.04)	(0.03)	(0.02)	(0.02)
Inv. year -3 × Log fund age	0.03	0.06	0.01	0.01	0.02	0.00
	(0.09)	(0.09)	(0.10)	(0.05)	(0.05)	(0.05)
Inv. year -2 × Log fund age	0.22***	0.24***	0.17***	0.09***	0.10***	0.06*
	(0.06)	(0.06)	(0.05)	(0.03)	(0.03)	(0.03)
Inv. year -1 × Log fund age	0.09*	0.10**	0.06	0.04	0.05	0.02
	(0.05)	(0.05)	(0.04)	(0.03)	(0.03)	(0.03)
Inv. year +1 × Log fund age	-0.05	-0.06	-0.05	-0.02	-0.02	-0.02
	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.02)
Inv. year +2 × Log fund age	-0.10**	-0.12***	-0.10**	-0.06**	-0.07**	-0.06**
	(0.05)	(0.05)	(0.04)	(0.03)	(0.03)	(0.03)
Inv. year +3 × Log fund age	-0.20***	-0.22***	-0.16***	-0.10***	-0.12***	-0.09**
	(0.06)	(0.06)	(0.06)	(0.04)	(0.04)	(0.04)
Inv. year +4 × Log fund age	-0.25***	-0.29***	-0.21***	-0.13***	-0.14***	-0.11***
	(0.06)	(0.06)	(0.07)	(0.04)	(0.04)	(0.04)
Inv. year +5 × Log fund age	-0.28***	-0.32***	-0.22***	-0.15***	-0.17***	-0.13***
	(0.07)	(0.07)	(0.07)	(0.04)	(0.04)	(0.04)
Constant	2.78***	1.65***	2.00***	1.27***	0.85***	0.92***
	(0.27)	(0.31)	(0.14)	(0.15)	(0.17)	(0.08)
Inv. year dummies × Fund controls	Yes	Yes	Yes	Yes	Yes	Yes
Inv. year dummies × Company controls	No	No	No	No	No	No
Company fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Inv. year fixed effects	No	Yes	No	No	Yes	No
Vintage fixed effects	Yes	No	No	Yes	No	No
PE firm fixed effects	Yes	Yes	No	Yes	Yes	No
Fund fixed effects	No	No	Yes	No	No	Yes
Observations	71,276	71,276	71,276	71,185	71,185	71,185
R-squared	0.44	0.44	0.44	0.41	0.41	0.41

PANEL B: Controlling for company's observable characteristics

	Log citations + 1			Log scaled citations + 1		
Log fund age	-0.14**	0.08	-0.16**	-0.04	0.05	-0.05
	(0.05)	(0.06)	(0.07)	(0.03)	(0.04)	(0.04)
Inv. year -3 × Log fund age	-0.02	0.07	-0.07	-0.03	0.01	-0.05
	(0.12)	(0.12)	(0.12)	(0.06)	(0.06)	(0.06)
Inv. year -2 × Log fund age	0.22***	0.30***	0.18**	0.09*	0.12***	0.05
	(0.08)	(0.08)	(0.08)	(0.05)	(0.05)	(0.04)
Inv. year -1 × Log fund age	0.10	0.14**	0.06	0.04	0.06	0.01
	(0.07)	(0.07)	(0.06)	(0.04)	(0.04)	(0.04)
Inv. year +1 × Log fund age	-0.07	-0.09	-0.07	-0.01	-0.02	-0.02
	(0.05)	(0.05)	(0.05)	(0.03)	(0.03)	(0.03)
Inv. year +2 × Log fund age	-0.13*	-0.17**	-0.13**	-0.07	-0.09*	-0.07*
	(0.07)	(0.07)	(0.06)	(0.05)	(0.05)	(0.04)
Inv. year +3 × Log fund age	-0.16**	-0.23***	-0.14**	-0.06	-0.10**	-0.06
	(0.07)	(0.07)	(0.07)	(0.04)	(0.04)	(0.04)
Inv. year +4 × Log fund age	-0.09	-0.18*	-0.09	-0.03	-0.08	-0.04
	(0.10)	(0.10)	(0.10)	(0.06)	(0.06)	(0.07)
Inv. year +5 × Log fund age	-0.13	-0.25***	-0.11	-0.08*	-0.14***	-0.07
	(0.08)	(0.08)	(0.08)	(0.05)	(0.05)	(0.05)
Constant	2.24***	1.44***	2.62***	0.92***	0.62**	1.16***
	(0.49)	(0.44)	(0.28)	(0.27)	(0.25)	(0.15)
Inv. year dummies × Fund controls	Yes	Yes	Yes	Yes	Yes	Yes
Inv. year dummies × Company controls	Yes	Yes	Yes	Yes	Yes	Yes
Company fixed effects	No	No	No	No	No	No
Inv. year fixed effects	No	Yes	No	No	Yes	No
Vintage fixed effects	Yes	No	No	Yes	No	No
PE firm fixed effects	Yes	Yes	No	Yes	Yes	No
Fund fixed effects	No	No	Yes	No	No	Yes
Observations	63,171	63,171	63,171	63,105	63,105	63,105
R-squared	0.16	0.17	0.20	0.13	0.13	0.17

Table 2.14: Fund horizon and subsequent sector excess returns

This table presents the result of the following investment level regression (investment made up to December 2006 are included):

$$ARet_{j,t+k} = \eta Age_{i,t} + \lambda Age_{i,t} \times SHARE_{j,t} + \beta SHARE_{j,t} + \delta X_{i,t} + \alpha_i + \epsilon_{i,j,t}$$

$ARet_{j,t+k}$ is the average five factor (market, small-minus-big, high-minus-low, momentum, liquidity) adjusted monthly return of the sector of company j in the k months following the investment by fund i in month t. $Age_{i,t}$ is the log of the age of fund i. $SHARE_{j,t}$ is the share of investments by private equity funds in sector j in month t. $X_{i,t}$ is a vector of time varying characteristics of fund i in month t including (i) the log number of exits, (ii) the number of investments involving fund i up to month t, (iii) the number of funds of the same class as fund i contemporaneously managed by the private equity firm managing fund i. α_i is a fund fixed effect. $\epsilon_{i,j,t}$ are residuals. Standard errors are corrected for month clustering and presented in parenthesis. ***, ** and * indicate significance at the 1, 5 and 10 percent level.

Dependent variable: Average monthly five factor adjusted returns				
Horizon	12 months	24 months	36 months	48 months
Log fund age	-0.04	-0.11*	-0.28***	-0.28***
	(0.09)	(0.06)	(0.05)	(0.04)
Log fund age x share of sector in total PE investments	0.38***	0.42***	0.45***	0.24***
	(0.14)	(0.13)	(0.11)	(0.09)
Share of sector in total PE investments	-0.32	0.01	-0.05	0.26
	(0.39)	(0.28)	(0.24)	(0.18)
Log number of exits	-0.14***	-0.06***	0.02	0.03*
	(0.04)	(0.02)	(0.02)	(0.02)
Log number of prior investments	0.01	0.02	0.05**	0.05***
	(0.04)	(0.02)	(0.02)	(0.02)
Log number of funds, same firm	0.22**	0.23***	0.22***	0.21***
	(0.09)	(0.06)	(0.05)	(0.04)
Constant	0.22*	0.21**	0.32***	0.29***
	(0.13)	(0.10)	(0.07)	(0.05)
Fund FE	yes	yes	yes	yes
Observations	19,351	19,351	19,351	19,351
R-squared	0.13	0.17	0.22	0.29

Table 2.15: Monthly calendar time portfolio excess returns

I build a series of monthly calendar time portfolios as follows. Each month I split investments in tertiles of age of the funds involved. I define young fund investments as investments of funds in the first tertile and old fund investments as investments of funds in the third tertile. I exclude sectors which are both invested by young and old funds. I first build a monthly calendar time portfolio of sectors picked by young private equity funds. Each month, I average the returns of sectors that have received investments by young private equity funds in the previous 12 months. I do the same at the 24, 36 and 48 months frequency. I build similar monthly calendar time portfolios of sectors which have received investments by old funds. I then compute the raw returns, CAPM alpha and five factor alpha generated by each of these portfolios. Finally, I build a portfolio long in sectors invested by young funds and short in sectors invested by old funds and compute abnormal returns of this portfolio using the CAPM and five factor model. Monthly returns are presented in percentages. P-values of a one-sided test are in parenthesis.

Investment holding period	12 months	24 months	36 months	48 months	60 months
Number of observations (months)	320	308	296	284	272
			Raw returns		
Long young funds	1.02***	1.05***	1.02***	1.05***	1.11***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Long old funds	1.05***	1.03***	0.96***	0.98***	1.03***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Long young funds short old funds	-0.03	0.02	0.06	0.06*	0.08*
	(0.67)	(0.34)	(0.12)	(0.08)	(0.06)
			CAPM alpha		
Long young funds	0.02	0.07	0.11	0.17	0.18
	(0.46)	(0.34)	(0.28)	(0.19)	(0.18)
Long old funds	0.06	0.07	0.07	0.12	0.13
	(0.36)	(0.35)	(0.35)	(0.25)	(0.25)
Long young funds short old funds	-0.04	0.00	0.04	0.04	0.05
	(0.77)	(0.49)	(0.21)	(0.18)	(0.15)
			Five factor model alpha		
Long young funds	0.034	0.051	0.066	0.091	0.143
	(0.373)	(0.316)	(0.275)	(0.204)	(0.102)
Long old funds	0.020	-0.003	-0.015	0.012	0.056
	(0.421)	(0.512)	(0.558)	(0.454)	(0.296)
Long young funds short old funds	0.014	0.054	0.081**	0.079**	0.087**
	(0.401)	(0.125)	(0.029)	(0.025)	(0.014)

Chapter 3

Heterogeneity in Retail Investors Behavior: Evidence from the Financial Crisis

Joint work with
Ron Kaniel (University of Rochester and CEPR) and
David Sraer (Princeton University and NBER)

Abstract

We use detailed brokerage account data to provide a quantitative exploration of the behavior of retail investors during the financial crisis of 2008. We identify ex ante dimensions of heterogeneity based on trading frequency, amount invested or type of securities traded that are associated with different trading behavior ex post. We show that investors who appear more sophisticated on these dimensions in the pre-crisis period were, in the post-crisis period, less likely to flee to safety, more likely to engage in liquidity provisions and to earn higher returns. Our analysis thus extend our understanding of retail investors' behavior and provide a new light on one particular mechanics of the financial crisis.

JEL classification: D10, G11

Keywords: financial crisis, individual investors, performance, liquidity provision

3.1 Introduction

Retail investors play an important role in the propagation of financial crisis. In a bank run, the lack of coordination among small depositors can lead to the depletion of a bank's capital in less than a day (see Diamond and Dybvig (1983) for a theoretical analysis and Iyer and Puri (2012) for an empirical analysis of bank runs). In a stock market crash, panic among retail investors can also lead to a dramatic fall in asset values. The episode of the Mississippi company represents a well-known example. In 1717, when Scottish banker John Law set up the Mississippi Company, which had control of trade between France and Louisiana, he also incidentally created one among the earliest stock markets by issuing stocks to the European public. Three years later, accumulating news that Louisiana had few precious metals triggered a rush of retail investors to convert their stocks to money, which ended in deadly riots and the collapse of the Mississippi Company.

The episode of the great financial crisis of 2008 appears to corroborate the view according to which retail investors flee to liquidity during times of financial market stress and thereby amplify the initial stress. Financial newspapers, both in Europe and in the U.S., have reported a massive exodus of small retail investors from the stock market following the financial crisis, with potentially worrisome consequences. According to the Wall Street Journal,[1] in the U.S., "from 2007 through 2009, [retail investors] withdrew money [from mutual funds that invest in U.S. stocks] for three consecutive years", which "marked the first three-year period of withdrawals since 1979-1981". "Their flight from stocks is changing the market dynamic. By adding money to mutual funds, individuals helped push stocks higher in the 1990s and to a lesser extent from 2003 through 2006. Now they are moving money out again on balance, making them a drag on the market." Similar concerns have been raised in France. According to the French newspaper Les Echos[2], the share of individual investors in the French stock market went from 13.8% in 2008 to 8.5% in 2012. In Figure

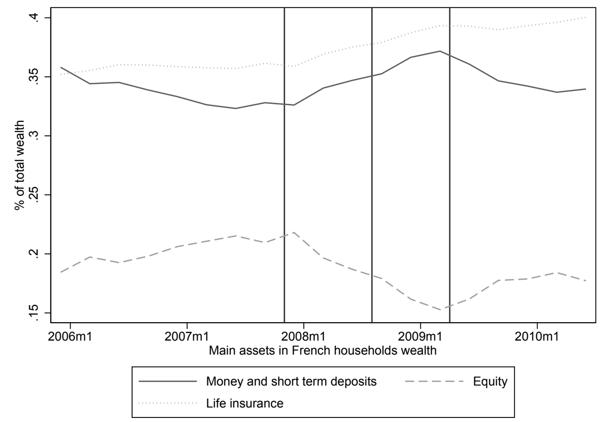
[1]See "Small Investors Flee Stocks, Changing Market Dynamics", Wall Street Journal, June 2010.
[2]See "Bourse : l'exode des investisseurs individuels", Les Echos, June 2012.

3.1, we use the households quarterly holdings data from Bank of France to construct the time-series of the portfolio allocation of French households. The top panel (a) shows that the bulk of French individual investors wealth is held through money, short term deposits, life insurance and equity. In the 18 months going from September 2007 to March 2009, the share of French households' wealth invested in equity went from 22% to 16%. A similar pattern emerges when we compare the share of listed stocks, mutual funds and bonds (which are the topic of this paper) in French households wealth in the middle panel (b). In the 18 months going from September 2007 to March 2009, the share of French households' wealth directly invested in listed stocks and equity mutual funds dropped sharply from 7% to 4%. Of course, since part of this trend could be driven by a price effect, we also look in the bottom panel (c) at the cumulative net quarterly inflows for French Households into risky assets (listed stocks and domestic equity mutual funds) and safe assets (bonds and domestic bonds mutual funds). This figure shows large and negative net outflows from risky assets relative to safe assets starting around the beginning of the crisis. As noted in the Les Echos article, this reduction in equity holding by small investors raises concerns about the long-run returns these investors might earn on their retirement savings.

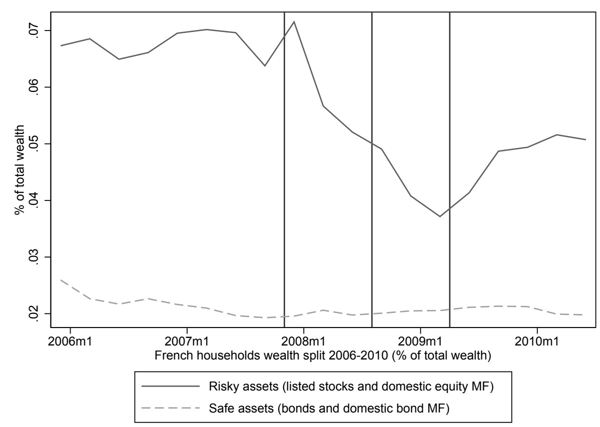
In the wake of the financial crisis of 2008, retail investors thus seem to have, in the aggregate, pulled back from stock markets around the world. In this paper, we use brokerage account data from a large French online broker[3] to provide a quantitative examination of the behavior of retail investors during the financial crisis of 2008/2009. From 1999 to 2010, this broker accounted for an average 15 percent of online brokers stock trades on Euronext Paris, which collectively represented 14 percent of all trades on this market, making our sample representative of the universe of retail investors in France. The data contains information on holdings and trades of all securities, including bonds, warrants, ETFs and mutual funds, for all the broker's customers between January 1999 to December 2010. It is thus particularly well-suited to analyze the investment behavior of individual investors, in particular in the

[3]See Foucault et al. (2011) for a detailed description of this dataset

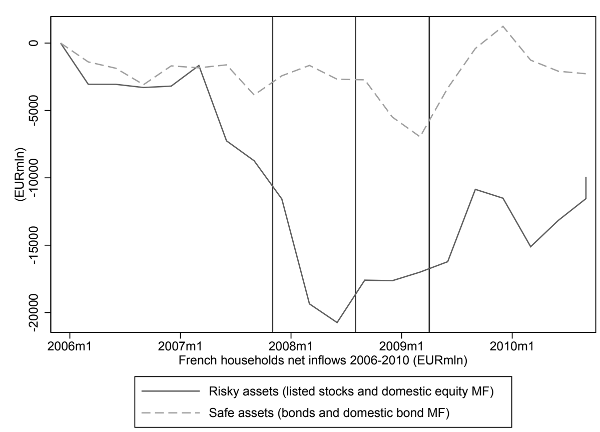
Figure 3.1: French households wealth split across asset classes, 2006-2010



(a) Wealth share



(b) Wealth share of safe and risky assets



(c) Cumulative net inflows of safe and risky assets

Notes: The three vertical lines indicate December 2007, September 2008 and April 2009.

midst of a major financial crisis.

Our analysis first confirms the aggregate flight to safety mentioned above. Splitting the set of tradable assets into risky assets (stocks, ETF and actively managed equity funds) and safe (bonds and actively managed bond funds) assets, we find that the average retail investor in our sample increased her trading of safe assets relative to risky assets during the crisis period (going from December 2007 to April 2009) by a significant 1.6%. This result is relative to the behavior of the same investors in the three years period preceding the crisis, which serves as a benchmark for the behavior of the retail investors in our sample. However, this average flight to safety is far from being uniform across traders. We define three arbitrary dimensions of ex ante heterogeneity among retail investors based on trading frequency, size of the trading account and type of securities traded. While these dimensions may capture many different features of the investors in our sample, they are loosely related to "sophistication" – we expect sophisticated investors to trade more frequently, have larger account or trade more in stocks relative to mutual funds. As it turns out, the flight to safety observed for the average investors is only present for these investors with less trading activity, smaller accounts or trading mostly in mutual funds in the pre-crisis period. To the contrary, more "sophisticated" investors – investors trading more frequently, with larger accounts or trading mostly individual stocks in the pre-crisis period – increase their trading of risky assets relative to safe assets in the crisis period relative to the pre-crisis period by a significant fraction going from 4.5% to 15.7%, depending on the type of heterogeneity considered. Therefore, the flight to safety observed on average during the crisis results from the portfolio reallocation of a specific set of investors, with easily identifiable ex ante characteristics.

In relation to this flight to safety, we then focus on investment in stocks and look at liquidity provision by retail investors. Kaniel et al. (2008) have documented on U.S. data that retail investors tend to buy stocks following recent declines in the stock price and sell following price increases. In doing so, they provide liquidity to meet institutional demand for immediacy and are rewarded for it with positive subsequent excess returns. We find similar evidence in our French sample: retail investors, in the aggregate, buy (sell) stocks in week t when the stock has experienced negative (positive) abnormal returns between week $t-2$ and week $t-1$. Interestingly, this contrarian behavior is left unchanged by the crisis – the elasticity of Net Individual Trading (NIT) [4] to past week returns does not significantly change once the crisis starts. Again, this aggregate result conceals a great deal of heterogeneity. Only the "sophisticated investors" – investors trading more frequently, with larger accounts or trading mostly in mutual funds – exhibit this contrarian behavior, and this both before and after the crisis. For the non-sophisticated investors, the aggregate Net Individual Trading is not significantly related to past returns[5]. These results emphasize how important the introduction of heterogeneity is to get a clear picture of the behavior of retail investors throughout the crisis. While a representative investor interpretation of our results would provide a somewhat inconsistent story – the average investor flees to safety but provide liquidity at the same time – the interpretation becomes much clearer once the heterogeneity is acknowledge: a group of retail investors increased its trading in risky assets and provided liquidity during the crisis; another group reduced its trading in risky assets, with a selling strategy unrelated to past week returns.

Our final set of results relates to the trading performance of retail investors during the crisis. We compute the Euro gains on transactions realized in both the pre-crisis and post crisis period made by each investor in our sample, normalized by the value of the investor's portfolio at the beginning of the period. We benchmark these returns against the aggregate stock market return. Three main results emerge from our performance analysis. First, the average retail investor in our sample is losing money relative to the market in the pre-crisis period: her realized yearly excess return is a significant -1.8% in the pre-crisis period. Second, this performance improves during the crisis: on average, the realized yearly excess returns of retail investors increase by .6 percentage points in the post-crisis period. Third, this

[4] As in Kaniel et al. (2008), NIT is defined as the ratio of weekly signed volume by retail investors to the average volume of the stock over the week.

[5] Note that these results confirm that the contrarian behavior highlighted by Kaniel et al. (2008) is unlikely to be driven by the execution of stale limit orders by inattentive investors

increase in performance during the crisis can be mostly attributed to the more "sophisticated" investors (who trade more frequently, have larger accounts or trade mostly in individual stocks as opposed to mutual funds): for instance, while the group of inactive investors experiences a reduction in yearly excess returns of about 1.9 percentage points during the crisis, the group of more active investors see its excess returns increase by a significant 5.4 percentage points during the crisis. In the light of our previous findings, these results have a natural interpretation. Providing liquidity can only be a profitable strategy in a market where liquidity is scarce. In the pre-crisis period, institutions can easily take on liquidity risk so that the returns of contrarian strategies are bound to be small. In the crisis period however, institutions might become more risk-averse, triggering an increase in the premium retail investors receive to provide liquidity. This view is consistent with the findings of Nagel (2012) who argues that that the recent market turmoil strained the inventory-absorption capacity of the market-making sector, thereby raising the expected return from liquidity provision. Although indirect, our results are thus consistent with a model in which *some* retail investors become substitute for financial institutions to provide liquidity in times of stress of financial markets.

As is evident from our analysis, to understand the behavior of retail investors, one needs to go beyond a representative investor analysis and to uncover the relevant dimension of heterogeneity in their investment behavior. But understanding this heterogeneity is important beyond what it teaches us about the nature and behavior of retail investors in general. One case in point is public policy. The efficiency of potential reforms to financial markets can strongly depend on such heterogeneity. The simplest example is a financial-transaction tax, which has been recently proposed by several governments of the E.U..[6]. Our analysis suggests that some easily observable characteristics of retail investors – such as past trading activity, the size of the trading account or the type of securities traded by the investor – predict sharply different investment behavior in times of crisis. Tagging a financial-transaction tax

[6]See for instance " Transaction tax proposal knocks shares", Financial Times, August 2011.

to any of these observable characteristics would be welfare improving relative to a uniform tax. For similar reasons, understanding this heterogeneity is useful for trading platforms when they design their transaction cost structure.

The remainder of the paper is organized in the following way. In Section 3.2, we describe the data used in the paper and how we construct the dimensions of heterogeneity used in the empirical analysis. In Section 3.3, we perform our main regression analyses to evaluate the behavior of retail investors in the crisis. Section 3.4 concludes.

3.2 Data and dimensions of heterogeneity

3.2.1 Description of the data

We use a large sample of French retail traders holdings and transactions between January 1999 and December 2010, provided by a leading European broker in personal investing and online trading. From 1999 to 2010, this broker accounted for an average 15 percent of online brokers stock trades on Euronext Paris, which collectively represented 14 percent of all trades in the market[7]. Hence we believe that our sample is fairly representative of the behavior of individual investors directly investing in the French stock market.

An individual is included in the sample if she posts at least one trade between January 2006 and April 2009. There are 67,553 investors in our final sample, each identified with a unique identifying number. For each of them, we obtain the year and the month of opening of the account, the year and month of birth, gender, county of residence as well as the type of accounts held (taxable, tax exempt, or both). At a monthly frequency, we obtain the quantity and value of each individual's position in each security, identified by its International Securities Identification Number (ISIN) thereafter. Unfortunately, neither we nor the data provider have any way to check whether individuals in the sample hold any other online or

[7] According to "Acsel", the association of French online brokers (see http://www.associationeconomienumerique.fr/) which collects monthly data on online trading.

offline account with other financial institutions. Not having access to a reliable estimate of the investor's wealth is a limitation of this dataset.

In addition to their holdings, we also have the full history of trades realized by each individual investor in the sample. For each order of stocks, bonds, warrants and ETFs, the dataset includes the identifier of the individual investor, the security's ISIN, the direction of the trade (buy or sell), the quantity and amount in Euros, and the day when the order was placed. From August 2001 onwards, we also have the type of order used (market versus limit orders). However, we do not observe unexecuted limit orders. Trades of mutual fund shares are only observed after 2006 – these trades are identified with the security ISIN, the direction of the trade and the amount traded[8].

To conduct our performance analysis, we obtain daily European stock and ETFs returns from EUROFIDAI.[9] Data for prices of other international stocks and ETFs are extracted from Datastream and when missing, they are inferred from our transaction file by dividing the value of a transaction by the number of securities transacted. Data for prices of bonds are from Datastream Bonds. Finally, data on prices of mutual funds shares come from EUROFIDAI MF. When prices are missing on a security, we use the last available price observed in the dataset. We also use EUROFIDAI historical exchange rates to convert foreign currency prices to Euros. While the coverage of price data for securities other than stocks is somewhat imperfect, it is important to keep in mind that stocks account for approximately 89 percent of the volume traded by investors in our sample.

For the purpose of our analysis, we also split the sample into two periods: a pre-crisis period, which runs from January 2006 (the date at which mutual fund data becomes available)

[8]We have not had access yet to monthly positions of mutual funds. To compute monthly holdings of mutual funds, we backlog positions from trades observed between 2006 and 2010. Hence, for instance, if the only trade posted by individual i between 2006 and 2010 is a sale of x shares of mutual fund n in 2009, then we assume that the position of individual i in mutual fund n in January 2006 was of x shares. This imperfect measure of mutual fund holdings might reduce the accuracy of our computation of performance, which is normalized by total portfolio value. However, we compute the average holding period of mutual funds in our sample and find that it is less than a year in our sample. Hence we believe that our results are unlikely to be severely affected by our computation of mutual fund holdings.

[9]EUROFIDAI is a research institute funded by the CNRS (French National Center for Scientific Research) whose mission is to develop European stock exchange databases for academic research.

to December 2007 and a crisis period which goes from December 2007 to April 2009. April 2009 corresponds to the beginning of the great stock rally of 2009, which saw the market index gain more than 80% in less than a year.

Our sample stands out in a number of ways. First, it covers the recent financial crisis and makes it possible to describe the behavior of a large number of retail investors around the period of high market stress running from December 2007 to April 2009. Second, the data does not cover only stocks trading but also contains individual level information on bonds, warrants[10], ETFs and mutual funds. This is somewhat unique in the universe of brokerage account data.

3.2.2 Dimensions of heterogeneity

As mentioned in the introduction, the main focus of this paper is on analyzing how some ex ante dimensions of heterogeneity among retail investors predict their ex post behavior during the crisis. Our purpose is to use observable persistent individual characteristics that are likely to be correlated with "sophistication": trading frequency, size of the trading account and type of securities traded. Precisely, to allocate investors to various categories, we examine their behavior in the pre-crisis period, which runs from January 2004 to December 2007. Our first category relates to the trading frequency. Each quarter, we sort investors into terciles of number of trades realized in the quarter. A retail investors is in the active (resp. inactive) category if she falls in the top (resp. bottom) tercile in more than two thirds of the quarters. Our second category relates to the size of the account. Each quarter, we sort investors in terciles of average account size. Investors are allocated to the large (resp. small) category if they fall in the top (resp. bottom) tercile of account size in more than two thirds of the quarters. Our third category relates to the type of securities mostly traded by the investors. Each quarter, we sort investors based on the ratio of stock volume they trade to the volume of stock *and* mutual funds they trade. Investors are allocated to the stock (resp.

[10] Note that we do not use warrant trading data in the analysis that follow.

mutual funds) category if they are above (resp. below) the median ratio in more than two thirds of the quarters. Importantly, since information on mutual fund trading is available only starting in January 2006, this split between mutual fund and stock traders has to be performed only on the period going from January 2006 to December 2007.

Of course, each of the variables we use to define these dimensions of heterogeneity is correlated with many different investors attributes, such as risk aversion, overconfidence, wealth, etc.... As a consequence, there are many different interpretations of our results in terms of a precise micro-mechanism. While the results we present below are consistent with one particular interpretation we favor – that the common component among these ex ante characteristics is related to sophistication – they are also consistent with a wide range of alternative interpretations. For instance, the frequency of trading might be directly related to over-confidence, and over-confidence can well explain part or all of the trader's behavior with respect to her trading strategy and her ultimate performance. However, whatever the micro-interpretation of our results, we still believe that our results are interesting in that they highlight how several dimensions of ex ante heterogeneity do predict the behavior of retail investors during the financial crisis.

Table 3.1 presents a transition matrix for investors for each of the three dimensions of heterogeneity used in the paper. In both the pre-crisis and the crisis period, we allocate investors to one of the group defined above. Of course, due to the definition we use for each category, a large fraction of investors cannot be allocated – for instance, an investor who is in the top tercile of numbers of trades in half of the quarters and in the bottom tercile in the remaining half of the quarters would neither be active nor inactive. Table 3.1 shows quite strikingly the persistence of the categories we have defined. For instance, in the trading frequency (resp. size and type of security) dimension, only 3% (resp. .7% and 2.2%) of investors switch from one category to the other from the pre-crisis to the crisis period. Table 3.1 thus makes us confident that these dimensions of heterogeneity are indeed capturing fixed-in-time attributes of retail investors.

In Table 3.2, we look at the overlap between the three dimensions of heterogeneity. Typically, our three categories tend to be highly correlated. For instance, among the Active investors, 58% are in the Large category and 74% are in the Stock category. At the same time, only 12% are in the Small category and 13% in the Mutual Fund category. These results are consistent with the premise of our analysis that our three dimensions of heterogeneity are capturing a common attribute of the retail investors in our sample, which we interpret as "sophistication". The remainder of our analysis tries to establish how this attribute relates to the investor's trading behavior.

3.3 Empirical analysis

3.3.1 Flight to safety

Our first analysis relates to the asset allocation choice of retail investors in the crisis period. The idea that retail investors flee to liquidity during the financial crisis is certainly consistent with the aggregate evidence presented on Figure 3.1. To investigate this question using our brokerage account data, we perform the following regression analysis. We first define a variable capturing the imbalance on a given security for each individual every week. This variable is defined for individual j and security i at date t:

$$Y_{it}^{j} = \begin{cases} log(1 + \text{avg weekly imbalance ('000 euro)) if imbalance>0} \\ -\,log(1 + |\text{avg weekly imbalance ('000 euro)) if imbalance<0} \end{cases},$$

where "avg weekly imbalance" is the mean weekly signed volume (in thousand euros) of transactions realized by investor j for period t (where t is either equal to the pre-crisis period or the crisis period so there are only two observations per individual/security) on security i. The reason why we take logs is that these average weekly imbalances have large outliers that are likely to influence the results in a linear specification. The reason why we that the log of 1 plus the average weekly imbalance is that for some traders, this average can be 0 when they don't trade over a large period of time. Taking directly the log would thus take these traders out of the sample. But the fact that they don't trade still contains relevant information. In any case, we also present a version of this regression analysis (see Table 3.7) where the dependent variable is directly the average weekly imbalance. The results are qualitatively similar to the ones we present here.

Once our imbalance variable is defined, we split the set of securities into two groups. Safe assets include bonds and actively managed bond funds. Risky assets include stocks, ETFs and actively managed equity funds. We then simply estimate the following linear model separately for the safe and risky assets:

$$Y_{it}^{j} = \alpha + \beta \times Crisis_t + \epsilon_{it}^{j},$$

Column 1 of Table 3.4 reports the α and β coefficients for the risky assets and Column 2 reports the coefficients for the safe assets. On average, the imbalance on risky assets remained constant in the crisis period relative to the pre-crisis period; the imbalance on safe assets increased by approximately 1.6 percentage points and this increase is significant at the 1% level. Thus, on average, retail investors in our data did significantly alter their portfolio allocation choice during the crisis to invest more in safe assets. However, Column 3 to 8 of Table 3.4 shows that this flight to safety is far from uniform across investors category. To establish this fact, we simply estimate a difference-in-difference equation:

$$Y_{it}^{j} = \alpha + \beta \times Crisis_t + \gamma \times X_j + \delta \times X_j \times Crisis_t + \epsilon_{it}^{j},$$

where X_j is a dummy variable equal to 1 if the investor is in the Active category (resp. Large or Stock) and 0 if she is in the Inactive category (resp. Small or Mutual Fund). When X is the Active dummy variable, the coefficient δ provides an estimation of the change in behavior of Active investors in the crisis relative to before the crisis, and this relative to the Inactive category. For instance, column 3 shows that Active investors *increased* their demand for risky asset by a significant 3.4% while *decreasing* at the same time their demand for safe assets by a significant 1.1%, relative to what Inactive investors did. On the other hand, the Inactive investors, the reference group, did experience a significant decline (-1.3%) in their demand for risky assets and a significant increase (+1.8%) in their demand for risky assets. Table 3.4 highlights a striking contrast between the investment behavior of Active and Inactive investors. Only those inactive investors flee to safety; Active investors actually increased their relative demand for risky assets in a significant way. This contrast is also apparent for the two other dimensions of heterogeneity. Large investors – relative to the Small ones – increased their relative demand for risky asset by about 2.6%. Investors in stocks – relative to investors in mutual funds – increased their relative demand of risky assets by about 5.8%.

3.3.2 Liquidity provision

We now turn to our second set of results on liquidity provision by retail investors during the crisis. We use a methodology similar to that in Kaniel et al. (2008). We first define, for each stock i, the Net Individual Trading (NIT) as the weekly signed volume by individual investors normalized by the market wide volume of the stock. To limit the role of outliers, we winsorize this variable at the 1% level. NIT captures the overall demand by retail investors in our sample for a given stock. Liquidity provision by retail investors on a given stock is then simply measured by the elasticity of this stock's NIT to the stock's past-week return. To avoid demand effects driven by market-wide movements, we adjust the stock past-week return for the overall market return over the same period. We call $CAR_{2,1}$ this past-week abnormal return variable. To analyze the average change in liquidity provision by retail investors during the crisis, we start by running the following linear regression:

$$NIT_{it} = a + b \times CAR_{2,1}^{it} + c \times Pre_t + d \times Pre_t \times CAR_{2,1}^{it} + \eta_{it},$$

where NIT_{it} is the NIT of stock i in week t and Pre is a dummy variable equal to 1 in the pre-crisis period. Column 1 of Table 3.5 reports the estimated coefficients. Two striking facts emerge from this regression. First, as in Kaniel et al. (2008), our retail investors are on average contrarian, i.e. in the aggregate, they buy (sell) stocks in week t when the stock has experienced negative (positive) abnormal returns between week $t-2$ and week $t-1$. Second, this contrarian behavior is left unchanged by the crisis. The elasticity of Net Individual Trading (NIT) to past week returns does not significantly change once the crisis starts – the d coefficient is not significantly different from 0. Quantitatively, this contrarian behavior is highly significant but has a limited magnitude. A 1 s.d. increase in the weekly abnormal return on a stock (i.e .072) leads to a 3% s.d. increase in NIT in the following week.

This result that retail investors tend to provide liquidity on average may appear surprising when related to our result in Section 3.3.1 that the average investor in the sample did flee to safety during the crisis. However, these results are easily reconciled when introducing our dimensions of heterogeneity in the analysis. To that end, we simply augment the previous regression by looking separately at the Net Individual Trading coming from investors in different categories. Call NIT_{it}^{X} the Net Individual Trading of stock i at date t computed at the level of investors with a similar characteristic X, where X is one of the three sources of heterogeneity defined in Section 3.2.2. For instance, $NIT_{it}^{X=active}$ is the weekly signed volume coming from Active investors normalized by the market wide volume of the stock in that week. We then simply estimate the following model:

$$\begin{aligned} NIT_{it}^{X} \;&=\; A + B \times CAR_{2,1}^{it} + C \times Pre_t + D \times Pre_t \times CAR_{2,1}^{it} + \\ &\qquad E \times X + F \times X \times CAR_{2,1}^{it} + G \times X \times Pre_t + H \times X \times Pre_t \times CAR_{2,1}^{it} + \nu_{it}, \end{aligned}$$

where X is equal to 1 if the NIT has been computed over the sample of investors with characteristics X and 0 if it has been computed over the sample of investors with characteristics

$-X$.

Column 2, 3 and 4 estimate the previous equation for the Active/Inactive split, the Large/Small split and the Stock/Mutual Fund split respectively. In all three regressions, we find the following two main results. First, only the Active, Large and Stock investors provide liquidity over the whole sample period. The elasticity of NIT to past-week abnormal returns is not significantly different from 0 when looking at the NIT coming from the investors trading infrequently, with small accounts or trading mostly in mutual funds. On the other hand, the elasticity of NIT to past-week abnormal return is negative and significant in the crisis period (the reference period in this regression) for the Active, Large and Stock Investors. Second, the crisis did not affect the nature of the trading strategies of the different group. We see that the elasticity of NIT to $CAR_{2,1}$ is not significantly different in the pre-crisis period than in the crisis period both for the "non-sophisticated" investors (the reference group in this regression) and the "sophisticated" investors.

The picture that emerges at this stage of the analysis is that "sophisticated" investors did not flee to safety during the crisis; to the contrary, they increased their relative demand for risky assets. Conversely, "non-sophisticated" investors increased their relative demand for safe assets, effectively fleeing to safety. At the same time, "sophisticated" investors were providing liquidity on the stock market, in a similar fashion as what they were doing in the period leading up to the financial crisis. The next section analyzes the financial gains made by these different categories of investors during the crisis.

3.3.3 Performance analysis

In this Section, we analyze the performance of our different categories of investors. To this end, we must first compute a performance measure for all the investors in our sample. Since we are primarily interested in assessing the rewards from liquidity provision, we define Euro gains on investments in the following way. We compute for each individual j, both for the crisis and the pre-crisis period:

$$Euro\ Gains^j = \frac{52}{W^j \times T} \sum_{n=1}^{N} \sum_{t=1}^{T} Q_{j,n,t} \times P_{n,t} \times AR_{n,t,t+1}$$

, where N is the number of assets, T is the number of weeks in each period, $Q_{i,n,t}$ is the cumulative number of units of asset n purchased/sold by individual i since the beginning of the period, $P_{n,t}$ is the price of asset n in week t and $AR_{n,t,t+1}$ is the return of asset n from week t to $t+1$, net of the market return over the same period. W^j is the value of individual j's portfolio at the beginning of the period. While our data includes the monthly position in each asset, we have not had access yet to monthly position of mutual funds. We backlog monthly mutual funds positions from trades observed between 2006 and 2010. Hence, for instance, if the only trade posted by individual i between 2006 and 2010 is a sale of x shares of mutual fund n in 2009, then we assume that the position of individual i in mutual fund n in January 2006 was of x shares. This imperfect measure of holdings might reduce the accuracy of our computation of performance, which is normalized by total portfolio value. However, we compute the average holding period of mutual funds in our sample and find that it is less than a year in our sample. Hence we believe that our results are unlikely to be severely affected by our computation of mutual fund holdings. Euro Gains are computed for each individual over the pre-crisis and the crisis period and are then winsorized at the 1% level. These Euro Gains can be interpreted as the returns on the initial portfolio value of the investor made on the active investment decisions.

To analyze the average performance in the sample, and how it was affected by the crisis, we simply estimate the following linear model:

$$Euro\ Gains_t^j = \lambda + \mu Crisis_t + \omega_{jt},$$

where $Crisis_t$ is a dummy variable equal to 1 in the crisis period and there is one observation per individual for each period (i.e. crisis and pre-crisis, the Euro gains are computed over these two periods). Column 1 of Table 3.6 presents the coefficient estimates from the previous equation. In the pre-crisis period, the average retail investor was losing money relative to the market, with a significant average yearly return of -1.8%. The average performance goes up significantly during the crisis period by .6 percentage points. One interpretation of these seemingly surprising facts is that, as we saw in Table 3.5, the average retail investor tends to provide liquidity. In the pre-crisis period, the liquidity premium must have been small, generating only low returns for the contrarian investors. Moreover, since retail investors tend to invest in high beta stocks[11], this contrarian behavior in a period where the market is mostly going up would have yielded negative abnormal returns. Column 2, 3 and 4 suggest that the liquidity provision behavior of retail investors may indeed be related to these performance results. In these columns, we simply interact our baseline equation with the three dimensions of heterogeneity defined in Section 3.2.2:

$$Euro\ Gains_t^j = \rho + \sigma Crisis_t + \tau X^j + \theta Crisis \times X^j + \Omega_{jt},$$

θ is then our coefficient of interest, which captures the relative performance increase, during the crisis, of investors with characteristics X. Column 2, 3 and 4 of Table 3.6 show that both the negative performance in the pre-crisis period and the significant increase in performance in the crisis period can be attributed only to those retail investors with "sophisticated" characteristics, i.e. investors with higher trading frequency, larger account or trading mostly in stocks as opposed to mutual funds. For instance, Inactive investors have actually positive and significant abnormal returns of about 1.3% in the pre-crisis period; their performance significantly declines during the crisis by -1.9 percentage points and becomes overall negative. Active investors, on the other hand, have negative and significant abnormal returns in the pre-period of about -5.4%. However, their performance increases significantly during the crisis period by 5.4%, leaving them with an overall return in this period approximately equal to the market return. An investigation of the other dimensions of heterogeneity leads to

[11]In our sample, the beta on the average retail investor's portfolio is 1.38.

qualitatively similar results. For these individuals who did not flee to safety (Table 3.4), or did provide liquidity throughout the sample period (Table 3.5), abnormal returns on their portfolio was negative in the pre-crisis period but improved significantly during the crisis period. One interpretation, albeit by far not the only one, is that providing liquidity turned out to be only a profitable strategy in the midst of the financial crisis, when institutions and a category of retail investors were pulling back from the stock market and fleeing to safety.

In Table 3.8, we detail these results by category of assets traded: for each retail investor in the sample, we compute the Euro gains separately on bonds, bond mutual funds, equity mutual funds, stocks and ETFs. We then compute the μ and θ coefficient for each of these performance measures. The results presented in Table 3.8 show that most of these performance results come from performance on stocks. This is not surprising this trading in stocks represent more than 89% of the active investment decisions of the retail investors in our sample.

3.4 Conclusion

Retail investors are not all alike. Accounting for the heterogeneity in their investment behavior is key to understand their behavior during the financial crisis. While many accounts of the crisis emphasize a flight to safety by retail investors, we show that a careful examination of brokerage account data reveals a more complicated pattern. A category of investors with more "sophisticated" characteristics (those investors trading frequently, with larger account or trading mostly in stocks as opposed to mutual funds), did actually increase its demand for risky assets during the crisis. At the same time, these investors were providing liquidity on stocks and got compensated for it. This behavior is relative to less "sophisticated" investors, who did flee to safety, were never providing liquidity and experienced a decline in their performance in the crisis period.

Understanding the heterogeneity in the behavior of retail investors has potentially important implications. For the design of public policies such as a transaction tax, it is important to understand how well-defined, observable, ex ante characteristics affect the ex post behavior of investors on financial market. For instance, an efficient design of a financial-transaction tax should tag individuals along the relevant dimensions of heterogeneity. Trading platforms can also benefit from a better understanding of this inherent heterogeneity by conditioning transaction fees on these same characteristics.

Bibliography

Diamond, Douglas W and Philip H Dybvig, "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy*, June 1983, *91* (3), 401–19.

Foucault, Thierry, David Alexandre Sraer, and David J. Thesmar, "Individual Investors and Volatility," *Journal of Finance*, 2011, *66* (4), 1369–1406.

Iyer, Rajkamal and Manju Puri, "Understanding Bank Runs: The Importance of Depositor-Bank Relationships and Networks," *American Economic Review*, June 2012, *102* (4), 1414–45.

Kaniel, Ron, Gideon Saar, and Sheridan Titman, "Individual Investor Trading and Stock Returns," *Journal of Finance*, 2008, *63* (1), 273–310.

Nagel, Stefan, "Evaporating Liquidity," *Review of Financial Studies*, 2012, *25* (7), 2005–2039.

Appendix

Table 3.1: Ex ante heterogeneity: persistence

This table presents the transition matrix of retail investors into the three categories defined in Section 3.2.2. For the three dimensions of heterogeneity, we allocate each investor in the sample both in the pre-crisis and the post-crisis period. ND corresponds to the set of retail investors who can't be allocated to one of the category. Reading: 0.6% of retail investors which were assigned to the Inactive group in the pre-crisis period are assigned to the Active group in the Crisis period.

		Crisis period			
		Inactive	ND	Active	Total
	Inactive	15.5%	29.9%	0.6%	46.1%
pre-crisis	ND	14.7%	17.6%	4.3%	36.5%
	Active	2.4%	6.8%	8.1%	17.4%
		Small	ND	Large	Total
	Small	9.6%	13.9%	0.2%	23.7%
pre-crisis	ND	7.6%	37.3%	3.3%	48.3%
	Large	0.5%	8.6%	18.9%	28.0%
		MF	ND	Stock	Total
	MF	6.0%	7.7%	1.2%	14.9%
pre-crisis	ND	2.9%	35.1%	6.1%	44.1%
	Stock	1.0%	20.4%	19.7%	41.0%

Table 3.2: Dimensions of heterogeneity: correlation

This table presents the overlap between the three dimensions of heterogeneity defined in Section 3.2.2. Reading: 36% of Large investors are in the Active category

Investor Type	Overlap with other types					
	Active	Inactive	Large	Small	Stock	MF
Active	100%	0%	58%	12%	74%	13%
Inactive	0%	100%	16%	34%	21%	16%
Large	36%	26%	100%	0%	62%	12%
Small	9%	65%	0%	100%	30%	10%
Stock	31%	24%	42%	17%	100%	0%
MF	15%	48%	23%	16%	0%	100%

Table 3.3: Summary statistics over the whole sample

Summary statistics for the dependent variables used in the regression analyses. Y_{it}^{j} is our imbalance variable and is defined as the log (1 + avg weekly imbalance ('000 euro)) if imbalance>0 and -log (1 + abs(avg weekly imbalance ('000 euro))) if imbalance<0, where "avg weekly imbalance" is the mean weekly signed volume (in thousand euros) of transactions realized by investor j at date t on security i. NIT is the net individual trading of a given stock and is defined as the weekly signed volume by individual investors normalized by the market wide volume of the stock. This variable is winsorized at the 1% level. Euro gains are defined in the following way. We compute for each individual i: $Euro\ Gains = \frac{52}{W^j \times T} \sum_{n=1}^{N} \sum_{t=1}^{T} Q_{j,n,t} \times P_{n,t} \times AR_{n,t,t+1}$, where N is the number of assets, T is the number of weeks, $Q_{i,n,t}$ is the cumulative number of units of asset n purchased/sold by individual i since the beginning of the period, $P_{n,t}$ is the price of asset n in week t and $AR_{n,t,t+1}$ is the return of asset n from week t to $t+1$, net of the market return over the same period. W^j is the value of individual j's portfolio at the beginning of the period. Euro gains are winsorized at the 1% level.

	Observations	Mean	Std. Dev.	p10	p50	p90
NIT	148,324	-0.0008	0.0503	-0.0177	0	0.0168
Y_{it}^{j} for risky assets	125,041	-0.0124	0.2382	-0.1593	0.0000	0.1174
Y_{it}^{j} for safe assets	125,041	0.0035	0.0725	0.0000	0.0000	0.0001
Euro Gains – All assets	123,427	-0.01507	0.2843	-0.16359	0	0.12796
Euro Gains – Bonds	124,084	-0.00137	0.99002	0	0	0
Euro Gains – Bond MF	124,145	0.00002	0.00016	0	0	0
Euro Gains – Equity MF	123,585	0.00012	0.0299	-0.00513	0	0.00886
Euro gains – Stocks	124,090	-0.01534	0.23177	-0.13668	0	0.0983
Euro Gains – ETFs	124,194	0.00004	0.00083	0	0	0

Table 3.4: Flight to safety and heterogeneity

The dependent variable is the log (1 + avg weekly imbalance ('000 euro)) if imbalance¿0 and -log (1 + abs(avg weekly imbalance ('000 euro))) if imbalance¡0. This table presents individual level OLS regressions of the log of mean weekly signed volume (in thousand euros) of safe or risky assets on a dummy (CRISIS) equal to one in the crisis period and on a dummy for each investor type. There are two observations for each individual: the average weekly imbalance in the pre-crisis and in the crisis period. Safe assets include bonds and actively managed bond funds. Risky assets include stocks, ETFs and actively managed equity funds. Active, Large and Stock are dummy variables corresponding to the categories defined in Section 3.2.2. *, **, and *** means statistically different from zero at 10, 5 and 1% level of significance.

	(1) Log risky	(2) Log safe	(3) Log risky	(4) Log safe	(5) Log risky	(6) Log safe	(7) Log risky	(8) Log safe
Crisis	0.000	0.016***	-0.013***	0.018***	-0.005	0.005***	-0.088***	0.050***
	(0.001)	(0.000)	(0.002)	(0.001)	(0.003)	(0.001)	(0.004)	(0.001)
Active			0.009***	0.002**				
			(0.003)	(0.001)				
Crisis × Active			0.034***	-0.011***				
			(0.004)	(0.001)				
Large					-0.022***	-0.004***		
					(0.003)	(0.001)		
Crisis x Large					0.032***	0.006***		
					(0.004)	(0.001)		
Stock							-0.017***	0.015***
							(0.003)	(0.001)
Crisis x Stock							0.110***	-0.047***
							(0.004)	(0.001)
Constant	-0.012***	-0.004***	-0.017***	-0.004***	-0.007***	-0.001***	0.008***	-0.016***
	(0.001)	(0.000)	(0.001)	(0.000)	(0.002)	(0.000)	(0.003)	(0.001)
Observations	125,041	125,041	78,660	78,660	64,346	64,346	72,624	72,624
R^2	0.000	0.012	0.003	0.013	0.002	0.005	0.013	0.036

Table 3.5: Liquidity provision and heterogeneity

NIT is the net individual trading of a given stock and is defined as the weekly signed volume by individual investors normalized by the market wide volume of the stock. This variable is winsorized at the 1% level. Pre is a dummy equal to 1 in the pre-crisis period and 0 in the crisis period. $CAR_{2,1}$ is the past-week return on the stock in excess of the market return over the same period. *, **, and *** means statistically different from zero at 10, 5 and 1% level of significance. Active, Large and Stock are dummy variables corresponding to the categories defined in Section 3.2.2. *, **, and *** means statistically different from zero at 10, 5 and 1% level of significance.

	Net Individual Trading			
	(1)	(2)	(3)	(4)
Pre	0.00	0.00	0.000***	0.00
	(0.000)	(0.000)	(0.000)	(0.000)
Pre x $CAR_{2,1}$	0.006	0.001**	0.001	-0.001
	(0.004)	(0.000)	(0.001)	(0.001)
$CAR_{2,1}$	-0.017***	0.000	0.000	-0.001
	(0.003)	(0.000)	(0.000)	(0.000)
Active		0.000**		
		(0.000)		
Pre x Active		0.000		
		(0.000)		
Pre x Active x $CAR_{2,1}$		-0.002		
		(0.002)		
Active x $CAR_{2,1}$		-0.008***		
		(0.001)		
Large			-0.000**	
			(0.000)	
Pre x Large			0.000	
			(0.000)	
Pre x Large x $CAR_{2,1}$			(0.001	
			(0.002)	
Large x $CAR_{2,1}$			-0.008***	
			(0.001)	
Stock				-0.001***
				(0.000)
Pre x Stock				0.000
				(0.000)
Pre x Stock x $CAR_{2,1}$				0.002
				(0.002)
Stock x $CAR_{2,1}$				-0.008***
				(0.001)
Constant	-0.001***	-0.000***	-0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Stock FE	Yes	Yes	Yes	Yes
Observations	148,324	296,648	296,648	296,648
R^2	0.022	0.009	0.01	0.01

Table 3.6: Performance analysis and heterogeneity

Euro gains are defined in the following way. We compute for each individual i: *Euro Gains* $= \frac{52}{W^j \times T} \sum_{n=1}^{N} \sum_{t=1}^{T} Q_{j,n,t} \times P_{n,t} \times AR_{n,t,t+1}$, where N is the number of assets, T is the number of weeks, $Q_{i,n,t}$ is the cumulative number of units of asset n purchased/sold by individual i since the beginning of the period, $P_{n,t}$ is the price of asset n in week t and $AR_{n,t,t+1}$ is the return of asset n from week t to $t+1$, net of the market return over the same period. W^j is the value of individual j's portfolio at the beginning of the period. Euro gains are computed for each individual over the pre-crisis and the crisis period and are then winsorized at the 1% level. Crisis is a dummy equal to 1 in the crisis period. Active, Large and Stock are dummy variables corresponding to the categories defined in Section 3.2.2. *, **, and *** means statistically different from zero at 10, 5 and 1% level of significance.

	Euro Gains			
	(1)	(2)	(3)	(4)
Crisis	0.006***	-0.019***	-0.009***	0.025***
	(0.002)	(0.002)	(0.003)	(0.004)
Active		-0.067***		
		(0.003)		
Crisis x Active		0.054***		
		(0.004)		
Large			-0.020***	
			(0.003)	
Crisis x Large			0.024***	
			(0.005)	
Stock				-0.030***
				(0.003)
Crisis x Stock				0.016***
				(0.005)
Constant	-0.018***	0.013***	0.002	-0.027***
	(0.001)	(0.002)	(0.002)	(0.003)
Observations	123,427	77,558	64,043	72,187
R^2	0.000	0.006	0.001	0.006

Table 3.7: Flight to safety and heterogeneity: using $ imbalances

The dependent variable is the average weekly imbalance (in '000 Euros). This table presents individual level OLS regressions of the mean weekly signed volume (in thousand euros) of safe or risky assets on a dummy (CRISIS) equal to one in the crisis period and on a dummy for each investor type. There are two observations for each individual: the average weekly imbalance in the pre-crisis and in the crisis period. Safe assets include bonds and actively managed bond funds. Risky assets include stocks, ETFs and actively managed equity funds.

	(1) Risky	(2) Safe	(3) Risky	(4) Safe	(5) Risky	(6) Safe	(7) Risky	(8) Safe
Crisis	-0.011**	0.021***	-0.023***	0.025***	-0.011	0.007***	-0.134***	0.068***
	(0.005)	(0.001)	(0.006)	(0.001)	(0.01)	(0.002)	(0.011)	(0.003)
Active			0.022***	0.002				
			(0.008)	(0.002)				
Crisis x Active			0.021*	-0.013***				
			(0.012)	(0.003)				
Large					-0.030***	-0.005***		
					(0.009)	(0.002)		
Crisis x Large					0.024*	0.008***		
					(0.013)	(0.003)		
Stock							-0.022**	0.020***
							(0.009)	(0.002)
Crisis x Large							0.147***	-0.064***
							(0.012)	(0.003)
Constant	-0.015***	-0.005***	-0.023***	-0.006***	-0.01	-0.002	0.011	-0.021***
	(0.003)	(0.001)	(0.004)	(0.001)	(0.006)	(0.001)	(0.007)	(0.002)
Observations	125,041	125,041	78,660	78,660	64,346	64,346	72,624	72,624
R^2	0	0.003	0.001	0.004	0	0.001	0.003	0.011

Table 3.8: Performance result by asset category

Euro gains are defined in the following way. We compute for each individual i: *Euro Gains* $= \frac{52}{W^j \times T} \sum_{n=1}^{N} \sum_{t=1}^{T} Q_{j,n,t} \times P_{n,t} \times AR_{n,t,t+1}$, where N is the number of assets, T is the number of weeks, $Q_{i,n,t}$ is the cumulative number of units of asset n purchased/sold by individual i since the beginning of the period, $P_{n,t}$ is the price of asset n in week t and $AR_{n,t,t+1}$ is the return of asset n from week t to $t+1$, net of the market return over the same period. W^j is the value of individual j's portfolio at the beginning of the period. Euro gains are computed for each individual over the pre-crisis and the crisis period and are then winsorized at the 1% level. The coefficient displayed on this table correspond to the μ coefficient (row 1) and the θ coefficient (row 2, 3 and 4) of the regression equations presented in Section 3.3.3, when estimated separately for each asset class. Active, Large and Stock are dummy variables corresponding to the categories defined in Section 3.2.2. *, **, and *** means statistically different from zero at 10, 5 and 1% level of significance.

	Euro Gains				
	(1)	(2)	(3)	(4)	
CRISIS-PRE	Bonds	Bond MF	Equity MF	Stocks	ETFs
All	-0.0035	-0.0000	0.0003	0.0048	-0.0001
	(-0.62)	(-16.34)***	(1.74)*	(3.67)***	(-11.77)***
Active-Inactive	-0.0021	0.0000	0.0037	0.0470	-0.0001
	(-0.92)	(4.33)***	(7.84)***	(13.08)***	(-10.09)***
Large-Small	-0.0030	-0.0000	0.0026	0.0170	-0.0001
	(-1.21)	(-5.35)***	(6.47)***	(4.53)***	(-5.53)***
Stock-MF	-0.0065	0.0001	-0.0076	0.0330	0.0002
	(-2.20)**	(26.77)***	(-15.30)***	(8.56)***	(11.38)***

Chapter 4

Households Learning in the Dark: Evidence from Retail Traders

Abstract

This paper develops the idea that households have an imprecise knowledge of their portfolio's exposure to systematic risk and that this leads them to make investment mistakes. This idea is tested in the context of the decision to actively trade rather than passively invest in the stock market. I show that the trading activity of individual investors increases (decreases) following high (low) gross performance. I carefully split individual investors performance into (i) the component of their performance related to the exposure of their trades to systematic risk factors and (ii) residual performance and show that their trading activity reacts to both. To account for these results, I contrast a story based on overconfidence against a simple model where individual investors have an imprecise knowledge of *both* their ability and systematic exposure ex ante, and learn about them as they trade. This model generates predictions consistent with the above evidence as well as additional predictions that are also borne out by the data.

JEL classification: D10, G11 Keywords: Learning, individual investor behavior, individual investor performance, household finance

4.1 Introduction

Financial planning is a challenge which households seem to handle unequally. Recent studies have shown that they invest sub-optimally due to a number of behavioral biases (Stango and Zinman, 2009) or lack of financial literacy (Van Rooij et al., 2007; Lusardi and Mitchell, 2008; Lusardi and Tufano, 2009). Other contributions to this literature argue that at least some households display a high level of rationality and sophistication (Calvet et al., 2006, 2009a,b). The reasons why some households do not seem to manage their financial wealth optimally is an important question with significant welfare implications. This paper shows that some investment mistakes might result from the imprecise knowledge that households have of the exposure of their portfolio to systematic risk.

I consider the decision to actively trade, rather than passively invest in the stock market. I use a 10 year long, unique and large sample of French retail traders provided by a leading European broker in personal investing and online trading. To identify the decision to actively trade, I use the "Differed Settlement Service" (henceforth "SRD") a specific feature of the French stock exchange (Euronext), launched in September 2000 and aimed at individual investors. This service enables them to initiate long or short positions on a list of eligible securities only putting a portion of the value of the position forward. The SRD is a service that brokers may or may not offer to their clients and which comes with related fees. The online broker which provided the data used in this paper has been offering this service since it was launched. Since individual investors posting trades on the SRD have to pay a fee to use the service, I am confident that only investors who are willing to actively trade will self select into this service.

Consistent with the literature, I observe that individual investors increase their active trading activity following good performance. More interestingly, I carefully split individual investors performance into (i) performance related to the exposure of their trades to systematic risk factors such as the market premium, the excess returns on small stocks and the excess returns on value stocks (henceforth "factor exposure performance") and (ii) residual performance (henceforth “alpha” or ”residual performance”) and show that trading activity reacts to both (i) and (ii). I ask whether this can be explained by market timing or by the ability of individual investors to predict the returns of systematic risk factors and to modify their exposure to those risk factors accordingly. Instead I find that the exposure of individual investors trades to systematic risk factors (henceforth “beta”) remains remarkably stable through time.

These results suggest that some individual investors are actively trading although they would be better off passively investing. Understanding why this happens is important. It could be that households suffer from a bias of overconfidence and overestimate the signal that past performance sends them about their own ability. This paper develops and tests an alternative story in which these patterns are due to individual investors' imprecise knowledge of their portfolio exposure to systematic risk. I present a simplified three period learning model in which individual investors allocate their wealth between active trading and passive investment in the stock market. When they passively invest in the stock market, they obtain the market return times their beta, i.e. their exposure to market returns, which I assume constant through time. When they actively trade, they obtain alpha, which I assume constant through time, in addition to the market return times their beta. Initially, they do not know alpha and beta. At the end of each period, they observe the market return and their own return and have to decide whether they should actively trade or passively invest in the next period. Ultimately, in the last period, once uncertainty is resolved, they actively trade only if they have a positive alpha. Before then, they have to trade to learn about their alpha. This model generates predictions consistent with the evidence highlighted in the data: active trading is sensitive to both alpha and beta related performance. Moreover, the model generates additional predictions regarding learning dynamics: (i) active trading sensitivity to past performance is smaller when factor returns have been large in absolute value, (ii) the sensitivity of active trading to alpha and factor exposure performance respectively increases and decreases as individual investors learn, and (iii) individuals with high factor exposure performance early in their active investment career stay active longer than individuals with low factor exposure performance early in their active investment career, independently of their respective alpha. I provide evidence supporting these three additional predictions in the sample.

Some individuals investors lose money by actively trading for too long despite a negative alpha or by quitting too early despite a positive alpha. Hence, understanding how they take their investment decision is important: if they trade to learn, then tools properly designed to help households figure out their risk exposure would probably save them time and money.

The learning model presented in the paper is close to those used by Mahani and Bernhardt (2007) and Linnainmaa (2011) who show that individuals find it optimal to trade even if they expect to lose money, as long as the expected short-term loss from trading is offset by the expected gain from learning. These papers motivate the fact that people have to trade to learn about their ability rather than run simulations before they start to trade ("paper learn") as follows. First, a major technical issue with active trading simulation is the difficulty to precisely estimate the actual price at which any given asset would actually be bought or when a limit order would actually be executed since order book data is hard to collect in real time. Second, investing real money is a strong commitment device to provide the best efforts while "paper learning" may lack one.

The addition of this paper to the frameworks developed in these two papers is the assumptions that individual investors performance depends on the exposure of their trades to factor returns and that they initially don't know their exposure to factor returns, i.e. their beta. The assumption that individual investors have imprecise knowledge of their portfolio beta and have to learn about it is consistent with survey evidence on individual investors, see for instance Glaser and Weber (2007a, 2009). It is also the assumption in a number of finance models. In Chevalier and Ellison (1997, 1999), for instance, mutual fund investors are shown to react to market-adjusted performance and less so to beta-adjusted performance. This suggests that individual investors use 1 as a benchmark for the beta of their investments.

Individual investors could ignore their true beta when they start trading either because they are financially illiterate or because it would be more costly for them to invest in the technology to compute beta than to trade for a while and incur the loss associated with negative alpha. Unfortunately, I'm unable to differentiate these two explanations in my data. Finally, recent findings in this field show that individual investors have preferred trading habitat, which means that they trade a limited number of assets with similar characteristics (Dorn and Huberman, 2010; Kumar et al., 2011).

This paper contributes to several strands of the literature. It first adds to the recent academic interest for the role of investors' experience and learning process on financial outcome. The fact that investors' own experience shape their future decisions has been shown in the context of IPOs (Kaustia et al., 2008; Chiang et al., 2011), retirement savings decisions (Choi et al., 2009) and mutual funds management (Greenwood and Nagel, 2009) with potential asset pricing implications. More generally, Malmendier and Nagel (2011) argue that individuals' experience of macro-economic outcome have long-term effects on their risk taking, which has an impact on aggregate stock price dynamics.

Learning may occur in a variety of ways. Investors may gradually discover the true value of model parameters by rationally updating their priors in a Bayesian way after each action (Pastor and Veronesi, 2009). Investors may as well update their beliefs in a non-Bayesian way as in Gervais and Odean (2001). Learning by doing is another way to model the evolution of investors' behavior (Nicolosi et al., 2009; Seru et al., 2009). This kind of learning is not the focus of this paper. Seru et al. (2009) show that individual investors behavior is consistent with both learning about one's ability and learning by doing but that the former is much more quantitatively significant than the later. Finally, List (2003), Agarwal et al. (2008) and Kaustia et al. (2008) show in various frameworks that with experience, investment behaviors tend to get closer to what full rationality would command. Finally, close to this paper, Kézdi and Willis (2011) find that households have different incentives to invest in the acquisition of financial knowledge, which determines their beliefs and, in turn, their stock market behavior.

This paper also adds to the literature on retail traders' behavior. A number of stylized facts have been established by this literature. On average, the average household trades in excess of what liquidity and hedging motives would command and loses money in the process (Barber and Odean, 2000; Barber et al., 2006; Grinblatt and Keloharju, 2000) especially when going online (Barber and Odean, 2002). This is generally attributed to behavioral biases such as overconfidence or gambling (Statman et al., 2006; Glaser and Weber, 2007b; Grinblatt and Keloharju, 2009; French, 2008). Some retail traders however manage to generate absolute performance (Barber et al., 2006; Grinblatt et al., 2009), with some persistence (Coval et al., 2005). Individual investors usually react to past performance by increasing their subsequent trading activity (Glaser and Weber, 2009; Nicolosi et al., 2009). Finally, individual investors have preferred trading habitat, which means that they trade a limited number of assets with similar characteristics (Dorn and Huberman, 2010; Kumar et al., 2011).

Although it is tested on a specific type of investment choice, namely the decision to actively trade on the stock market, the idea developed in this paper is likely to apply to various alternative contexts. It could apply, for instance, to the behavior of retail mutual fund investors.

The rest of this paper is organized as follows. Section 4.2 develops the simple trading model. In section 4.3, I define the measures of performance and trading activity used in the empirical analysis. Section 4.4 presents the data and section 4.5 describes the results. Predictions regarding learning dynamics are tested in section 4.6. Section 4.7 concludes.

4.2 A simple trading model

In this section I present a simplified three period learning model which predicts that active trading is sensitive to past performance. More specifically, it predicts that active trading is sensitive to *both* performance related to the exposure to systematic risk *and* residual performance. Moreover, it delivers additional predictions which distinguishes it from a story based on overconfidence.

Individual investors allocate their wealth between active trading and passive investment in the stock market with only one risk factor, the market factor. When individual investors passively invest in the stock market returns, they obtain the market return times their beta, i.e. their exposure to the market, which I assume constant through time. When they actively trade, they obtain alpha, which I assume constant through time, in addition to the their market exposure performance. Initially, they do not know alpha and beta. At the end of each period, they observe the market return and their own returns and have to decide whether they should actively trade or passively invest in the next period. Ultimately, in the last period, uncertainty is resolved and they actively trade only if they have a positive alpha. Before then, they have to trade to learn about their alpha. The model time line is presented in figure 4.1 and the formal solution of the model is derived in appendix.

Model set up

Individuals are risk neutral[1] and live for three periods. At the beginning of period 1, they have an initial wealth of 1 to invest in the stock market. At each period t individual i gets returns $1 + R_{it}$. Since R_{it} is small[2], I make the simplifying assumption that individual i maximizes:

$$E(\sum_{t=1}^{3} R_{it}) = \sum_{t=1}^{3} E(R_{it}) \tag{4.1}$$

Individual i's returns are characterized by factor exposure performance (beta) and excess performance (alpha). Beta is assumed to be constant across three periods, consistently with evidence that retail traders usually focus on a limited number of stocks with similar characteristics. Results would be unaffected if exposure was correlated over time. I provide evidence in that this is indeed the case in the empirical section of this paper.

[1]Preliminary computations show that results are unaffected in a CARA or CRRA framework

[2]In the empirical sections of the paper, monthly returns are used

Hence the exposure of her portfolio to the market factor at time t is measured by:

$$\beta_i \sim N(1, \sigma_\beta)$$

Individual i's alpha is also assumed to be fixed over time. Results would be unaffected if ability was correlated over time. Hence individual i's ability is given by:

$$\alpha_i \sim N(0, \sigma_\alpha)$$

α_i and β_i are independent.

The market factor return at time t which is called M_t:

$$M_t \sim N(0, \sigma_M)$$

M_t and M_{t+1} are independent.

To invest her wealth in the stock market, individual i chooses each period between:

- A fund which returns the market exposure performance of her portfolio :$\beta_i M_t$
- Active trading which returns the market exposure performance of her portfolio as well as alpha, the excess performance which she is able to generate: $\alpha_i + \beta_i M_t$

At the end of each period t, individual i 's information set, θ_{it}, contains the history of realizations of M_t and R_{it}.

Hence uncertainty on α_i and β_i is resolved at the end of the second period if individual i has actively traded at least once: she just has to solve a system of two equations with two unknown.

I present the time line of the model in figure 4.1. At each node, individual i chooses between actively trading (arrows going up) or passively investing (arrows going down) depending on her information set (described in boxes).

Solving the model

As apparent in figure 4.1, this model may be solved backward for the decision at each node, starting from period 3. I solve it in appendix and present the main intuitions here.

Individual investors always trade in period 1 to seize the option value of actively trading. In other words, there is more upside from trading that not trading in period 1: as you trade in period 1, you obtain more information about your market exposure and your alpha. Just like in any one-armed bandit problem, there is a value to exploration that drives all individual investors into trading in period 1.

In period 2 individual i actively trades only if her expectation of alpha conditional on the information she obtained from trading in period 1 is positive. From her active trading in period 1, she observes the market return, m_1 and her own active trading return r_{i1}. She knows that r_{i1} is the sum of her alpha plus her beta times m_1. She also knows that whether or not she trades in period 2, she will observe m_2 and r_{i2} after trading in period 2, so that she will be able to infer her true alpha and trade accordingly in period 3. Hence whether or not she trades in period 2 will not affect her decision for period 3. Hence she trades in period 2 only if her expected return from doing so is larger than passively invest, that is if her expectation of alpha conditional on her information set at time 1, $E(\alpha_i/\theta_{i1})$ is superior to 0.

Finally, in period 3, uncertainty about α_i and β_i is resolved: since individual i has actively traded for at least one period she just has to solve a system of two equations with two unknowns to compute α_i and β_i . Hence positive alpha individual investors actively trade in period 3 while negative alpha individual investors passively invest.

Predictions on the sensitivity of active trading to past performance

In this paragraph I present the predictions generated by this model. In particular, I focus on the decision that individual investors take at the beginning of period 2, after having observed market returns and their own active trading returns in period 1. At this point, individuals i has observed the realizations of the market m_1 and her own return r_{i1}. She decides to actively trade in period 2 only if:

$$E(\alpha_i/m_1, r_{i1}) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + m_1^2\sigma_\beta^2}(r_{i1} - m_1) > 0$$

I average this expression across groups and plot it as a function of m_1. I derive the following predictions both analytically in appendix and graphically in figures 4.2, 4.3 and 4.4, by using sample estimates of σ_α and σ_β.

Prediction 1: individual investors active trading decision is sensitive to their past performance.

Intuitively, for any realization m_1, individual investors who obtain a performance $r_{i1} > 0$ are more likely to have high alpha. Hence they are more likely to have $E(\alpha_i/m_1, r_{i1}) > 0$ and hence to trade in period 2. Figure 4.2 presents, for each realization m_1 of M_1, the average of $E(\alpha_i/m_1, r_{i1})$ for (i) individuals who experienced positive performance and (ii) individuals who obtained negative performance in period 1. It is clear that on average, individual investors who obtained positive performance will trade in period 2 while others won't. This is simply because those that had positive performance in period 1 are also those that are more likely to have high alpha on average.

Prediction 2: individual investors active trading decision is sensitive to their alpha.

Intuitively, individual investors with higher alpha are more likely to have $E(\alpha_i/m_1, r_{i,1}) > 0$ and hence to trade in period 2. Figure 4.3 presents, for each realization m_1 of M_1, the average of $E(\alpha_i/m_1, r_{i1})$ for (i) individuals who have a high alpha and (ii) individuals who have a negative alpha. It is clear that on average, individual investors who have a high alpha will trade in period 2 while others won't.

Prediction 3: individual investors active trading decision is sensitive to their factor exposure performance.

Recall that individual investors know nothing about their β_i to start with, but the distribution of β_i which is normal and centered around 1. Suppose they experience a very large positive market return m_1 in period 1. Investors with high β_i will obtain a very high r_{i1} while investors with low β_i will obtain a very low r_{i1}. Then investors with high β_i are more likely than low β_i to have $E(\alpha_i/m_1, r_{i1}) > 0$. Hence high β_i are on average more likely than low β_i to actively trade in period 2 in this case, regardless of their true α_i. Figure 4.4 presents, for each realization m_1 of M_1, the average of $E(\alpha_i/m_1, r_{i1})$ for (i) individuals who have a beta larger than 1 and (ii) individuals who have beta lower than 1. It is clear that on average, individual investors who have a high (low) beta will actively trade in period 2 when m_1 has been positive (negative).

4.3 Defining individual investors' active trading and performance

In this section, I lay out the methodology used to measure active trading and individual investors active trading performance.

Measuring individual investors activity

This paragraph presents the way active trading is measured in the rest of this paper.

Identifying active trading from standard buy-and-hold strategies is a non trivial task. The literature has used different method to do so, such as for instance restricting the analysis to round-trip trades (Linnainmaa, 2011). Instead, I identify active trading using the "Differed Settlement Service" (henceforth "SRD") a specific feature of the French stock exchange (Euronext) and aimed at individual investors to enable them to initiate long position or short positions on a list of eligible securities only putting a portion of the value of the position forward, with a settlement of positions at the end of the month. In other words, the main reasons why an individual might use this service is the ability to short sell and the ability to lever up. The SRD is a service that brokers may or may not offer to their clients and which comes with related fees. The online broker which provided the data used in this paper has been offering this service since it was launched. Since individual investors posting trades on the SRD have to pay a fee to use the service, I am confident that only investors who are willing to actively trade will self select into this service. In particular, the SRD is widely used among individual investors. In the data, the SRD is used for 46% of stock trades and 64% of volume of French stocks.

Euronext launched the SRD in September 2000. As detailed in Foucault et al. (2011), the SRD was introduced as part of a reform of the French stock exchange related to European harmonization. Before September 2000, individual investors could trade large stocks on a future markets and small stocks on the spot market. After September 2000, individual investors can trade all stocks spot and can use the SRD service to take positions on eligible stocks. To be eligible to the SRD, a stock must either have a market cap larger than 1 billion euros and a daily volume of at least 1 million euro, or belong to the SBF 120, the French index of the 120 largest stocks. There were 120 eligible stocks initially, and 233 as of 2010. Individual investors using the SRD can take positions until the fifth business day before the end of the month. When an individual reaches the end of the month, either she has liquidated all her positions and she will receive the difference between the selling and the buying prices. Either she still has some nonzero positions in various stocks and she can choose to (i) liquidate these positions or (ii) roll them over to the next month in exchange of a "roll-over fee".

The fees that brokers usually charge has two components on top of usual brokerage fees. First, 0.02% to 0.03% are charged each day on the amount held through the SRD. Second, brokers charge a roll-over fee of 0.2% to 0.3% on the amount that individual investors wish to roll-over to the next month.

The amount of leverage that individual investors can take depends on their outstanding portfolio, which serves as a collateral to their purchases through the SRD. The rules for the collateral are set by the French regulator (AMF)[3]. Individual investors are allowed to lever up to five times the amount of cash and French government bonds and monetary funds they hold, four times the amount of bonds (traded on any EU regulated market, other European government bonds and bond funds) they hold and two and a half times the stocks (traded on any EU regulate market) and European equity funds they hold.

Restricting the sample to SRD trades is particularly well suited to capture active trading for a variety of reasons. First, investors using the SRD are very unlikely to do so for liquidity or hedging motives as it comes with additional fees. Hence we are left exclusively with active traders intending to generate performance by exploiting private information or acting as liquidity providers by taking leverage or shorting stocks. Moreover, the monthly settlement procedure alleviates liquidity constraints, since investors do not need to actually pay for the assets they trade before the date of the monthly settlement. This mitigates the concern that inferences may be biased by the fact that past performance influences subsequent trading by relieving or increasing liquidity constraints. Also, the monthly settlement procedure generates a clear sequence of actions and outcomes which is ideally suited to test any learning model. At the end of each month, retail investors may close their positions or roll them over in exchange for a "roll-over fee". Hence there is no doubt that individual investors using the SRD do analyze their positions and performance at the end of each month to take subsequent decisions. Finally, in figure 4.5, I present the investment objective of 833 individuals in the dataset either using the SRD service or not using it. SRD traders answer "Directly investing in the stock market" approximately 70% of the time, which is twice as often as non SRD traders. This confirms that the main characteristic of SRD traders is the desire to actively trade in the stock market.

I track the "active trading" career of individuals from the first day they start using the SRD. I stop tracking the "active trading" career of individuals if they do not post any trade using the SRD for three successive months. In other words, I am not keeping in the sample

[3]However each financial intermediary is free to request higher minimum collateral rates.

months during which an individual investor uses the SRD after she has not posted any trades using the SRD for three successive months. However, two thirds of individual investors in the sample never use the SRD again once they stop trading for three successive months.

I measure the intensity of individual investor i 's trading activity in month t using a variety of variables: the log of the number of orders, the log of the euro volume traded, the log euro volume per stock traded, the excess volume purchased over volume sold normalized by total volume, the log of the average order size and a continuation dummy equals to one if individual i is active in month $t+1$ or if individual i stops trading.

Measuring individual investors performance

Measuring investors' stock performance is uneasy for at least three reasons. First, the horizon at which the performance should be measured is unclear and largely varies across studies. Second, it is unclear how performance should be computed when investors hold on to their positions and do not unwind them, since the time when positions are liquidated contributes to the overall performance of the trade. Third, investors may build up and liquidate positions in a sequential way which makes it difficult to know which entry price and exit price to consider for the trade.

The framework I use enables me to handle the first point. The horizon of a trader using the SRD service is clearly the end of the month, where she'll have to decide whether to roll-over her position (at a cost) or not. To handle the second and third point, I apply the approach used by Linnainmaa (2011) to measure the monthly performance of individual investors in my sample.

For each stock s traded by individual i in month t, I computed its performance $\pi_{i,s,t}$ as follows:

$$
\begin{aligned}
&\pi_{i,s,t} \\
&= p^s_{i,s,t}/p^b_{i,s,t} - 1, \text{ if } v^s_{i,s,t} = v^b_{i,s,t} \\
&= (p^s_{i,s,t}/p^b_{i,s,t} - 1) \times v^s_{i,s,t}/v^b_{i,s,t} + (p^c_{s,t}/p^b_{i,s,t} - 1) \times (v^b_{i,s,t} - v^s_{i,s,t})/v^b_{i,s,t}, \text{ if } v^s_{i,s,t} < v^b_{i,s,t}
\end{aligned}
$$

$= (p^s_{i,s,t}/p^b_{i,s,t} - 1) \times v^b_{i,s,t}/v^s_{i,s,t} + (p^s_{i,s,t}/p^c_{s,t} - 1) \times (v^s_{i,s,t} - v^b_{i,s,t})/v^s_{i,s,t}$, if $v^s_{i,s,t} > v^b_{i,s,t}$

where n is the number of stocks investor i trades in month t, $p^b_{i,s,t}$ and $p^s_{i,s,t}$ are the investor's average purchase and sale prices in stock s, where the purchase and sale price of stock s in day d is the closing price of stock s in day d. $v^b_{i,s,t}$ and $v^s_{i,s,t}$ are the number of shares purchased and sold, and $p^c_{s,t}$ is the stock's price on the last day of the trading month. If an investor buys and sells different amounts, the remaining position is marked-to-market at the end of the trading month.

From there I compute an unweighted and a weighted measure of individual i 's performance in month t , $\pi_{i,t}$ and $\pi^w_{i,t}$ as

$$\pi_{i,t} = \tfrac{1}{n} \textstyle\sum_{s=1}^{n} \pi_{i,s,t}$$

$$\pi^w_{i,t} = \textstyle\sum_{s=1}^{n} \tfrac{vol_{i,s,t}}{vol_{i,t}} \pi_{i,s,t}$$

where $vol_{i,t}$ and $vol_{i,s,t}$ are respectively the total euro volume and the euro volume of stock s traded in month t by individual i. In the paragraphs that follow, I will only consider $\pi_{i,t}$ to save notations. But all empirical results will be run using both weighted and unweighted measures of performance.

At the end of each trading month, if they have a still unwound long (short) position in a stock, investors have the choice between obtaining delivery (payment) of the stock or rolling the position over to the next trading month. One limitation of the data is that I do not observe whether a position is rolled-over. What I observe, however, is whether an investor pays any roll-over fees in any given month. Hence, I consider that investors (not) paying roll-over fees in a given month roll-over all (none of) their positions to the following month. Rolled-over positions are then valued at the closing price of the first day of the next month and included in the computation of the next month performance.

Measuring the alpha and factor exposure performance of individual investors

This paper intends to isolate the performance of active trading by individual investors related to usual risk factors: Mkt, the market return minus the risk free rate, SMB (the excess return of small firms) and HML (the excess return of value firms). The method used here compares the performance of an individual i in month t trading n stocks s to the performance she would have obtained by trading n indices replicating the exposure of each stock s to each of the three factors $f \in \{Mkt, SMB, HML\}$ instead. I do this in two steps. I first compute the exposure (beta) of each stock s in month t to each factor f. Then I compute the performance individual i would have obtained if she had traded n indices replicating the returns of each risk factor instead of n stocks.

Computing risk factor exposure for each stock

I first construct the series of daily returns of each of the three factors. Mkt is the daily value weighted average return on all stocks traded by individual investors, minus the interest rate on French 10 year treasury bonds. SMB is computed by sorting firms according to the past year market capitalization. Big firms are the 20% largest firms and small firms are the 20% smallest ranked in the last month of the previous year. To determine SMB, I subtract, each month, the value weighted monthly returns of the largest firms from the value weighted monthly returns of the 20% largest firms. To compute HML, I sort firms by past year book-to-market value of assets. "Value firms" are firms with the 20% highest book-to-market in the previous year, and "glamor," firms with the lowest 20%. HML is the difference in value-weighted monthly returns between the value and the glamor portfolio.

Then in each month t, for each stock s, I run the following Fama French three factor model:

$$R_{s,d} = \beta_{s,t}^{Mkt} \times Mkt_d + \beta_{s,t}^{SMB} \times SMB_d + \beta_{s,t}^{HML} \times HML_d + \epsilon_{s,d}$$

where $R_{s,d}$ is the return of stock s minus the risk free rate in day d, Mkt_d, SMB_d and HML_d are the daily returns of each of the three factors. The coefficient obtained for each factor f, $\beta^f_{s,t}$, measures the sensitivity of the returns of stock s to factor f as measured in month t.

Replicating performance with factor indices

Then for each stock s and each factor $f \in \{Mkt, SMB, HML\}$, I compute $\pi^f_{i,s,t}$, the profit individual investor i would have realized in month t if instead of trading stock s she had traded an index replicating the returns of factor f.

$\pi^f_{i,s,t}$

$= i^s_{i,f,t}/i^b_{i,f,t} - 1$, if $v^s_{i,s,t} = v^b_{i,s,t}$

$= (i^s_{i,f,t}/i^b_{i,f,t} - 1) \times v^s_{i,s,t}/v^b_{i,s,t} + (i^c_{f,t}/i^b_{i,f,t} - 1) \times (v^b_{i,s,t} - v^s_{i,s,t})/v^b_{i,s,t}$, if $v^s_{i,s,t} < v^b_{i,s,t}$

$= (i^s_{i,s,t}/i^b_{i,s,t} - 1) \times v^b_{i,s,t}/v^s_{i,s,t} + (i^s_{i,s,t}/i^c_{s,t} - 1) \times (v^s_{i,s,t} - v^b_{i,s,t})/v^s_{i,s,t}$, if $v^s_{i,s,t} > v^b_{i,s,t}$

where n is the number of stocks investor i trades in month t, $i^b_{i,f,t}$ and $i^s_{i,f,t}$ are the investor's average purchase and sale prices in the index replicating the performance of factor f, $v^b_{i,s,t}$ and $v^s_{i,s,t}$ are the number of shares purchased and sold of stock s, and $i^c_{f,t}$ is the index's price on the last day of the month. If an investor buys and sells different amounts, the remaining position is marked-to-market at the end of the trading month.

From there I get for each individual i in each month t its factor exposure performance for each factor $f \in \{Mkt, SMB, HML\}$, $\beta\pi^f_{i,t}$, computed as:

$$\beta\pi^f_{i,t} = \tfrac{1}{n} \textstyle\sum_{s=1}^{n} \beta^f_{s,t-1\times}\pi^f_{i,s,t}$$

I also compute the weighted factor exposure performance for each factor f using the same weights as above, i.e. the ratio of the euro volume traded of stock s in month t to the total euro volume traded in month t.

For each individual i in month t, I compute the average beta of individual investor i's trades in month t with respect to each of the three factors f , $\beta^f_{i,t}$, as follows:

$$\beta_{i,t}^{f} = \frac{1}{n}\sum_{s=1}^{n}\beta_{s,t-1}^{f}$$

I also compute a weighted version of that beta for each factor f using the same weights as above, i.e. the ratio of the euro volume traded of stock s in month t to the total euro volume traded in month t.

Finally, I recover the residual performance of each individual i ("alpha"), in each month t stripped from the factor exposure performance, $\alpha_{i,t}$ as follows:

$$\alpha_{i,t} = \pi_{i,t} - \beta\pi_{i,t}^{Mkt} - \beta\pi_{i,t}^{SMB} - \beta\pi_{i,t}^{HML}$$

It should be noted that alpha is not a pure measure of excess performance and also absorbs some noise. It should be interpreted as the share of individual i's performance not attributable to systematic risk factors. I also compute the weighted alpha using the same weights as above, i.e. the ratio of the euro volume traded of stock s in month t to the total euro volume traded in month t.

In the empirical part that follows, I will measure the sensitivity of individual i active trading in month t to the various components of her performance.

4.4 Data

I use a 10 year long, unique and large sample of French retail traders provided by a leading European broker in personal investing and online trading.

From 1999 to 2010, this broker accounted for an average 15% of online brokers trades on Euronext Paris, which collectively represented 14% of all trades in the market[4]. This broker provided the complete daily trading records of their active retail client base from January 1999 to December 2010. For a sub-sample of 833 traders, the answers to a regulatory ("Mifid") questionnaire are available, with information regarding wealth and investment objectives.

[4]According to Acsel, an association of online brokers (see http://www.associationeconomienumerique.fr/) which collects monthly data on online trading

To identify the decision to actively trade, I use the "Differed Settlement Service" (henceforth "SRD") a specific feature of the French stock exchange (Euronext), launched in September 2000 and aimed at individual investors. When restraining individual investors actively trading using the SRD, I am left with 14,151 individual investors.

I obtain exhaustive and reliable market data from EUROFIDAI, which enables me to compute accurate stock and factor returns. I obtain book-to-market ratios from Datastream.

Summary statistics are presented in table 4.1. There are 14,151 distinct individual investors actively trading in 7.7 successive periods on average. On average, these individual investors post 6 trades per month using the SRD, with an average volume traded of 21,590 euros and an average trade size of 3,463 euros. Note that there are very few short positions observed in the dataset, although the cost of taking a short position through the SRD is symmetric to the cost of taking a long position. Hence individual investors seem to use the SRD to take levered position rather than take sort positions. Note that there is no way to make sure that individual investors in this sample are not using another broker before, after or contemporaneously to their investment career as observed in this sample.

4.5 Sensitivity of trading to past performance

In this section I first check that, consistent with existing literature, individual investors' active trading in the sample is sensitive to their past performance. Then I ask to what extent the activity of individual investors is sensitive to the various components of their performance: alpha and factor exposure performance.

Individual investors trading is sensitive to past performance

The first test I run intends to check that, consistent with previous literature, individual investors' active trading is sensitive to past performance. The following regression is run:

$$T_{i,t+1} = \lambda_1 \pi_{i,t} + \lambda_2 P_{i,t} + \kappa_i + \delta_e + \mu_t + \epsilon_{i,t}$$

where $T_{i,t+1}$ is a measure of trading activity of individual i in month t. I consider successively the log of the number of orders, the log euro volume traded, the log euro volume per stock traded, the excess volume purchased over volume sold normalized by total volume, the log of the average order size, and a dummy equals to one if individual i is active in month $t+1$ or if individual i stops trading. $\pi_{i,t}$ is the performance of individual i in month t as computed in section 4.3, $P_{i,t}$ is the increase in the log euro value of individual i's portfolio in month t (including positions taken on the spot market) and κ_i, δ_e and μ_t are respectively individual, experience (the number of trading month since individual i started active trading) and month fixed effects. Standard errors are clustered at the monthly level. The same regression is run with the weighted performance of individual i in month t where the performance on each stock traded by individual i in month t is weighted by the ratio of the euro volume traded of stock s in month t to the total euro volume traded in month t. Table 4.2 presents the results of this regression. All measures of trading activity are significantly positively correlated with past performance. A one standard deviation in performance increases the probability of continuation by 0.5%.

The second test intends to disentangle the sensitivity of trading to the various components of individual investors' performance. The following regression is run:

$$T_{i,t+1} = \lambda_1 \alpha_{i,t} + \lambda_2 \beta\pi_{i,t}^{Mkt} + \lambda_3 \beta\pi_{i,t}^{SMB} + \lambda_4 \beta\pi_{i,t}^{HML}$$
$$+\lambda_5 \pi_{i,t}^{Mkt} + \lambda_6 \pi_{i,t}^{SMB} + \lambda_7 \pi_{i,t}^{HML} + \lambda_8 \beta_{i,t}^{Mkt} + \lambda_9 \beta_{i,t}^{SMB} + \lambda_{10} \beta_{i,t}^{HML} + \lambda_{11} P_{i,t} + \kappa_i + \delta_e + \mu_t + \epsilon_{i,t}$$

where $T_{i,t+1}$ is a measure of trading activity of individual i in month . I consider successively the log of the number of orders, the log euro volume traded, the log euro volume per stock traded, the excess volume purchased over volume sold normalized by total volume, the log of the average order size, and a dummy equals to one if individual i is active in month $t+1$ or if individual i stops trading. $\beta\pi_{i,t}^{Mkt}$, $\beta\pi_{i,t}^{SMB}$and $\beta\pi_{i,t}^{HML}$ are the returns obtained from the exposure of individual i's trades to the three risk factors. The third component of the regression includes the following controls. $\pi_{i,t}^{Mkt}$, $\pi_{i,t}^{SMB}$and $\pi_{i,t}^{HML}$ are the returns on the indices replicating the risk factors. $\beta_{i,t}^{Mkt}$ $\beta_{i,t}^{SMB}$and $\beta_{i,t}^{HML}$ are the average exposure of the stocks traded by individual i to each of the three risk factors. $P_{i,t}$ is the increase in the log euro value of individual i's overall portfolio in month t (including positions taken on the spot market) and κ_i, δ_e and μ_t are respectively individual, experience (the number of trading month since individual i started active trading) and month fixed effects. Standard errors are clustered at the monthly level. The same regression is then run with variables weighted by the ratio of the euro volume traded of stock s in month t to the total euro volume traded in month t.

The coefficients of interest are λ_1, λ_2, λ_3 and λ_4. Results are presented in table 4.3. Surprisingly, perhaps, all measures of investor trading strongly react to both alpha and factor exposure performance. Coefficients on factor exposure performance are even slightly higher than coefficients on alpha. Results are of the same magnitude than those in table 4.2.

Do investors time the market?

This section investigates whether the result obtained in the previous section could be explained by individual investors ability to forecast factor returns. Suppose that factor returns are correlated from one period to another. Knowing this, individual investors in month t could guess what factor returns will be in month $t+1$ and adjust their factor exposure accordingly to benefit from higher factor exposure performance. For instance, suppose that market returns are auto-correlated. Hence if you observe high market returns in month t you can extrapolate that market returns will be high in month $t+1$ and decide to increase your exposure to market returns and trade more in month $t+1$.

First, I check whether factor returns are auto-correlated. To do so, I compute the monthly auto-correlation in the returns of each of the three risk factors: Mkt, SMB and HML. It appears that Mkt and SMB are not auto-correlated and that HML is weakly auto-correlated. These correlations are presented in table 4.4. From there, it seems implausible that any individual can reliably infer month $t+1$ factor returns by observing month t factor returns.

However, individuals may be able to forecast factor returns. Hence I check whether their factor exposure moves in the same direction as factor returns. I run the following regression for each factor $f \in \{Mkt, SMB, HML\}$:

$$R_t^f = \lambda_1 \beta_{i,t}^f + \kappa_i + \delta_e + \epsilon_{i,t}$$

where R_t^f is the return of risk factor f in month t. κ_i and δ_e are respectively individual and experience (the number of trading month since individual i started active trading) fixed effects. Standard errors are clustered at the monthly level. Results are presented in table 4.5. The exposure to the market is negatively correlated to the market factor returns, while the exposure to the two other factors are uncorrelated with these factors. Hence individual investors do not seem to strategically adjust their factor exposure to increase their performance.

Finally, several recent papers (Dorn and Huberman, 2010; Kumar et al., 2011) have established that households keep trading a subset of stocks with similar characteristics over time. I check that this translates into a persistence of the characteristics of retail investors actively traded portfolio. I run the following regression for each factor $f \in \{Mkt, SMB, HML\}$:

$$\beta_{i,t+1}^f = \lambda_1 \beta_{i,t}^f + \delta_e + \mu_t + \epsilon_{i,t}$$

where $\beta_{i,t+1}^f$ is the exposure of individual i to factor f in month t. δ_e and μ_t are respectively experience (the number of trading month since individual i started active trading) and month fixed effects. Standard errors are clustered at the monthly level. Results presented in table 4.6 show that the exposure to risk factors is quite stable through time. The auto-correlation coefficient on alpha is positive though not significant, probably because the measure of alpha includes some noise as mentioned above.

4.6 Learning dynamics

In this section, I highlight and test three additional predictions of the model presented above. These predictions are specific to this model and would be harder to justify in an story based on overconfidence. First, I show that the sensitivity of trading to past performance decreases with the absolute value of factor returns. Second, I check whether the sensitivity of trading to past performance evolves through time as it is described in the model: trading should become less sensitive to the three risk factor past performance and more sensitive to the residual performance. Third, I check whether the speed at which individuals learn and decide to quit active trading is influence by their early experience. For instance individuals with low market betas experiencing low stock market returns early in their life keep actively trading for a longer period of time than individuals with high market betas and low stock market returns, everything else equal.

Prediction 4: individual investors' active trading sensitivity to past performance is decreasing in absolute factor returns.

In the simplified trading model presented above, individual investors decide whether to actively trade or not in period 2 if:

$$\mathrm{E}(\alpha_i/m_1, r_{i1}) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}(r_{i1}\text{-}m_1) > 0$$

This expectation is increasing in the value of r_{i1}, the performance in period 1. However, when m_1 has been very high or very low, the sensitivity of the expectation to r_{i1} is lower. Intuitively, if the market return in period 1 has been close to zero, individual investor i infers alpha more easily from the performance she obtained in period 1. Hence a small but positive performance in period 1 will signal a positive alpha and conversely. However, if the market return in period 1 has been very large or very low, individual i has a much harder time inferring alpha from her performance in period 1. Hence her decision is relatively less sensitive to past performance. This prediction can be tested quite easily. I run the following regression in the panel of individual investors:

$$T_{i,t+1} = \lambda_1\pi_{i,t} + \lambda_2 AbsR_t^{Mkt} \times \pi_{i,t} + \lambda_3 AbsR_t^{SMB} \times \pi_{i,t} + \lambda_4 AbsR_t^{HML} \times \pi_{i,t} + \lambda_5 AbsR_t^{Mkt} + \lambda_6 AbsR_t^{SMB} + \lambda_7 AbsR_t^{HML} + \lambda_8 P_{i,t} + \kappa_i + \delta_e + \epsilon_{i,t}$$

where $T_{i,t+1}$ is a measure of trading activity of individual i in month. I consider successively the log of the number of orders, the log euro volume traded, the log euro volume per stock traded, the excess volume purchased over volume sold normalized by total volume, the log of the average order size, and a dummy equals to one if individual i is active in month $t+1$ or if individual i stops trading. $\pi_{i,t}$ is the performance of individual i in month t as computed in section 2, $AbsR_t^f$ is the absolute value of the return of risk factor f in month t, $P_{i,t}$ is the increase in the log euro value of individual i's portfolio in month t, and κ_i and δ_e are respectively individual and experience fixed effects. Standard errors are in parenthesis and are clustered at the monthly level. In Panel B, the same regression is run with the weighted performance of individual i in month t where the performance on each stock traded by individual i in month t is weighted by the ratio of the euro volume traded of stock s in month t to the total euro volume traded in month t. Results in table 4.7 show that consistent with the intuition highlighted in the model, individual investors' trading activity seem to be less sensitive to past performance when past factor returns have been large in absolute value.

Prediction 5: individual investors' active trading sensitivity to alpha decreases with experience while individual investors' active trading decision sensitivity to factor exposure performance decreases.

In the simple three period model presented above, individual i's choice to actively trade in period 2 is sensitive to past factor exposure performance as well as to her alpha. In period 3, her choice to actively trade is fully sensitive to her alpha, but not to her past factor exposure performance anymore. Hence we should observe that the sensitivity of trading activity to past factor exposure performance decreases through time while the sensitivity of trading activity to alpha increases. I define experience as the number of months since individual i started actively trading. Then I run the following regression in the cross-section of actively trading individual investors at each experience month e:

$$T_{i,e+1} = \lambda \alpha_{i,e} + \epsilon_{i,e}$$

$$T_{i,e+1} = \lambda\gamma_{i,e} + \epsilon_{i,e}$$

where $T_{i,e+1}$ is a dummy equals to one if individual i is active in month $e+1$ or zero if individual i stopped trading. $\alpha_{i,e}$ is the alpha of investor i in experience month e and $\gamma_{i,e} = \pi_{i,e} - \alpha_{i,e}$ is the aggregate factor exposure performance over all three risk factors of individual i in experience month e.

I keep the first four months of experience after the initial trading month because half of investors in the sample stop trading on the SRD after 4 months. I then plot the coefficients obtained in these regressions and confidence intervals in figure 4.6. As it turns out, the sensitivity to alpha seems to increase through time while the sensitivity to the aggregate factor exposure performance seems to decrease through time.

Prediction 6: individual investors with high factor exposure performance early in their investment life keep on actively trading longer and exit later, everything else equal.

Another prediction of the model regards the speed at which individual investors decide to stop actively trading. In the model presented above, investors with high alpha are more likely to actively trade. Moreover, high (low) beta investors are more likely to keep on actively trading in period 2 when they have experienced positive (negative) factor returns in period 1, independently of their alpha. I run the following regression on the sample of investors for which I observed an exit before December 2010:

$$S_i = \lambda_1\alpha_{i,1} + \lambda_2\gamma_{i,1}$$

where S_i is the inverse of the number of successive active trading periods of individual i, $\gamma_{i,1} = \pi_{i,1} - \alpha_{i,1}$ is the aggregate factor exposure performance over all three risk factors of individual i in the first month she trades. The regression is run with unweighted and weighted measures of performance. Results presented in table 4.8 provide some evidence, consistent with the model predictions, that the initial factor exposure performance of individual investors has an impact on the length of their active trading career and the speed at which they stop actively trading.

4.7 Conclusion

This paper develops the idea that households have an imprecise knowledge of their portfolio's exposure to systematic risk and that this leads them to make investment mistakes. This idea is tested in the context of the decision to actively trade on the stock market rather than passively invest.

I use a 10 year long, unique and large sample of French retail traders provided by a leading European broker in personal investing and online trading. To identify the decision to actively trade, I use a specific feature of the French stock exchange (Euronext).

Using a number of measures of individual investors trading activity, I observe that individual investors increase their active trading activity following good performance. More interestingly, I carefully split individual investors performance into (i) excess performance and (ii) performance related to the exposure of their trades to usual risk factors such as the market premium, the excess returns on small stocks and the excess returns on value stocks and show that trading activity reacts to both. I ask whether this can be explained by market timing or the ability of individual investors to predict the returns of risk factors and to modify their exposure to those risk factors accordingly. Instead I find that the exposure of individual investors trades to risk factors remains remarkably stable.

I present a simple trading model which accounts for these findings. In the model, individual investors have to decide whether to passively invest or actively trade with limited information on their alpha and beta. It predicts that active trading should be sensitive to both alpha and factor exposure performance. The model also generates additional predictions regarding learning dynamics: (i) active trading sensitivity to past performance is smaller when factor returns have been large in absolute value, (ii) the sensitivity of active trading to alpha and factor exposure performance respectively increases and decreases as individual investors learn, and (iii) individuals with high factor exposure performance early in their active investment career stay active longer than individuals with low factor exposure performance early in their active investment career, independently of their respective alpha. I provide some empirical evidence supporting these three additional predictions.

Since some individuals investors lose money by actively trading for too long despite a negative alpha or by quitting too early despite a positive alpha, understanding how they take their trading decisions is important. If they learn as they trade, then tools properly designed to help households figure out their risk exposure would probably save them time and money. Moreover, whether the fact that individual investors are ex ante uncertain about their ability and risk exposure is the result of financial illiteracy or of the cost associated with computing factor exposure is an important question.

The mechanism highlighted in this paper may be insightful in the context of other households financial decisions. In particular, it may well apply to households investments in mutual funds, a topic that I am currently exploring.

Bibliography

Agarwal, Sumit, John C. Driscoll, Xavier Gabaix, and David Laibson, "Learning in the Credit Card Market," *NBER Working Paper*, February 2008, (13822).

Barber, Brad M and Terrance Odean, "Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors," *The Journal of Finance*, 2000, *55* (2), 773–806.

— **and** — , "Online Investors: Do the Slow Die First?," *Review of Financial Studies*, 2002, *15* (2), 455–487.

— **, Yi-Tsung Lee, Yu-Jane Liu, and Terrance Odean**, "Just How Much Do Individual Investors Lose by Trading?," *Review of Financial Studies*, 2006, *22* (2), 609–632.

Calvet, Laurent E, John Y Campbell, and Paolo Sodini, "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes," *Journal of Political Economy*, 2006, *115* (5), 707–747.

— **,** — **, and** — , "Fight Or Flight? Portfolio Rebalancing by Individual Investors," *Quarterly Journal of Economics*, 2009, *124* (1), 301–348.

— **,** — **, and** — , "Measuring the Financial Sophistication of Households," *American Economic Review*, 2009, *99* (2), 393–398.

Chevalier, Judith and Glenn Ellison, "Risk Taking by Mutual Funds as a Response to Incentives," *Journal of Political Economy*, 1997, *105* (6), 1167–1200.

— **and** — , "Career Concerns of Mutual Fund Managers," *Quarterly Journal of Economics*, 1999, *114* (2), 389–432.

Chiang, Y M, D Hirshleifer, Y Qian, and A E Sherman, "Do Investors Learn from Experience? Evidence from Frequent IPO Investors," *Review of Financial Studies*, 2011, *24* (5), 1560–1589.

Choi, James J, David Laibson, Brigitte C Madrian, and Andrew Metrick, "Reinforcement Learning and Savings Behavior.," *The Journal of finance*, 2009, *64* (6), 2515–2534.

Coval, Joshua D, David A Hirshleifer, and Tyler Shumway, "Can Individual Investors Beat the Market ? Can Individual Investors Beat the Market ?," *Social Science Research*, 2005, *1001* (04), 1–37.

Dorn, Daniel and Gur Huberman, "Preferred risk habitat of individual investors," *Journal of Financial Economics*, 2010, *97* (1), 155–173.

Foucault, Thierry, David Sraer, and David J Thesmar, "Individual Investors and Volatility," *Journal of Finance*, 2011, *LXVI* (4), 1369–1406.

French, Kenneth R, "Presidential Address: The Cost of Active Investing," *The Journal of Finance*, 2008, *63* (4), 1537–1573.

Gervais, Simon and Terrance Odean, "Learning to be overconfident," *Review of Financial Studies*, 2001, *14* (1), 1–27.

Glaser, M and M Weber, "Why inexperienced investors do not learn: They do not know their past portfolio performance," *Finance Research Letters*, 2007, *4* (4), 203–216.

Glaser, Markus and Martin Weber, "Overconfidence and trading volume," *The Geneva Risk and Insurance Review*, 2007, *32* (1), 1–36.

_ **and** _ , "Which past returns affect trading volume?," *Journal of Financial Markets*, 2009, *12* (1), 1–31.

Greenwood, Robin and Stefan Nagel, "Inexperienced investors and bubbles," *Journal of Financial Economics*, 2009, *93* (2), 239–258.

Grinblatt, Mark and Matti Keloharju, "The investment behavior and performance of various investor types: a study of Finland's unique data set," *Journal of Financial Economics*, 2000, *55* (1), 43–67.

_ **and** _ , "Sensation Seeking, Overconfidence, and Trading Activity," *The Journal of Finance*, 2009, *64* (2), 549–578.

_ **,** _ **, and Juhani T Linnainmaa**, "Do Smart Investors Outperform Dumb Investors ?," *Social Science Research*, 2009, (09).

Kaustia, Markku, Eeva Alho, and Vesa Puttonen, "How Much Does Expertise Reduce Behavioral Biases? The Case of Anchoring Effects in Stock Return Estimates," *Financial Management*, 2008, *37* (3), 391–412.

Kumar, Alok, Jeremy Page, and Oliver Spalt, "Investor Clienteles and Habitat-Based Return Comovements," *AFA 2011 Denver Meetings Paper.*, 2011.

Kézdi, Gábor and Robert J. Willis, "Household Stock Market Beliefs and Learning," *NBER working paper*, November 2011, (17614).

Linnainmaa, J T, "Why Do (Some) Households Trade So Much?," *Review of Financial Studies*, 2011, *24* (5), 1630–1666.

List, John A, "Does Market Experience Eliminate Market Anomalies?," *Quarterly Journal of Economics*, 2003, *118* (1), 41–71.

Lusardi, Annamaria and Olivia S Mitchell, "Planning and Financial Literacy: How Do Women Fare?," *American Economic Review*, 2008, *98* (2), 413–417.

_ **and Peter Tufano**, "Debt literacy, financial experiences, and overindebtedness," *NBER working paper*, 2009, *54* (14808), 258.

Mahani, Reza and Dan Bernhardt, "Financial Speculators' Underperformance: Learning, Self-Selection, and Endogenous Liquidity," *The Journal of Finance*, 2007, *62* (3), 1313–1340.

Malmendier, Ulrike and Stefan Nagel, "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?*," *The Quarterly Journal of Economics*, 2011, *126* (1), 373–416.

Nicolosi, Gina, Liang Peng, and Ning Zhu, "Do individual investors learn from their trading experience?," *Journal of Financial Markets*, 2009, *12* (2), 317–336.

Pastor, Lubos and Pietro Veronesi, "Learning in Financial Markets," *Annual Review of Financial Economics*, 2009, *1* (1), 361–381.

Seru, A, T Shumway, and N Stoffman, "Learning by Trading," *Review of Financial Studies*, 2009, *23* (2), 705–739.

Stango, Victor and Jonathan Zinman, "Exponential Growth Bias and Household Finance," *The Journal of Finance*, 2009, *64* (6), 2807–2849.

Statman, M, S Thorley, and K Vorkink, "Investor Overconfidence and Trading Volume," *Review of Financial Studies*, 2006, *19* (4), 1531–1565.

Van Rooij, Maarten, Annamaria Lusardi, and R J M Rob Alessie, "Financial Literacy and Stock Market Participation," *NBER working paper*, 2007, *No. 13565* (April), 1–48.

Appendices

Maths appendix

Solving the model

In the following paragraphs, I solve for the optimal decision at each node of the tree, given the expected information set.

Period 3

In this paragraph I solve for the optimal decision of investor i whether or not to actively trade at the beginning of period 3. After period 2, either individual i has actively traded at least one time in period 1 and 2, or she has never actively traded before.

If she has actively traded once or twice in periods 1 and/or 2, she has enough information to discover α_i and β_i after period 2. Thus uncertainty is resolved. The expected payoff from actively trading in period 3 is given by $\alpha_i + \beta_i E(M_3)$ while the expected payoff from passively investing in period 3 is given by $\beta_i E(M_3)$. Hence individual i will actively trade if $\alpha_i > 0$ and passively invest if $\alpha_i \leq 0$.

If she has never actively traded in periods 1 and 2, individual i knows the value of β_i but she has no information about α_i. The payoff from actively trading is given by $E(\alpha_i) + \beta_i E(M_3) = \beta_i E(M_3)$ which equals the expected payoff from passively investing, $\beta_i E(M_3)$. Hence individual i is indifferent between actively trading and passively investing in period 3.

Period 2

In this paragraph I solve for the optimal decision of investor i whether or not to actively trade at the beginning of period 2. After period 1, either individual i has actively traded in period 1 or not.

If she has actively traded in period 1, she has observed the realizations r_{i1} and m_1 of R_{i1} and M_1 (Let θ_{i1} be her information set at that point). Moreover, she knows that whichever her decision to actively trade or passively invest in period 2, uncertainty will be resolved in period 3 so that she will trade only if $\alpha_i > 0$. Her decision in period 2 has no impact on her expected period 3 payoff. Hence individual i will actively trade if $E(\alpha_1/\theta_{i1}) > 0$ and passively invest if $E(\alpha_i/\theta_{i1}) \leq 0$.

If she has passively invested in period 1, she has observed the realizations $r_{i1}$and m_1 of R_{i1} and M_1,θ_{i1}. With respect to the beginning of period 1, she has no additional information regarding α_i. However, she knows the exact value of β_i. If she actively trades in period 2, uncertainty will be resolved in period 3 so that she will trade only if $\alpha_i > 0$. In this case, she expects to obtain $E(\alpha_i/\alpha_i > 0) + E(\beta_i M_3/\alpha_i > 0) = E(\alpha_i/\alpha_i > 0)$ if she actively trades in period 3 and $E(\beta_i M_3/\alpha_i \leq 0) = 0$ if she passively invests in period 3.

Hence individual i always actively trades in period 2 if she has not traded in period 1, so as to seize the opportunity to trade in period 3 when uncertainty is resolved.

In other words,

If individual i has actively traded in period 1

- The expected payoff from actively trading in period 2 is given by:

$$\begin{aligned}&\{E(\alpha_i/\theta_{i1}) + E(\beta_i M_2/\theta_{i1})\}\\ &\quad + \{P(\alpha_i/\alpha_i > 0)\,[E(\alpha_i/\alpha_i > 0) + E(\beta_i M_3/\alpha_i > 0)] + P(\alpha_i/\alpha_i < 0)E(\beta_i M_3/\alpha_i > 0)\}\\ &\quad = E(\alpha/\theta_{i1}) + \tfrac{1}{2}E(\alpha_i/\alpha_i > 0)\end{aligned}$$

- The expected payoff from passively investing in period 2 is given by:

$$\begin{aligned}&E(\beta_i M_2/\theta_{i1})\\ &\quad + \{P(\alpha_i/\alpha_i > 0)\,[E(\alpha_i/\alpha_i > 0) + E(\beta_i M_3/\alpha_i > 0)] + P(\alpha_i/\alpha_i < 0)E(\beta_i M_3/\alpha_i > 0)\}\\ &\quad = \tfrac{1}{2}E(\alpha_i/\alpha_i > 0)\end{aligned}$$

Hence she actively trades if $E(\alpha_i/\theta_{i1}) > 0$

If individual i has passively invested in period 1

- The expected payoff from actively trading in period 2 is given by :

$$\{E(\alpha_i/\theta_{i1}) + E(\beta_i M_2/\theta_{i1})\}$$
$$+\{P(\alpha_i/\alpha_i > 0)\,[E(\alpha_i/\alpha_i > 0) + E(\beta_i M_3)] + P(\alpha_i/\alpha_i < 0)E(\beta_i M_3)\}$$
$$= \tfrac{1}{2}E(\alpha_i/\alpha_i > 0)$$

- The expected payoff from passively investing in period 2 is given by :

$$\left\{E(\alpha_i/\theta_{i1}) + E(\beta_i M_2/\theta_{i1})\right\}$$
$$+\left\{P(\alpha_i/\alpha_i>0)[E(\alpha_i)+E(\beta_i M_3)]+P(\alpha_i/\alpha_i<0)E(\beta_i M_3)\right\}$$
$$= 0$$

Hence she always actively trades.

Period 1

In this paragraph I solve for the optimal decision of investor i whether or not to actively trade in period 1. I show that she always actively trades in period 1 in order to seize the option value of learning about her α_i.

The intuition goes as follows. Individual i takes her decision at the beginning of period 1 by comparing the expected payoffs of actively trading versus passively investing over the 3 periods. Suppose that individual i has a negative α_i. If she passively invests in period 1, we know that she will trade in period 2 (from the above paragraph) and that she will refrain from trading in period 3, when uncertainty is resolved. Hence she will have been hurt by actively trading with a negative α_i in 1 period only. On the contrary, if she decides to actively invest in period 1, she will refrain from trading in period 2 as she observes poor portfolio performance and in period 3 when uncertainty is resolved. Hence she will have been hurt by actively trading with a negative α_i in one period only.

Now suppose that individual i has a positive α_i. The same reasoning shows that by passively investing in period 1 she makes money by actively trading in 2 out of 3 periods only. However, if she actively trades in period 1, she makes money by actively trading in 3 out of 3 periods.

Hence, active trading provides limited downside and a clear upside to individual i. Thus she always decides to actively trade rather than passively invest in period 1.

In other words,

- The expected payoff from actively trading in period 1 is given by :

$$E(\alpha_i) + E(\beta_i M_1)$$
$$+P(\alpha_i/E(\alpha_i/R_{i1}, M_1) > 0)\,\{E(\alpha_i/E(\alpha_i/R_{i1}, M_1) > 0) + E(\beta_i M_2/E(\alpha_i/R_{i1}, M_1) > 0)\} + P(\alpha_i/E(\alpha_i/R_{i1}, M_1) \leq 0)E(\beta_i M_2/E(\alpha_i/R_{i1}, M_1) \leq 0)$$
$$+\tfrac{1}{2}E(\alpha_i/\alpha_i > 0)$$
$$= P(\alpha_i/E(\alpha_i/R_{i1}, M_1) > 0)E(\alpha_i/E(\alpha_i/R_{i1}, M_1) > 0) + \tfrac{1}{2}E(\alpha_i/\alpha_i > 0)$$

- The expected payoff from passively investing in period 1 is given by :

$$E(\alpha_i) + E(\beta_i M_1)$$
$$+P(\alpha_i/E(\alpha/R_{i1}, M_1) > 0)\,\{E(\alpha_i/E(\alpha_i/R_{i1}, M_1) > 0) + E(\beta_i M_2/E(\alpha_i/R_{i1}, M_1) > 0)\} + P(\alpha_i/E(\alpha_i/R_{i1}, M_1) \leq 0)E(\beta_i M_2/E(\alpha_i/R_{i1}, M_1) \leq 0)$$
$$+\tfrac{1}{2}E(\alpha_i/\alpha_i > 0)$$
$$= \tfrac{1}{2}E(\alpha_i/\alpha_i > 0)$$

Hence individual i actively trades in period 1 if:

$P(\alpha_i/E(\alpha_i/R_{i1}, M_1) > 0)E(\alpha_i/E(\alpha_i/R_{i1}, M_1) > 0) > 0$

Let us first show that $P(\alpha_i/E(\alpha_i/R_{i1}, M_1) > 0)$ is strictly positive. Since it is positive, it is sufficient to show that it is not equal to zero. Note that

$E(\alpha_i/E(\alpha_i/R_{i1}, M_1)) = 0$. Since $E(\alpha/R_{i1}, M_1)$ takes positive and negative values (it is not always zero), we need to have $P(\alpha_i/E(\alpha_i/R_{i1}, M_1) > 0) > 0$.

Let us now show that $E(\alpha_i/E(\alpha_i/R_{i1}, M_1) > 0) > 0$. Let us rewrite:

$E(\alpha_i/E(\alpha_i/R_{i1},M_1)>0) = E(\alpha_i 1_{E(\alpha_i/R_{i1},M_1)>0}) = E(E(\alpha_i 1_{E(\alpha_i/R_{i1},M_1)>0}/R_{i1},M_1)) = E(1_{E(\alpha_i/R_{i1},M_1)>0}\times E(\alpha_i/R_{i1},M_1))$

This last expression is strictly positive. It can be shown indeed that $E(f\times 1_{f>0})>0$ unless f is everywhere negative.

To see this, suppose$f$ is not everywhere negative. $\{f>0\}=\bigcup_n \{f>\frac{1}{n}\}$.

Now note that $\exists n$ such that $E(\{f\times 1_{\{f>0\}}\})>\frac{1}{n}P(\{f>\frac{1}{n}\})>0$.

Hence $P(\alpha_i/E(\alpha_i/R_{i1},M_1)>0)E(\alpha_i/E(\alpha_i/R_{i1},M_1)>0)>0$ and individual i always actively trades in period 1.

Prediction 1

In period 2, individual i has observed the market return, m_1, and her portfolio return, r_{i1}. Individual i actively trades in period 2 if the expectation of his α_i conditional on his information set is positive, that is if (according to the projection theorem for normal variables):

$E(\alpha_i/Z_{i1}=r_{i1}) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}(r_{i1}-m_1)>0$, where $Z_{i1}=\alpha_i+\beta_i m_1$

Taking this expression conditional on $r_{i1}>0$ gives:

$E(\frac{\sigma_\alpha^2}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}(\alpha_i+(\beta_i-1)m_1)/\alpha_i+\beta m_1>0) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}E(\alpha_i+\beta m_1/\alpha_i+\beta m_1>0) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}\sqrt{\frac{2}{\pi}}\sqrt{\sigma_\alpha^2+m_1^2\sigma_\beta^2}$

Taking this expression conditional on $r_{i1}<0$ gives:

$E(\frac{\sigma_\alpha^2}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}(\alpha_i+(\beta_i-1)m_1)/\alpha_i+\beta m_1<0) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}E(\alpha_i+\beta m_1/\alpha_i+\beta m_1<0) = -\frac{\sigma_\alpha^2}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}\sqrt{\frac{2}{\pi}}\sqrt{\sigma_\alpha^2+m_1^2\sigma_\beta^2}$

Hence the result

Prediction 2

In period 2, individual i has observed the market return, m_1, and her portfolio return, r_{i1}. Individual i actively trades in period 2 if the expectation of his α_i conditional on his information set is positive, that is if:

$E(\alpha_i/Z_{i1}=r_{i1}) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}(r_{i1}-m_1)>0$, where $Z_{i1}=\alpha_i+\beta_i m_1$

Taking this expression conditional on $\alpha_i>0$ gives:

$E(\frac{\sigma_\alpha^2}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}(\alpha_i+(\beta_i-1)m_1)/\alpha_i>0) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}E(\alpha_i/\alpha_i>0) = \frac{\sigma_\alpha^3}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}\sqrt{\frac{2}{\pi}}$

Taking this expression conditional on $\alpha_i>0$ gives:

$E(\frac{\sigma_\alpha^2}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}(\alpha_i+(\beta_i-1)m_1)/\alpha_i\leq 0) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}E(\alpha_i/\alpha_i\leq 0) = -\frac{\sigma_\alpha^3}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}\sqrt{\frac{2}{\pi}}$

Hence the result.

Prediction 3

In period 2, individual i has observed the market return, m_1, and her portfolio return, r_{i1}. Individual i actively trades in period 2 if the expectation of his α_i conditional on his information set is positive, that is if:

$E(\alpha_i/Z_{i1}=r_{i1}) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}(r_{i1}-m_1)>0$, where $Z_{i1}=\alpha_i+\beta_i m_1$

Taking this expression conditional on $\beta_i>1$ gives:

$E(\frac{\sigma_\alpha^2}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}(\alpha_i+(\beta_i-1)m_1)/\beta_i>1) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}E(m_1(\beta_i-1)/\beta_i>1) = \frac{\sigma_\alpha^2 m_1}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}\sigma_\beta\sqrt{\frac{2}{\pi}}$

Taking this expression conditional on $\beta_i\leq 1$ gives:

$E(\frac{\sigma_\alpha^2}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}(\alpha_i+(\beta_i-1)m_1)/\beta_i>1) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}E(m_1(\beta_i-1)/\beta_i>1) = \frac{\sigma_\alpha^2 m_1}{\sigma_\alpha^2+m_1^2\sigma_\beta^2}(-\sigma_\beta\sqrt{\frac{2}{\pi}})$

Hence the result.

Appendix

Figure 4.1: Model time line

This figure displays the time line of the model. At the beginning of each period, the arrow going up means that individual chooses to actively trade in period t and get $\alpha_i + \beta_i \times M_t$ where M_t is the market return. The flat arrow means that individual chooses to passively invest in period t and get $\beta_i \times M_t$. Information set at each node is briefly described inside boxes. The dotted line shows the path individuals never take.

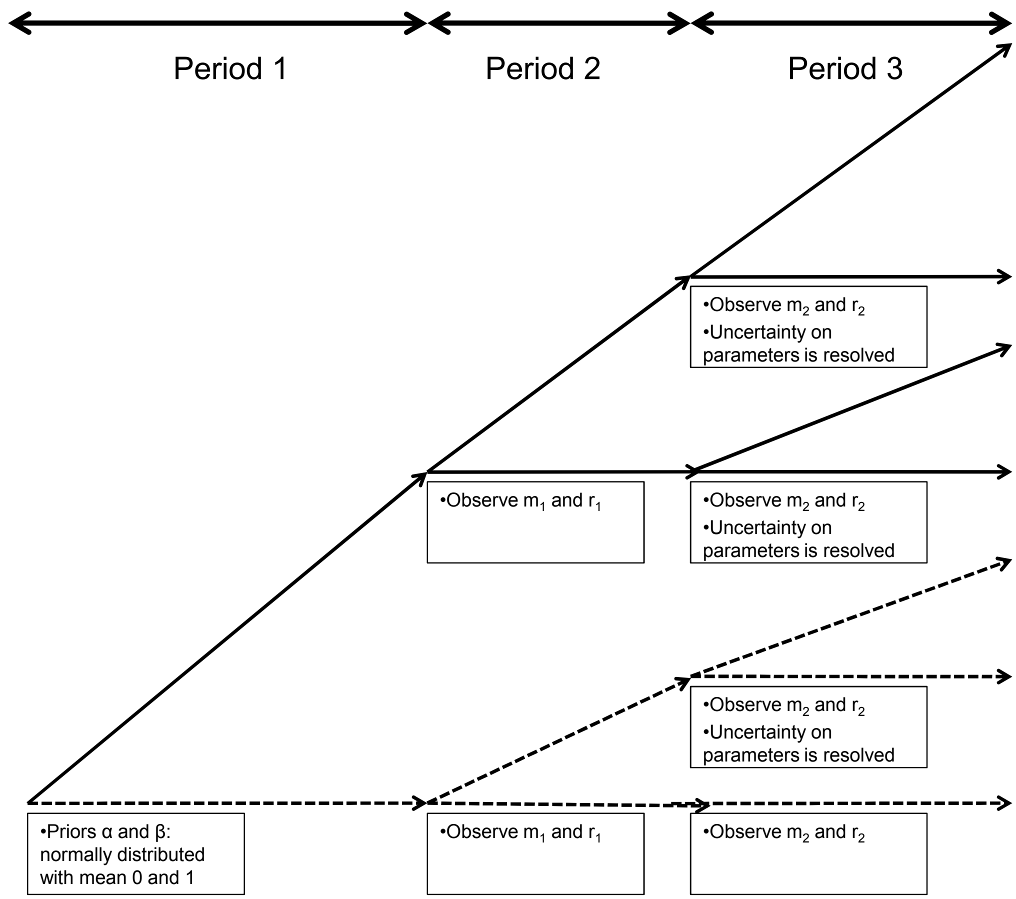


Figure 4.2: Average period 2 conditional expectation of α_i by period 1 performance

This figure displays the value of $E(\alpha_i/\theta_{i1})$, the expectation of α_i at the beginning of period 2, conditional on portfolio returns and market returns observed in period 1. $E(\alpha_i/\theta_{i1})$ is a function of period 1 market return. As derived in section 5, individual i will actively trade in period 2 if $E(\alpha_i/\theta_{i1}) > 0$ and passively invest otherwise. The black line in the figure below is the average value of $E(\alpha_i/\theta_{i1})$ across individuals who obtained positive performance in period 1, that is $r_{i1} > 0$; the grey line is the average value of $E(\alpha_i/\theta_{i1})$ across individuals who obtained negative performance in period 1, that is $r_{i1} \leq 0$. On average, individual investors who obtained positive performance in period 1 will actively trade in period 2 while individual investors who obtained negative performance in period 1 won't. I use some arbitrary values of σ_α and σ_β that I obtain from the dataset used in this paper.

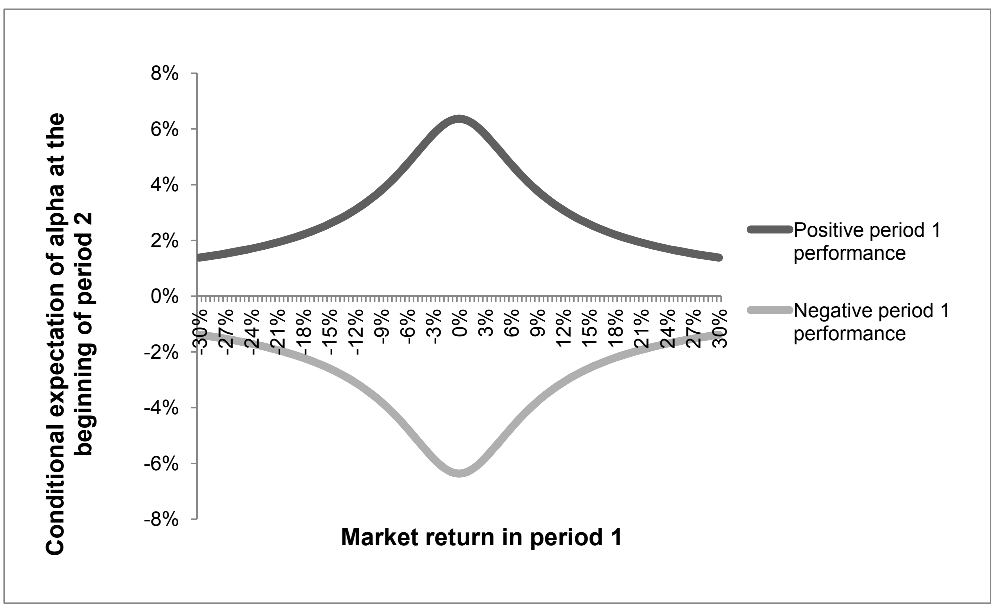


Figure 4.3: Average period 2 conditional expectation of α_i by Alpha

This figure displays the value of $E(\alpha_i/\theta_{i1})$, the expectation of α_i at the beginning of period 2, conditional on portfolio returns and market returns observed in period 1. $E(\alpha_i/\theta_{i1})$ is a function of period 1 market return. As derived in section 5, individual i will actively trade in period 2 if $E(\alpha_i/\theta_{i1}) > 0$ and passively invest otherwise. The black line in the figure below is the average value of $E(\alpha_i/\theta_{i1})$ across individuals with $\alpha_i > 0$; the grey line is the average value of $E(\alpha_i/\theta_{i1})$ across individuals with $\alpha_i \leq 0$. On average, high alpha individuals choose to actively trade in period 2. I use some arbitrary values of σ_α and σ_β that I obtain from the dataset used in this paper.

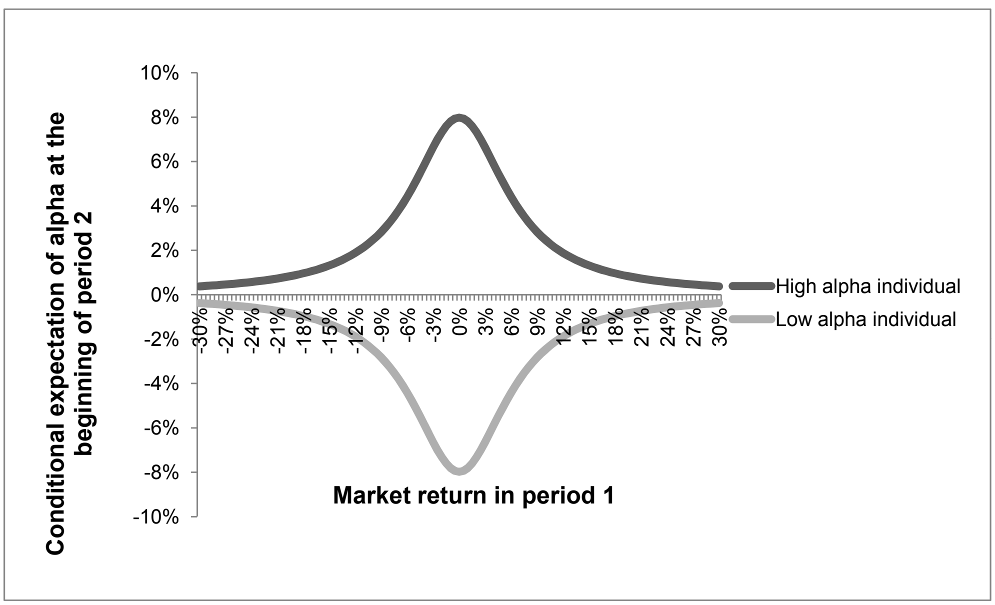


Figure 4.4: Average period 2 conditional expectation of α_i by market exposure (Beta)

This figure displays the value of $E(\alpha_i/\theta_{i1})$, the expectation of α_i at the beginning of period2, conditional on portfolio returns and market returns observed in period 1. $E(\alpha_i/\theta_{i1})$ is a function of period 1 market return. As derived in section 5, individual i will actively trade in period 2 if $E(\alpha_i/\theta_{i1}) > 0$ and passively invest otherwise. The black line in the figure below is the average value of $E(\alpha_i/\theta_{i1})$ across individuals with $\beta_i > 1$; the grey line is the average value of $E(\alpha_i/\theta_{i1})$ across individuals with $\beta_i \leq 1$. On average, high beta individuals actively trade in period 2 when period 1 market returns have been positive. I use some arbitrary values of σ_α and σ_β that I obtain from the dataset used in this paper.

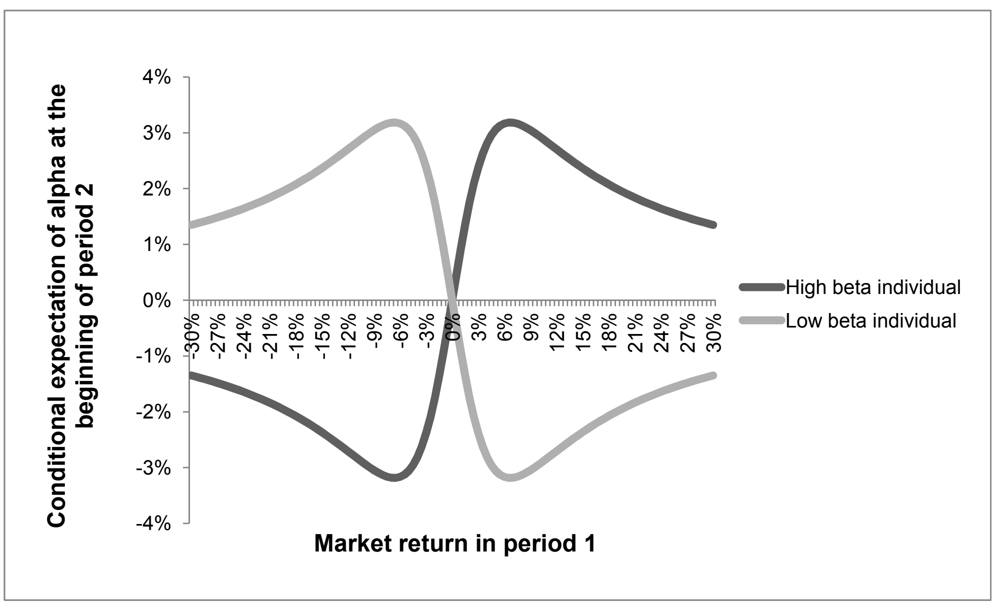


Figure 4.5: Answers to Mifid questionnaire by SRD versus NO SRD Investors

This figure presents the average answer by 833 individuals in the dataset either using the SRD service or not using it to the question "What are your main investment objectives?".

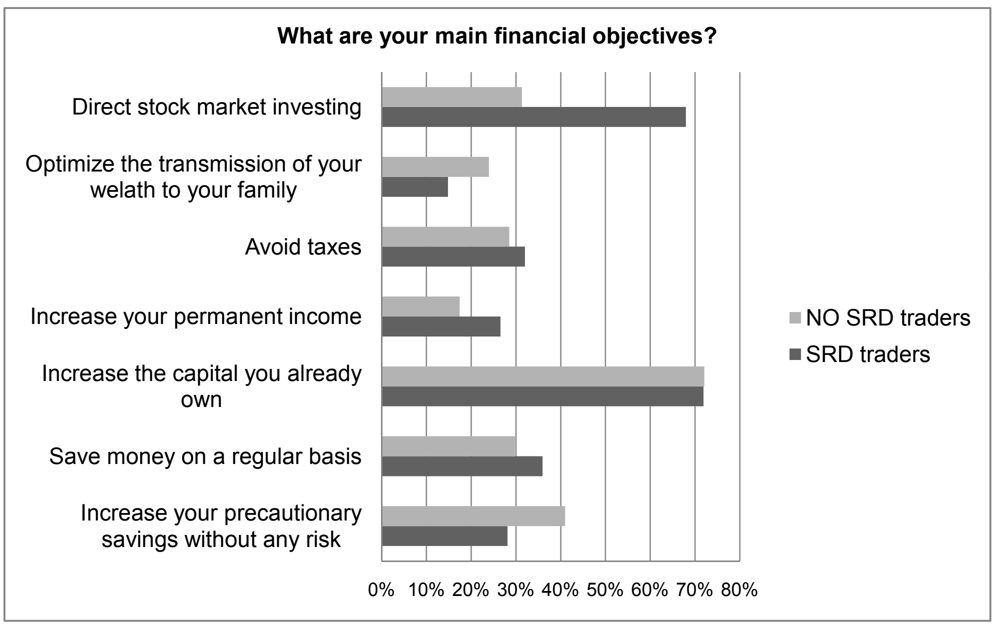


Figure 4.6: Experience and the sensitivity of active trading to past performance

These two figures plot the confidence interval of the coefficients of the following series of regressions which are run in the cross-section of actively trading investors at each experience month e:

$$T_{i,e+1} = \lambda \alpha_{i,e} + \epsilon_{i,e}$$

$$T_{i,e+1} = \lambda \gamma_{i,e} + \epsilon_{i,e}$$

where $T_{i,e+1}$ is a dummy equals to one if individual i is active in month $e+1$ or zero if individual i stopped trading. $\alpha_{i,e}$ is the alpha of investor i in experience month e and $\gamma_{i,e} = \pi_{i,e} - \alpha_{i,e}$ is the aggregate factor exposure performance over all three risk factors of individual i in experience month e.

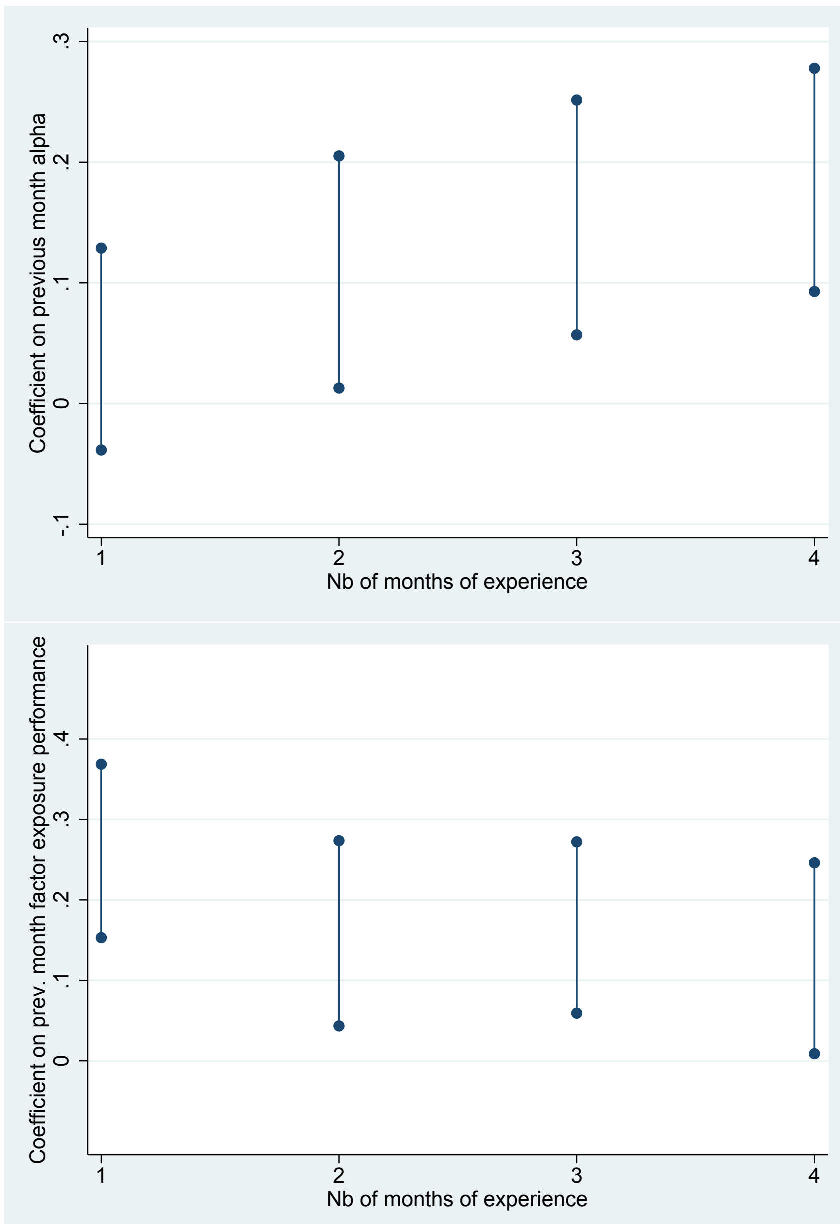


Table 4.1: Summary statistics

This table presents summary statistics for the sample of active traders used in this paper.

	Obs.	Mean	Std. dev.
Individual investors characteristics			
Number of active trading months	14151	7.73	13.82
Number of different stocks traded	14151	12.46	16.38
Trading measures (monthly)			
Log number of trades	109394	1.83	1.25
Log eur volume traded	109394	9.98	1.75
Log average volume per stock traded	109394	7.88	1.49
Eur signed volume over eur volume	109394	0.08	0.56
Log average size of trades	109394	8.15	0.98
Continuation dummy	109394	0.91	0.29
Unweighted performance measures (monthly)			
Performance	109394	-0.001	0.083
Residual performance (Alpha)	109394	-0.001	0.071
Mkt exposure performance, $\beta\pi_{i,t}^{Mkt}$	109394	-0.001	0.065
SMB exposure performance, $\beta\pi_{i,t}^{SMB}$	109394	0.002	0.025
HML exposure performance, $\beta\pi_{i,t}^{HML}$	109394	0.000	0.020
Mkt exposure, $\beta_{i,t}^{Mkt}$	109394	1.590	0.889
SMB exposure, $\beta_{i,t}^{SMB}$	109394	0.468	0.831
HML exposure, $\beta_{i,t}^{HML}$	109394	0.097	0.563
Mkt performance, $\pi_{i,t}^{Mkt}$	109394	0.001	0.029
SMB performance, $\pi_{i,t}^{SMB}$	109394	0.003	0.023
HML performance, $\pi_{i,t}^{HML}$	109394	0.004	0.031
Weighted performance measures (monthly)			
Performance	109394	-0.002	0.095
Residual performance (Alpha)	109394	-0.001	0.083
Mkt exposure performance, $\beta\pi_{i,t}^{Mkt}$	109394	-0.002	0.074
SMB exposure performance, $\beta\pi_{i,t}^{SMB}$	109394	0.001	0.028
HML exposure performance, $\beta\pi_{i,t}^{HML}$	109394	0.000	0.023
Mkt exposure, $\beta_{i,t}^{Mkt}$	109394	1.223	1.224
SMB exposure, $\beta_{i,t}^{SMB}$	109394	0.354	0.892
HML exposure, $\beta_{i,t}^{HML}$	109394	0.068	0.565
Mkt performance, $\pi_{i,t}^{Mkt}$	109394	0.000	0.032
SMB performance, $\pi_{i,t}^{SMB}$	109394	0.002	0.025
HML performance, $\pi_{i,t}^{HML}$	109394	0.003	0.034

Table 4.2: Sensitivity of individual investors active trading to past performance

This table presents the result of the following regression in the panel of 14151 distinct individual investors actively trading in 7.7 successive periods on average.

$$T_{i,t+1} = \lambda_1 \pi_{i,t} + \lambda_2 P_{i,t} + \kappa_i + \delta_e + \mu_t + \epsilon_{i,t}$$

where $T_{i,t+1}$ is a measure of trading activity of individual i in month $t+1$. I consider successively the log of the number of orders (column 1), the log euro volume traded (column 2), the log euro volume per stock traded (column 3), the excess volume purchased over volume sold normalized by total volume (column 4), the log of the average order size (column 5), and a dummy equals to one if individual i is active in month $t+1$ or if individual i stops trading (column 6). $\pi_{i,t}$ is the performance of individual i in month t as computed in section 3.2.2, $P_{i,t}$ is the increase in the log euro value of individual i's portfolio in month t and κ_i, δ_e and μ_t are respectively individual, experience (the number of trading month since individual i started active trading) and month fixed effects. Standard errors are in parenthesis and are clustered at the monthly level. In Panel B, the same regression is run with the weighted performance of individual i in month t where the performance on each stock traded by individual i in month t is weighted by the ratio of the euro volume traded of stock s in month t to the total euro volume traded in month t.

	PANEL A: unweighted measure of performance					
	(1)	(2)	(3)	(4)	(5)	(6)
Prev. month perf.	0.569***	0.898***	0.270***	0.408***	0.328***	0.086***
	(0.083)	(0.111)	(0.048)	(0.062)	(0.037)	(0.018)
Δ Log portfolio value	-0.001	-0.002*	-0.001	0.002**	-0.002**	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Constant	1.492***	9.573***	8.285***	0.084***	8.080***	0.742***
	(0.049)	(0.056)	(0.046)	(0.028)	(0.023)	(0.018)
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Experience FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	109,394	109,394	109,394	109,394	109,394	109,394
R-squared	0.551	0.666	0.692	0.283	0.776	0.469

	PANEL B: weighted measure of performance					
	(1)	(2)	(3)	(4)	(5)	(6)
Prev. month perf.	0.304***	0.469***	0.103***	0.279***	0.165***	0.059***
	(0.070)	(0.089)	(0.036)	(0.055)	(0.031)	(0.014)
Δ Log portfolio value	-0.001	-0.003*	-0.001	0.002**	-0.002**	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Constant	1.485***	9.561***	8.281***	0.080***	8.076***	0.742***
	(0.049)	(0.057)	(0.046)	(0.028)	(0.023)	(0.018)
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Experience FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	109,394	109,394	109,394	109,394	109,394	109,394
R-squared	0.550	0.665	0.692	0.282	0.776	0.469

Table 4.3: Sensitivity of individual investors active trading to their ability and market exposure

This table presents the result of the following regression in the panel of 14151 distinct individual investors actively trading in 7.7 successive periods on average.

$$T_{i,t+1} = \lambda_1\alpha_{i,t} + \lambda_2\beta\pi_{i,t}^{Mkt} + \lambda_3\beta\pi_{i,t}^{SMB} + \lambda_4\beta\pi_{i,t}^{HML}$$
$$+\lambda_5\pi_{i,t}^{Mkt} + \lambda_6\pi_{i,t}^{SMB} + \lambda_7\pi_{i,t}^{HML} + \lambda_8\beta_{i,t}^{Mkt} + \lambda_9\beta_{i,t}^{SMB} + \lambda_{10}\beta_{i,t}^{HML} + \lambda_{11}P_{i,t} + \kappa_i + \delta_e + \mu_t + \epsilon_{i,t}$$

where $T_{i,t+1}$ is a measure of trading activity of individual i in month $t+1$ .I consider successively the log of the number of orders (column 1), the log euro volume traded (column 2), the log euro volume per stock traded (column 3), the excess volume purchased over volume sold normalized by total volume (column 4), the log of the average order size (column 5), and a dummy equals to one if individual i is active in month $t+1$ or if individual i stops trading (column 6). $\beta\pi_{i,t}^{Mkt}$, $\beta\pi_{i,t}^{SMB}$and $\beta\pi_{i,t}^{HML}$ are the returns obtained from the exposure of individual i's trades to the three risk factors. The third component of the regression includes the following controls. $\pi_{i,t}^{Mkt}$, $\pi_{i,t}^{SMB}$and $\pi_{i,t}^{HML}$ are the returns on the indices replicating the risk factors. $\beta_{i,t}^{Mkt}$ $\beta_{i,t}^{SMB}$and $\beta_{i,t}^{HML}$ are the average exposure of the stocks traded by individual i to each of the three risk factors. $P_{i,t}$ is the increase in the log euro value of individual i's portfolio in month t and κ_i, δ_e and μ_t are respectively individual, experience (the number of trading month since individual i started active trading) and month fixed effects. Standard errors are in parenthesis and are clustered at the monthly level. In Panel B, the same regression is then run with variables weighted by the ratio of the euro volume traded of stock s in month t to the total euro volume traded in month t.

	PANEL A: unweighted measure of performance					
	(1)	(2)	(3)	(4)	(5)	(6)
Residual performance, $\alpha_{i,t}$	0.612***	0.972***	0.250***	0.424***	0.359***	0.086***
	(0.093)	(0.124)	(0.055)	(0.065)	(0.045)	(0.023)
Mkt exposure performance, $\beta\pi_{i,t}^{Mkt}$	0.841***	1.336***	0.514***	0.432***	0.495***	0.130***
	(0.155)	(0.188)	(0.108)	(0.122)	(0.069)	(0.033)
SMB exposure performance, $\beta\pi_{i,t}^{SMB}$	0.702***	1.206***	0.573***	0.492***	0.504***	0.170***
	(0.253)	(0.325)	(0.183)	(0.137)	(0.122)	(0.049)
HML exposure performance, $\beta\pi_{i,t}^{HML}$	0.814***	1.057***	-0.163	0.777***	0.243**	0.113**
	(0.247)	(0.300)	(0.176)	(0.186)	(0.095)	(0.045)
Mkt exposure, $\beta_{i,t}^{Mkt}$	0.008	0.006	0.015*	-0.004	-0.002	-0.003
	(0.008)	(0.011)	(0.008)	(0.005)	(0.005)	(0.002)
SMB exposure, $\beta_{i,t}^{SMB}$	-0.001	-0.016	-0.034***	0.005	-0.015***	0.002
	(0.009)	(0.014)	(0.008)	(0.005)	(0.006)	(0.002)
HML exposure, $\beta_{i,t}^{HML}$	0.014	0.018	-0.002	-0.008	0.004	0.003
	(0.010)	(0.014)	(0.010)	(0.006)	(0.005)	(0.002)
Mkt performance, $\pi_{i,t}^{Mkt}$	-0.978**	-1.552***	-0.436	-0.209	-0.574***	-0.095
	(0.471)	(0.592)	(0.296)	(0.352)	(0.181)	(0.066)
SMB performance, $\pi_{i,t}^{SMB}$	-0.663	-0.901	-0.016	-0.399	-0.238	-0.065
	(0.547)	(0.667)	(0.300)	(0.348)	(0.183)	(0.060)
HML performance, $\pi_{i,t}^{HML}$	0.112	0.121	-0.090	-0.231	0.010	-0.025
	(0.281)	(0.369)	(0.191)	(0.300)	(0.118)	(0.035)
ΔLog portfolio value	-0.001	-0.002	-0.001	0.002**	-0.002**	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Constant	1.476***	9.567***	8.287***	0.078***	8.090***	0.743***
	(0.053)	(0.061)	(0.047)	(0.028)	(0.024)	(0.018)
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Experience FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	109,394	109,394	109,394	109,394	109,394	109,394
R-squared	0.551	0.666	0.692	0.283	0.776	0.470

	PANEL B: weighted measure of performance					
	(1)	(2)	(3)	(4)	(5)	(6)
Residual performance , $\alpha_{i,t}$	0.303***	0.453***	0.045	0.306***	0.150***	0.050***
	(0.067)	(0.091)	(0.046)	(0.056)	(0.040)	(0.019)
Mkt exposure performance, $\beta\pi_{i,t}^{Mkt}$	0.216	0.395*	0.092	0.280**	0.179***	0.073**
	(0.159)	(0.204)	(0.082)	(0.108)	(0.061)	(0.036)
SMB exposure performance, $\beta\pi_{i,t}^{SMB}$	0.410	0.736**	0.332**	0.462***	0.326***	0.143***
	(0.255)	(0.327)	(0.135)	(0.124)	(0.113)	(0.047)
HML exposure performance, $\beta\pi_{i,t}^{HML}$	0.823***	0.992***	-0.185	0.612***	0.169**	0.113***
	(0.235)	(0.289)	(0.138)	(0.179)	(0.084)	(0.043)
Mkt exposure, $\beta_{i,t}^{Mkt}$	-0.132***	-0.169***	-0.035***	-0.011***	-0.037***	-0.010***
	(0.013)	(0.015)	(0.005)	(0.004)	(0.004)	(0.001)
SMB exposure, $\beta_{i,t}^{SMB}$	0.067***	0.073***	-0.004	0.009*	0.006	0.005***
	(0.013)	(0.016)	(0.005)	(0.005)	(0.005)	(0.002)
HML exposure, $\beta_{i,t}^{HML}$	-0.000	0.003	-0.006	-0.006	0.004	0.003
	(0.017)	(0.021)	(0.006)	(0.006)	(0.006)	(0.002)
Mkt performance, $\pi_{i,t}^{Mkt}$	-0.672	-0.891	0.054	-0.149	-0.219	-0.043
	(0.495)	(0.601)	(0.230)	(0.295)	(0.141)	(0.060)
SMB performance, $\pi_{i,t}^{SMB}$	-1.022*	-1.250**	-0.022	-0.228	-0.228	-0.071
	(0.528)	(0.630)	(0.226)	(0.287)	(0.145)	(0.051)
HML performance, $\pi_{i,t}^{HML}$	0.209	0.209	0.016	-0.263	0.000	-0.006
	(0.264)	(0.327)	(0.129)	(0.242)	(0.090)	(0.027)
ΔLog portfolio value	-0.001	-0.003*	-0.001	0.002**	-0.002**	0.000
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Constant	1.583***	9.693***	8.320***	0.085***	8.110***	0.749***
	(0.055)	(0.061)	(0.046)	(0.028)	(0.023)	(0.018)
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Experience FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	109,394	109,394	109,394	109,394	109,394	109,394
R-squared	0.556	0.670	0.692	0.283	0.777	0.470

Table 4.4: Auto-correlation of risk factor returns

This table the presents the auto-correlation of the three risk factors monthly returns.

	R_t^{Mkt}	R_t^{SMB}	R_t^{HML}
Auto-correlation coefficient	-0.104	0.001	0.166*
P-value	0.2534	0.988	0.068
Nb of observations	122	122	122

Table 4.5: Co-movement of risk factors and individual investors risk factor exposure

I run the following regression for each factor $f \in \{Mkt, SMB, HML\}$:

$$R_t^f = \lambda_1 \beta_{i,t}^f + \kappa_i + \delta_e + \epsilon_{i,t}$$

where R_t^f is the return of risk factor f in month t. κ_i, δ_e and μ_t are respectively individual, experience (the number of trading month since individual i started active trading) and month fixed effects. Standard errors are in parenthesis and are clustered at the monthly level.

	R_t^{Mkt}	R_t^{SMB}	R_t^{HML}
Mkt exposure, $\beta_{i,t}^{Mkt}$	-0.958***		
	(0.343)		
SMB exposure, $\beta_{i,t}^{SMB}$		0.0262	
		(0.273)	
HML exposure, $\beta_{i,t}^{HML}$			0.251
			(0.473)
Constant	1.704**	-0.760*	-5.684***
	(0.790)	(0.420)	(0.690)
Individual FE	Yes	Yes	Yes
Experience FE	Yes	Yes	Yes
Observations	109,394	109,394	109,394
R-squared	0.141	0.126	0.259

Table 4.6: Persistence in ability and factor exposure over time

I run the following regression for each factor $f \in \{Mkt, SMB, HML\}$:

$$\beta^f_{i,t+1} = \lambda_1 \beta^f_{i,t} + \delta_e + \mu_t + \epsilon_{i,t}$$

where $\beta^f_{i,t+1}$ is the exposure of individual i to factor f in month t. $delta_e$and μ_t are respectively experience (the number of trading month since individual i started active trading) and month fixed effects. Standard errors are in parenthesis and are clustered at the monthly level.

	$\alpha_{i,t+1}$	$\beta^{Mkt}_{i,t+1}$	$\beta^{SMB}_{i,t+1}$	$\beta^{HML}_{i,t+1}$
Residual performance, $\alpha_{i,t}$	0.0139			
	(0.009)			
Mkt exposure, $\beta^{Mkt}_{i,t}$		0.133***		
		(0.019)		
SMB exposure, $\beta^{SMB}_{i,t}$			0.106***	
			(0.022)	
HML exposure, $\beta^{HML}_{i,t}$				0.0680***
				(0.018)
Constant	0.0155***	1.310***	0.835***	0.183***
	(0.000)	(0.017)	(0.009)	(0.004)
Month FE	Yes	Yes	Yes	Yes
Experience FE	Yes	Yes	Yes	Yes
Observations	100,632	100,632	100,632	100,632
R-squared	0.082	0.244	0.181	0.257

Table 4.7: Sensitivity of active trading to performance interacted with absolute factor returns
This table presents the result of the following regression in the panel of 14151 distinct individual investors.

$$T_{i,t+1} = \lambda_1 \pi_{i,t} + \lambda_2 AbsR_t^{Mkt} \times \pi_{i,t} + \lambda_3 AbsR_t^{SMB} \times \pi_{i,t} + \lambda_4 AbsR_t^{HML} \times \pi_{i,t}$$
$$+\lambda_5 AbsR_t^{Mkt} + \lambda_6 AbsR_t^{SMB} + \lambda_7 AbsR_t^{HML} + \lambda_8 P_{i,t} + \kappa_i + \delta_e + \epsilon_{i,t}$$

where $T_{i,t+1}$ is a measure of trading activity of individual i in month $t+1$. I consider successively the log of the number of orders (column 1), the log euro volume traded (column 2), the log euro volume per stock traded (column 3), the excess volume purchased over volume sold normalized by total volume (column 4), the log of the average order size (column 5), and a dummy equals to one if individual i is active in month $t+1$ or if individual i stops trading (column 6). $\pi_{i,t}$ is the performance of individual i in month t as computed in section 3.2.2, $AbsR_t^f$ is the absolute value of the return of risk factor f in month t, $P_{i,t}$ is the increase in the log euro value of individual i's portfolio in month t, and κ_i and δ_e are respectively individual and experience fixed effects . Standard errors are in parenthesis and are clustered at the monthly level. In Panel B, the same regression is run with the weighted performance of individual i in month t where the performance on each stock traded by individual i in month t is weighted by the ratio of the euro volume traded of stock s in month t to the total euro volume traded in month t.

	PANEL A: unweighted measure of performance					
	(1)	(2)	(3)	(4)	(5)	(6)
Previous month performance, $\pi_{i,t}$	1.457***	2.099***	0.246	1.001***	0.642***	0.115**
	(0.240)	(0.302)	(0.152)	(0.192)	(0.125)	(0.056)
$AbsR_t^{Mkt} \times \pi_{i,t}$	-5.047***	-5.204**	1.186	-2.990**	-0.157	-0.215
	(1.836)	(2.251)	(1.066)	(1.361)	(0.889)	(0.287)
$AbsR_t^{SMB} \times \pi_{i,t}$	-9.168***	-10.241***	2.224	-4.537	-1.073	-0.500
	(2.962)	(3.752)	(1.938)	(2.975)	(1.594)	(0.530)
$AbsR_t^{HML} \times \pi_{i,t}$	-0.042	-2.838	-2.456*	0.224	-2.796***	0.476
	(1.801)	(2.123)	(1.402)	(1.588)	(1.046)	(0.357)
$AbsR_t^{Mkt}$	0.412	-0.597	-0.494*	-0.136	-1.009***	-0.008
	(0.471)	(0.598)	(0.256)	(0.263)	(0.283)	(0.097)
$AbsR_t^{SMB}$	1.568**	1.448*	-0.429	0.056	-0.120	0.221**
	(0.643)	(0.776)	(0.348)	(0.314)	(0.295)	(0.105)
$AbsR_t^{HML}$	-0.009	-0.262	-0.342	-0.043	-0.253	0.022
	(0.455)	(0.546)	(0.219)	(0.219)	(0.227)	(0.065)
Δ Log portfolio value	-0.001	-0.002	-0.001	0.002**	-0.001*	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Constant	1.423***	9.669***	8.518***	0.065***	8.246***	0.583***
	(0.048)	(0.054)	(0.038)	(0.023)	(0.019)	(0.018)
Individual and Experience FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	109,394	109,394	109,394	109,394	109,394	109,394
R-squared	0.535	0.653	0.689	0.261	0.767	0.455

	PANEL B: weighted measure of performance					
	(1)	(2)	(3)	(4)	(5)	(6)
Previous month performance, $\pi_{i,t}$	1.101***	1.706***	0.386***	0.783***	0.605***	0.108***
	(0.166)	(0.180)	(0.112)	(0.100)	(0.081)	(0.032)
$AbsR_t^{Mkt} \times \pi_{i,t}$	-1.688	-1.896	0.447	-1.763**	-0.208	-0.065
	(1.060)	(1.283)	(0.686)	(0.733)	(0.609)	(0.173)
$AbsR_t^{SMB} \times \pi_{i,t}$	-6.207***	-7.528***	-0.509	-2.373	-1.322	-0.508
	(2.250)	(2.433)	(1.511)	(1.843)	(1.041)	(0.345)
$AbsR_t^{HML} \times \pi_{i,t}$	-0.547	-2.441	-2.122*	1.029	-1.895**	0.382
	(1.498)	(1.752)	(1.135)	(1.313)	(0.861)	(0.256)
$AbsR_t^{Mkt}$	0.538	-0.473	-0.526**	-0.086	-1.012***	-0.003
	(0.464)	(0.577)	(0.254)	(0.246)	(0.276)	(0.095)
$AbsR_t^{SMB}$	1.550**	1.425*	-0.417	0.039	-0.126	0.221**
	(0.665)	(0.795)	(0.348)	(0.313)	(0.296)	(0.104)
$AbsR_t^{HML}$	-0.035	-0.291	-0.340	-0.049	-0.256	0.022
	(0.460)	(0.546)	(0.223)	(0.217)	(0.230)	(0.065)
Log portfolio value	-0.000	-0.002	-0.001	0.002**	-0.001*	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Constant	1.419***	9.665***	8.520***	0.062***	8.246***	0.583***
	(0.048)	(0.055)	(0.038)	(0.023)	(0.020)	(0.018)
Individual and Experience FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	109,394	109,394	109,394	109,394	109,394	109,394
R-squared	0.535	0.653	0.689	0.261	0.767	0.455

Table 4.8: First period experience and the speed of exit

I run the following regression on the sample of investors for which I observed an exit before December 2010:

$$S_i = \lambda_1 \alpha_{i,1} + \lambda_2 \gamma_{i,1}$$

where S_i is the inverse of the number of successive active trading periods of individual i, $\gamma_{i,1} = \pi_{i,1} - \alpha_{i,1}$ is the aggregate factor exposure performance over all three risk factors of individual i in the first month she trades. Regressions are run for weighted and unweighted measures of performance respectively in columns 1 and 2. Standard errors are in parenthesis.

	Unweighted measures of perf.	Weighted measures of perf.
	Speed of exit, S_i	Speed of exit, S_i
Residual performance , $\alpha_{i,1}$	0.008	0.028
	(0.042)	(0.035)
Factor exposure performance, $\gamma_{i,1}$	-0.220***	-0.116***
	(0.053)	(0.045)
Constant	0.466***	0.467***
	(0.003)	(0.003)
Observations	13,754	13,754
R-squared	0.001	0.001

Conclusion

I briefly conclude this dissertation by emphasizing some directions in which I would like to take my research further.

There is still a lot to learn about trade credit provision, especially at a time when trade credit regulation reforms are likely to take place in Europe. Dimensions which were left out of the scope of Chapter 1 include the impact of the 2006 trade credit restriction on the financial and supply chain policies of borrowers. Conversation with practitioners indicate that small firms using external transportation services suffered a great deal from the cut in their suppliers' credit provision. Hence, more work is needed to apprehend the welfare implications of trade credit. More generally, firms operate within complicated supply chains and are likely to be affected by the behavior of other firms along these supply chains. Understanding how corporate policy is influenced by firms interactions with their clients, suppliers and competitors is a very promising research path to which I intend to contribute in the future.

Policies attempting to promote innovation and entrepreneurship have been on the top of policy makers agenda in the past few years, with very few instances of success. In the last decade, many traditional sources of entrepreneurial finance such as venture capital have seen poor returns. At the same time, new forms of financing have emerged such as corporate venture capital, crowd-funding, angel groups and various incubator-like facilities. Innovation is a complex and multi-dimensional phenomenon, which makes it a fascinating topic of research.

In Chapter 2 of this dissertation, I considered innovation from a finance standpoint. I believe there is much more to learn about the relationships between finance and innovation. In his Presidential Address at the annual meetings of the American Finance Association in January 2012, Raghuram Rajan delivered a thoughtful framework to think about this question. He highlighted the fundamental friction between the need for a company to differentiate and innovate to find its way through competition, and the need to be standardized enough to raise outside finance to fund development. I look forward to seeing and doing more research in this direction.

Evidence is currently accumulating that households make investment mistakes. Part of these are due to financial illiteracy, some others to behavioral biases, others to the predatory behaviors of financial intermediaries. All have diverging policy implications. In Chapter 4 of this dissertation, I hypothesize that households have a limited knowledge of their exposure to systematic risk factors, and that they consequently make investment mistakes. In a companion, yet very preliminary project with Yang Sun (MIT Sloan), we investigate whether the same pattern is observed in retail mutual fund flows. If it is the case that retail investors do not adjust for the exposure to systematic risk factors in evaluating their mutual fund or financial intermediaries performance, this probably leads them to make costly decision. This also potentially carries asset pricing implications. According to this hypothesis, high beta mutual funds should experience larger outflows when the market has gone down. High beta stocks could thus incur excess price pressure in bear markets as high beta mutual funds have to liquidate their holdings to face retail investors redemptions. More generally, I plan to carry on more research to understand how systematic risk is factored in households financial decisions, and how households invest in the acquisition of financial knowledge.

Essays in Empirical Financial Economics

This dissertation is made of four distinct chapters. In the first chapter, I consider an exogenous restriction on the ability of French trucking firms to extend payment terms to their clients. I find that they provide trade credit at the cost of lower investment, lower return on assets, and higher default risk. In the second chapter, I show that private equity funds with a longer horizon select younger companies at an earlier stage of their development. Companies which receive funding from funds with a longer horizon increase their patent stock significantly more than companies which receive funding from investors with a shorter horizon. The third chapter presents a joint work with Ron Kaniel and David Sraer. We use detailed brokerage account data to provide a quantitative exploration of the behavior of retail investors during the financial crisis of 2008. We show that investors who appear more sophisticated on these dimensions in the pre-crisis period were, in the post-crisis period, less likely to flee to safety, more likely to engage in liquidity provisions and to earn higher returns. In the fourth chapter, I develop the idea that households have an imprecise knowledge of their portfolio's exposure to systematic risk and that this leads them to make investment mistakes. This idea is tested in the context of the decision to actively trade rather than passively invest in the stock market.

Keywords: horizon, innovation, patents, venture capital, private equity, learning, individual investor behavior, household finance, trade credit, crisis.

Essais en Economie Financière Empirique

Cette thèse est constituée de quatre chapitres distincts. Dans le premier chapitre, j'utilise une restriction exogène de la capacité des entreprises de transport routier à consentir des délais de paiement à leurs clients. Je montre que certaines entreprises prêtent à leurs clients au détriment de leurs investissements, de leur rentabilité et en s'exposant au risque de défaillance. Dans le second chapitre, je montre que les fonds d'investissement dont l'horizon est long choisissent des entreprises plus jeunes, à un stade moins avancé de leur développement. Les entreprises investies par des fonds dont l'horizon est plus long accroissent leur stock de brevets plus rapidement que celles qui sont investies par des fonds dont l'horizon est plus court. Le troisième chapitre est le résultat d'une collaboration avec Ron Kaniel et David Sraer. Nous utilisons des données détaillées de courtier et entreprenons une exploration quantitative du comportement des investisseurs individuels pendant la crise financière de 2008. Nous montrons que les investisseurs qui ont l'air les plus sophistiqués dans la période antérieure à la crise ont une propension moins grande à fuir vers les actifs sans risque, et une propension plus grande à être apporteurs de liquidité et à obtenir des rendements élevés pendant la crise. Dans le quatrième chapitre, j'explore l'idée selon laquelle les ménages ont une connaissance limitée de l'exposition de leur portefeuille aux facteurs de risque systématique, ce qui les conduit à faire des erreurs. Cette idée est appliquée à la décision des investisseurs individuels d'intervenir activement plutôt que d'investir passivement sur les marchés d'actions.

Mots clefs: horizon, innovation, brevets, capital investissement, apprentissage, investisseurs individuels, finance des ménages, délais de paiement, crise.