N O T A S E C O N Ó M I C A S 46

Paul Slovic
Affect, Reason, Risk and Rationality

João Eira
An Introduction to Other-Regarding Preferences with an Application to Contract Design

António Afonso / André Albuquerque
Sovereign Credit Rating Mismatches

Fernando Alexandre / Pedro Bação, Carlos Carreira / João Cerejeira / Gilberto Loureiro, António Martins / Miguel Portela
Investimento Empresarial em Portugal: Crise e Recuperação

Sara Cerdeira
The Business Plan in the Context of Corporate Entrepreneurship: A Literature Review

Imprensa da Universidade de Coimbra
Coimbra University Press

Faculdade de Economia da
Universidade de Coimbra

(Página deixada propositadamente em branco)

NOTAS

ECONÓMICAS 46

Propriedade
Faculdade de Economia
Universidade de Coimbra
Av. Dias da Silva, 165
3004-512 Coimbra

Apoios
Esta revista é financiada por Fundos Nacionais através da FCT – Fundação para a Ciência e a Tecnologia no âmbito do projeto PEst-OE/EGE/UI0031/2014

Número avulso
Portugal 7,5 €
Outros países 15 €

Condições de assinatura bienal
Portugal: 20 €
Outros países: 40 €

ISSN 0872-4733

ISSN Digital 2183-203X

DOI https://doi.org/10.14195/2183-203X_46

Depósito Legal 64932/93

(Página deixada propositadamente em branco)

Sumário

Artigos

Paul Slovic
Affect, Reason, Risk and Rationality • 7

João Eira
An Introduction to Other-Regarding Preferences
with an Application to Contract Design • 17

António Afonso / André Albuquerque
Sovereign Credit Rating Mismatches • 49

Fernando Alexandre / Pedro Bação / Carlos Carreira / João Cerejeira,
Gilberto Loureiro / António Martins / Miguel Portela
Investimento Empresarial em Portugal: Crise e Recuperação • 71

Sara Cerdeira
The Business Plan in the Context of Corporate
Entrepreneurship: A Literature Review • 99

(Página deixada propositadamente em branco)

Paul Slovic
University of Oregon, USA
pslovic@uoregon.edu

DOI: https://doi.org/10.14195/2183-203X_46_1

Affect, Reason, Risk and Rationality[1]

Paul Slovic

Received for publication: December 12, 2017
Accepted for publication: December 12, 2017

ABSTRACT
Our understanding of the psychology of human judgment, preference and choice is continually evolving. Behavioral researchers are coming to recognize that there is an aspect of information-processing that has been rather neglected. This is the, experiential, affect-based side of our mental life, which appears every bit as important as the analytic/deliberative side that has been the focus of much prior research and the foundation for multi criteria decision analysis. This essay will briefly describe new research demonstrating the powerful influence of affect on decision-making. Reliance on affect is essential to rational behavior yet it sometimes misleads us. In such circumstances we need to ensure that reason also is employed.
Keywords: Psychology of human judgment; affect; behavioral research; multi criteria decision analysis.

JEL Classification: C91; D81; D91.

Acknowledgements: This paper is based upon the author's joint work with Ellen Peters, Melissa Finucane, and Donald MacGregor. Preparation of the manuscript was supported by grant # SES-0241313 from the National Science Foundation.

[1] This is a reprint of the essay published in the *Newsletter of the European Working Group* "*Multicriteria Aid for Decisions*," Series 3, 13, Spring 2006.

Background and Theory: The Importance ff Affect

Although the visceral emotion of fear certainly plays a role in risk as feelings, we shall focus here on a "faint whisper of emotion" called affect. As used here, "affect" refers to specific feelings of "goodness" or "badness" experienced with or without conscious awareness. Affect plays a central role in what have come to be known as dud-process theories of thinking. As Epstein (1994) observed,

> *There is no dearth of evidence… that people apprehend reality in two fundamentally different ways, one variously labeled intuitive, automatic, natural, non-verbal, narrative, and experiential, and the other analytical, deliberative, verbal, and rational.* (p. 710)

Table 1 compares these two systems. One of the main characteristics of the experiential system is its affective basis. Although analysis is certainly important in some decision-making circumstances, reliance on affect is a quicker, easier, and more efficient way to navigate in a complex, uncertain, and sometimes dangerous world. Many theorists have given affect a direct and primary role in motivating behavior. Pleasant feelings motivate actions and thoughts anticipated to reproduce the feelings. Unpleasant feelings motivate actions and thought anticipated to avoid the feelings.

Table 1. Two modes of thinking: Comparison of experiential and analytic systems

System 1 Experiential System	System 2 Analytic System
Affective: pleasure-pain oriented	Logical: reason oriented (what is sensible)
Connections by association	Connections by logical assessment
Behavior mediated by feelings from past experiences	Behavior mediated by conscious appraisal of events
Encodes reality in concrete images, metaphors, and narratives	Encodes reality in abstract symbols, words, and numbers
More rapid processing: oriented toward immediate action	Slower processing: oriented toward delayed action
Self-evidently valid: "experiencing is believing"	Requires justification via logic and evidence

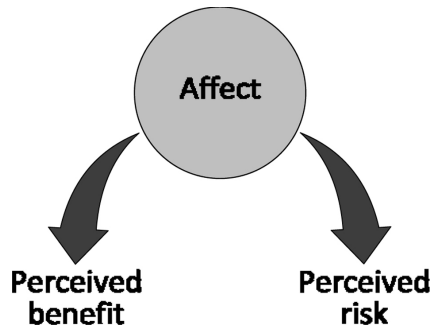
Source: Adapted from Epstein (1994).

There are strong elements of rationality in both systems. The experiential system enabled human beings to survive during their long period of evolution. Long before there was probability theory, risk assessment, and decision analysis, there were intuition, instinct, and gut feeling to tell us whether an animal was safe to approach or the water was safe to drink.

As life became more complex and humans gained more control over their environment, analytic tools were invented to "boost" the rationality of experiential thinking.

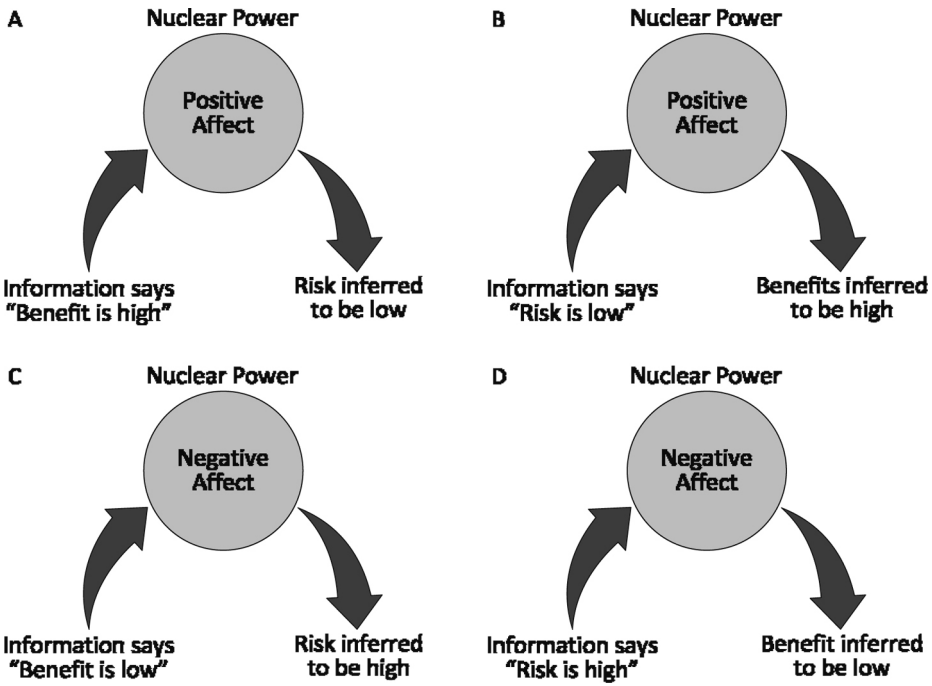
Studies of risk perception have demonstrated that, whereas risk and benefit tend to be positively correlated in the world, they are negatively correlated in people's minds and judgments (Fischhoff et al., 1978). The significance of this finding was not realized until a study by Alhakami and Slovic (1994) found that the inverse relationship between perceived risk and perceived benefit of an activity (e.g., using pesticides) was linked to the strength of positive or negative affect associated with that activity as measured by rating the activity on bipolar scales such as good/bad, nice/awful, etc. This implies that people judge a risk not only by what they think about it but also by how they feel about it. If their feelings towards an activity are favorable, they are moved toward judging the risks as low and the benefits as high; if their feelings toward it are unfavorable, they tend to judge the opposite – high risk and low benefit. Finucane et al. (2000) called this process "the affect heuristic" (see Figure 1).

Figure 1. A model of the affect heuristic explaining the risk/benefit confounding observed by Alhakami and Slovic (1994). Judgments of risk and benefit are assumed to derive by reference to an overall affective evaluation of the stimulus item. Source: Finucane et al. (2000)



If affect guides perceptions of risk and benefit, then providing information about benefit should change perception of risk and vice-versa (see Figure 2). For example, information stating that benefit is high for a technology such as nuclear power should lead to more positive overall affect which should, in turn, decrease perceived risk (Figure 2A).

Figure 2. Model showing how information about benefit (A) or information about risk (B) could increase the positive affective evaluation of nuclear power and lead to inferences about risk and benefit that coincide affectively with the information given. Similarly, information could make the overall affective evaluation of nuclear power more negative as in C and D, resulting in inferences about risk and benefit that are consistent with this more negative feeling. Support for this model was found by Finucane et al. (2000)



Finucane et al. (2000) tested the predictions outlined in Figure 2, providing four different kinds of information designed to manipulate affect by increasing or decreasing perceived benefit or by increasing or decreasing perceived risk. This was done for each of three technologies. The predictions were confirmed. Further support for the affect heuristic came from a second experiment by Finucane et al. who found that the inverse relationship between perceived risks and benefits increased greatly under time pressure, when opportunity for analytic deliberation was reduced. These two experiments demonstrate that affect influences judgment directly and is not simply a response to a prior analytic evaluation.

Failures of the Experiential System

The affect heuristic has been portrayed as the centerpiece of the experiential mode of thinking, the dominant mode of risk assessment and survival during the evolution of the

human species. However, like other heuristics that provide efficient and generally adaptive responses but occasionally get us into trouble, reliance on affect can also mislead us, as will be shown below. Indeed, if it were always optimal to follow our affective and experiential instincts, there would have been no need for the rational/analytic system of thinking to have evolved and become so prominent in human affairs.

Judgments of Probability, Relative Frequency, and Risk

The experiential system of thinking encodes reality in images, metaphors, and narratives to which affective feelings have become attached. To demonstrate this system, Denes-Raj and Epstein (1994) showed that, when offered a chance to win $1.00 by drawing a red jelly bean from an urn, individuals often elected to draw from a bowl containing a greater absolute number, but a smaller proportion, of red beans (e.g., 7 in 100) than from a bowl with fewer red beans but a better probability of winning (e.g., 1 in 10). These individuals reported that, although they knew the probabilities were against them, they felt they had a better chance when there were more red beans.

We can characterize Epstein's subjects as following a mental strategy of "imaging the numerator" (i.e., the number of red beans) and neglecting the denominator (the number of beans in the bowl). Consistent with the affect heuristic, images of winning beans convey positive affect that motivates choice.

Although the jelly bean experiment may seem frivolous, imaging the numerator brings affect to bear on judgments in ways that can be both non-intuitive and consequential. Slovic, Monahan, and MacGregor (2000) demonstrated this by asking experienced forensic psychologists and psychiatrists to judge the likelihood that a hospitalized mental patient would commit an act of violence within 6 months after being discharged from the facility. An important finding was that clinicians who were given another expert's assessment of a patient's risk of violence framed in terms of relative frequency (e.g., "of every 100 patients similar to Mr. Jones, 10 are estimated to commit an act of violence to others") subsequently labeled Mr. Jones as more dangerous than did clinicians who were shown a statistically "equivalent" risk expressed as a probability (e.g., "Patients similar to Mr. Jones are estimated to have a 10% chance of committing an act of violence to others").

Not surprisingly, when clinicians were told that "20 out of every 100 patients similar to Mr. Jones are estimated to commit an act of violence," 41% refused to discharge the patient. But when another group of clinicians was given the risk as "patients similar to Mr. Jones are estimated to have a 20% chance of committing an act of violence," only 21% refused to discharge the patient. Follow-up studies showed that representations of risk in the form of individual probabilities of 10% or 20% led to relatively benign images of one person, unlikely to harm anyone, whereas the "equivalent" frequentistic representations created frightening images of violent patients (example: "Some guy going crazy and killing someone"). These affect-laden images likely induced greater perceptions of risk in response to the relative frequency frames.

Insensitivity to Probability (Probability Neglect)

When the consequences of an action or event carry strong affective meaning, as is the case with a lottery jackpot or a cancer, the probability of such consequences often carries too little weight. As Loewenstein et al. (2001) observe, one's images and feelings toward winning the lottery are likely to be similar whether the probability of winning is one in ten million or one in ten thousand. They further note that responses to uncertain situations appear to have an all-or-none characteristic that is sensitive to the *possibility* rather than the *probability* of strong positive or negative consequences, causing very small probabilities to carry great weight. Empirical support for these arguments comes from Rottenstreich and Hsee (2001) who show that, if the potential outcome evokes strong positive or negative affect, its attractiveness or unattractiveness is relatively insensitive to changes in probability as great as from .99 to .01.

Legal scholar Cass Sunstein (2003; p. 122) labels this insensitivity *probability neglect* and argues that this phenomenon causes extreme overreaction to terrorist threats by both public officials and private citizens.

> [P]eople are prone to… *probability neglect*, especially when their emotions are intensely engaged. Probability neglect is highly likely in the aftermath of terrorism….When probability neglect is at work, people's attention is focused on the bad outcome itself, and they are inattentive to the fact that it is unlikely to occur.

Managing Affect, Reason, and Risk

Affect misguides us in many important ways resulting from the natural limitations of the experiential system and the existence of stimuli in the environment that are simply not amenable to valid affective representation. We have seen above the way that perceptions of risk can be confused by positive feelings (e.g., benefits). Risk perceptions and decision making can also be inappropriate when the *presence of strong affect* leads us to be insensitive to probabilities. Moreover, the affective system seems designed to sensitize us to small changes in our environment (e.g., the difference between 0 and 1 deaths) at the cost of making us less able to appreciate and respond appropriately to larger changes further away from zero (e.g., the difference between 87 deaths and 88 deaths). Fetherstonhaugh et al. (1997) referred to this insensitivity as "psychophysical numbing." Nobel-prize winning biochemist Albert Szent-Gyorgi put it another way as he struggled to comprehend the enormity of the consequences of nuclear war: "I am deeply moved if I see one man suffering and would risk my life for him. Then I talk impersonally about the possible pulverization of our big cities, with a hundred million dead. I am unable to multiply one man's suffering by a hundred million."

Now that we are beginning to understand the complex interplay between emotion, affect, and reason that is wired into the human brain and essential to rational behavior, the challenge before us is to think creatively about what this means for managing risk and making good decisions. On the one hand, how do we apply reason to temper the strong emotions engendered by some risk events? On the other hand, how do we infuse needed

"doses of feeling" into circumstances where lack of experience may otherwise leave us too "coldly rational?"

Can Generation of Reasons Degrade Decision Quality?

Daniel Kahneman (2003) in his Nobel Prize Address argues that highly accessible impressions produced by the experiential system (he calls it System 1) control judgments and decisions, unless modified or overridden by the deliberate operations of the analytic system (called System 2). This suggests that deliberative, reason-based analysis generally will improve decision quality. This view also implies that errors of intuitive judgment involve failures of both systems—System 1, which generates the error, and System 2, which fails to detect and correct it. The corrective operations of System 2 may be impaired by time pressure (Finucane et al., 2000), by cognitive load (Shiv and Federikhan, 1999; Gilbert, 2002), by stress, by age, or by individual cognitive imitations (Peters et al., 2005).

But what happens when System 2 is brought into play early, as when an individual is asked to generate reasons to support a judgment or decision? Research by Wilson and colleagues demonstrates that, when affect is important, an attempt by the decision maker to provide reasons might sometimes produce an inferior decision by interfering with the affective feelings (Epstein, 1994; see Table 1) that subsequently determine how we will experience the consequences of the decision (Wilson and Schooler, 1991; Wilson *et al.*, 1993). For example, Wilson *et al.* found that people who gave numerous reasons for liking an art poster prior to choosing it were subsequently less satisfied with it than those who chose without explicitly considering reasons. Similar degrading of decision performance due to introspection is reported by Tordesillas and Chaiken (1999). Could this pose problems for decision analysis, which depends heavily on introspective judgments?

Can Analysis Benefit from Experiential Thinking?

The answer to this question is almost certainly yes. Even such prototypical analytic exercises as proving a mathematical theorem or selecting a move in chess benefit from experiential guidance. The mathematician senses whether the proof "looks good" and the chessmaster gauges whether a contemplated move "feels right," based upon stored knowledge of a large number of winning patterns (de Groot, 1978). Analysts attempting to build a model to solve a client's decision-making problem are instructed to rely upon the client's sense of unease about the results of the current model as a signal that further modeling may be needed (Phillips, 1984). A striking example of failure because an analysis was devoid of feeling was perpetrated by Philip Morris. The company commissioned an analysis of the costs to the Czech government of treating diseased smokers. Employing a very narrow conception of costs, the analysis concluded that smokers benefited the government by dying young. The analysis created so much hostility that Philip Morris was forced to issue an apology ("Philip Morris," 2001). Another example of the need to respect "experiential wisdom" comes from

the inquiry into the causes of the Columbia Space Shuttle disaster, which pointed to the failure of NASA's risk assessment protocols to give weight to the worries and hunches of personnel who had observed suspicious damage to heat-shielding tiles on previous flights. An article in *Aviation Week* asserted that lack of hard data prevented the input of common sense analysis into the risk-assessment process (Covault, 2003).

Elsewhere I have argued that risk analysis needs to be sensitive to the "softer" values underlying such qualities as dread, equity, controllability, etc. that underlie people's concerns, as well as to degrees of ignorance or scientific uncertainty (Slovic, 1987; 2000). A blueprint for doing this is sketched in the National Academy of Sciences report *Understanding Risk: Decision Making in a Democratic Society* (National Research Council, 1996).

Conclusion

Reliance on affect is a sophisticated cognitive mechanism that helps us to respond quickly and effectively in many decision situations. In other circumstances, affect may lead us to judge probabilities and consequences and make decisions in ways that are not beneficial. We need to understand the circumstances in which affect improves our decision making and the circumstances in which it leads us astray. Additional research on affect and decision making will be essential to this understanding.

References

Alhakami, A. S.; Slovic, P. (1994) A psychological study of the inverse relationship between perceived risk and perceived benefit, *Risk Analysis*, 14(6), 1085-1096.

Covault, C. (2003) Echoes of challenger, *Aviation Week and Space Technology*, 158(17), April 28, 29-31.

de Groot, A. D. (1978) *Thought and Choice in Chess*, New York, Monton.

Epstein, S. (1994) Integration of the cognitive and the psychodynamic unconscious, *American Psychologist*, 49, 709-724.

Fetherstonhaugh, D.; Slovic, P.; Johnson, S. M.; Friedrich, J. (1997) Insensitivity to the value of human life: A study of psychophysical numbing, *Journal of Risk and Uncertainty*, 14, 283-300.

Finucane, M. L.; Alhakami, A.; Slovic, P.; Johnson, S. M. (2000) The affect heuristic in judgments of risks and benefits, *Journal of Behavioral Decision Making*, 13(1), 1-17.

Fischhoff, B.; Slovic, P.; Lichtenstein, S., Read, S.; Combs, B. (1978) How safe is safe enough? A psychometric study of attitudes toward technological risks and benefits, *Policy Sciences*, 9, 127-152.

Kahneman, D. (2003) A perspective on judgment and choice: Mapping bounded rationality, *American Psychologist*, 58(9), 697-720.

Kahneman, D.; Slovic, P.; Tversky, A. (Eds.) (1982) *Judgment under Uncertainty: Heuristics and Biases*, New York, Cambridge University Press.

Loewenstein, G. F.; Weber, E. U.; Hsee, C. K.; Welch, E. S. (2001) Risk as feelings, *Psychological Bulletin*, 127(2), 267-286.

National Research Council. Committee on Risk Characterization (1996) Understanding Risk: Informing Decisions in a Democratic Society, Washington, DC, National Academy Press.

Peters, E.; Västfjäll, D.; Slovic, P.; Mertz, C.K.; Mazzocco, K.; Dickert, S. (2005) Numeracy and Decision Making (Decision Research Report No. 05-01).

Philip Morris issues apology for Czech study on smoking. (2001, July 27) *The New York Times*, p. C12.

Phillips, L. D. (1984) A theory of requisite decision models, *Acta Psychologica*, 56, 29-48.

Rottenstreich, Y.; Hsee, C. K. (2001) Money, kisses and electric shocks: On the affective psychology of probability weighting, *Psychological Science*, 12(3), 158-190.

Shiv, B.; Fedorikhin, A. (1999) Heart and mind in conflict: The interplay of affect and cognition in consumer decision making, *Journal of Consumer Research*, 26, 278-292.

Slovic, P. (1987) Perception of risk, *Science*, 236, 280-285.

Slovic, P. (2000) *The Perception of Risk*, London, Earthscan.

Slovic, P.; Monahan, J.; MacGregor, D. G. (2000) Violence risk assessment and risk communication: The effects of using actual cases, providing instruction, and employing probability versus frequency formats, *Law and Human Behavior*, 24, 271-296.

Sunstein, C. R. (2003) Terrorism and probability neglect, *The Journal of Risk and Uncertainty*, 26(2/3), 121-136.

Tordesillas, R. S.; Chaiken, S. (1999) Thinking too much or too little? The effects of introspection on the decision-making process, *Personality and Social Psychology*, 25(5), 623-629.

Wilson, T. D.; Lisle, D. J.; Schooler, J. W.; Hodges, S. D.; Klaaren, K. J.; LaFleur, S. J. (1993) Introspecting about reasons can reduce postchoice satisfaction, *Personality and Social Psychology Bulletin*, 19, 331-339.

Wilson, T. D.; Schooler, J. W. (1991) Thinking too much: Introspection can reduce the quality of preferences and decisions, *Journal of Personality and Social Psychology*, 60(2), 181-192.

(Página deixada propositadamente em branco)

João Eira
simbyotic@gmail.com

DOI: https://doi.org/10.14195/2183-203X_46_2

An Introduction to Other-Regarding Preferences with an Application to Contract Design

João Eira

Received for publication: February 22, 2018
Revision accepted for publication: March 19, 2018

ABSTRACT
Economic models of individual behavior often make the assumption that in evaluating between competing alternatives agents are only concerned with how each alternative impacts their own payoffs. This simple, yet reasonable, assumption postulates that agents are self-regarding, that is, agents are not concerned with how their decisions affect other people. This study casts doubt on this assumption.
There has been a steady accumulation of experimental evidence from games such as the ultimatum game and the gift exchange game where the observed behavior is not explained by assuming that agents have self-regarding preferences. Agents often make decisions that lower their payoffs if by doing so other agents are better off. In contrast to self-regarding preferences, agents are said in this case to have other-regarding preferences.
Most of the evidence discussed in this study was gathered by the use of laboratory experiments. The issue of external validity of this evidence has long been a point of contention. Lab experiments are highly artificial environments that place strong constraints on individual behavior. While this imbues them with their source of methodological strength, it is also a weakness. Evidence gathered in the lab does not necessarily generalize to the real world, and lab experiments are often compared with field studies which are assumed to provide evidence that is more externally valid. We examine the question of the external validity of lab experiments and conclude they are a valid tool for gathering scientific evidence about human behavior.
Inequity aversion is presented as a method of modeling other-regarding preferences. The model is promptly used to explain the behavior documented in the ultimatum game. An example on how to use other-regarding preferences to study real world economic interactions is provided in the study of contract design under moral hazard.
Keywords: Self-regarding preferences; social preferences; ultimatum game, contract design.

JEL Classification: D01; C70; B41.

1. Introduction

Economics is a social science and the economic behavior that is its subject of study is human behavior. As social scientists, economists are interested in studying agents, their actions, the reasons behind them, and the consequences that result from them. It is through the use of tractable models that economists perform their studies and, as a necessary step to develop these models, it is required that the goals and motivations that precede and drive human behavior be formalized.

We are regularly faced with situations where we have to choose between multiple possible courses of action. Before entering college we must decide between majoring in Physics or Economics. When lunch hour arrives, and we find ourselves at a mall, we have to choose one restaurant out from possible dozens. Economists deal with this basic fact of everyday life by introducing the concept of preferences, which are defined as rankings that express the subjective comparative evaluations of alternatives (Hausman, 2011).

An agent's behavior can be summarized as the maximization of an abstract utility function. While this utility function does not necessarily take into account the underlying psychological processes that underlie preference formation, it has become standard in Economics to take this function as being the result of an evaluation that takes into account as its sole parameter how each alternative impacts the agent's payoff. Under this behavioral assumption, the maximizing behavior corresponds to the idea that in the presence of competing alternatives people seek to maximize their own expected payoffs. Agents are then said to be self-regarding.

This assumption was put forth in 1881, when the political economist and philosopher Francis Edgeworth asserted that "the first principle of Economics is that every agent is actuated only by self-interest" (Edgeworth, 1881). More famously, we see it in Adam Smith's concept of the invisible hand, where the market is able to turn what are private vices into public virtues.[1]

This work intends to show that there is sufficient experimental evidence showing that the assumption that agents are self-regarding is insufficient. It will be shown that there is a wide range of behavior which the self-regarding assumption is not able to explain and that instead one needs to take into account that agents are concerned not only with themselves but also take into account the well-being of others.

We will therefore contrast other-regarding preferences with self-regarding preferences. An agent is said to be self-regarding if he is only preoccupied with how an action impacts himself, while an other-regarding individual is not only preoccupied with himself but also with other people.

Note that an individual with other-regarding preferences does not imply that he is not preoccupied with himself. For example, one can be honest because one does not wish to impose costs on others by deceiving them, but honesty can also be a self-regarding behavior if practiced in order to be the kind of person one wants to be. Thus, the distinction between the two preferences is not that other-regarding preferences are counter-preferential, in the

[1] Contrary to what one might infer from the Invisible Hand concept, Adam Smith never believed humans are only driven by self-interest (Smith, 1822).

sense of behavior not following from the maximization of a utility function, but that agents are motivated by a concern about the effects of one's actions on others.

The present study is structured as follows. In Section 2 we will survey the evidence accumulated through the use of experimental games such as the ultimatum game and the gift exchange game that proves the existence of behaviors which the assumption of agents having self-regarding preferences is not able to explain. Section 3 provides a methodological defense against critics who argue against the use of laboratory evidence, such as that described in Section 2, to infer the determinants of human behavior. Following that, Section 4 introduces the Fehr-Schmidt model of inequity aversion, a model of other-regarding preferences that is able to predict the perplexing behavior documented in Section 2. Finally, in Section 5 we provide a motivating example for the use of other-regarding preferences by showing that their inclusion is able to explain the optimal choice between competing contracts under the existence of moral hazard.

2. Other-Regarding Preferences: Experimental Evidence

2.1. The Ultimatum and Dictator Games

The ultimatum game

The ultimatum game is a one-shot game between two players, a proposer and a responder. The proposer is given an integer amount of tokens, x, by the experimenter and must offer a share of it to the responder. If the responder accepts, the proposer's offer is implemented and both part ways with their respective payout. If the responder rejects the offer both players part ways with nothing.

If both players have self-regarding preferences the proposer's optimal strategy will be to propose the lowest possible amount that he is allowed to offer. Accordingly, the responder should accept whatever amount the proposer is willing to part ways with because otherwise he will be left with nothing rather than something. The subgame perfect Nash equilibrium for the ultimatum game is one where the payoffs are $(x - p, p)$, where p is the lowest possible amount that the proposer is allowed to offer. The experimental evidence, however, does not support this prediction.

Camerer (2011a) provides a detailed summary of the main results from a number of experiments using the ultimatum game. The main conclusions are as follows: The mean offer made by the proposer falls between 30% and 40% of the initial endowment. The median offer is 40 to 50%. There are rarely any unfair offers, that is, offers that fall in the 0 to 10% range. Offers that are too generous (i.e. more than half of the endowment) are also rarely observed. Low offers are often rejected, with offers below 20% being rejected about half of the time.

An increase in the stakes involved does not change the nature of the results. A possible objection might be that the stakes, or the monetary amount at stake in the interaction, are too low to elicit the required mental effort for players to play in the 'appropriate' manner. That is, if the stakes involved are low it is possible that players will not take

the game seriously. However, when the stakes are increased players continue behaving in ways that do not conform to the self-regarding prediction.

For example, Cameron (1999) conducted experiments using the ultimatum game in Indonesia where the largest monetary amount at stake was equivalent to about three times the average monthly expenditure of the participants. The authors conclude in this case that "significant deviations from game-theoretic behavior persist even in high stakes games." The one change in player behavior that the authors were able to observe was that responders were willing to accept a lower percentage offer, while there was no behavioral change from proposers.

Andersen et al. (2011) employ the ultimatum game in a poor village in Northeast India to study the effect of an increase in stakes on responder behavior. They are motivated by the finding that an increase in stakes does not elicit lower offers from proposers, which makes difficult the study of the effect an increase in stakes has in how responders deal with low, or unfair, offers. The authors increase the stakes by a factor of 1,000 — 20 to 20,000 rupees (1.6 to 16,000 hours of work) — and they alter the standard experimental instructions to elicit lower offers than usual from proposers. They find that responders play more closely to their predicted equilibrium response as stakes increase, usually as the amounts offered are equivalent to 30-40 days of wages or more. Rejection rates approach zero as the amount of money that responders must forgo with a rejection increases, meaning that stakes have their predicted effect. The authors point out that their finding confirms rather than rejects previous results given that one does not typically encounter situations where such high stakes are involved and the bulk of everyday market transactions are low-stakes affairs.

Slonim and Roth (1998) combine learning and increased stakes. Subjects from Slovak Republic play 10 rounds of the ultimatum game with stakes that are between 60 and 1,500 Slovak crowns. Their results confirm previous findings, that behavior in the ultimatum game does not confirm the equilibrium predictions.

A possible objection might be that the observation of behavior not consistent with the self-regarding equilibrium prediction rises from the reliance of sterile laboratory experiments with college students, implying that the results do not generalize to a wider population. Early cross-cultural experiments with college students from Israel, the United States, Japan, and Yugoslavia, confirmed the standard finding in ultimatum experiments where the predicted equilibrium is never met, though the results did show substantial differences between countries regarding the distribution of offers made by the proposer (Roth et al. 1991). These results provided some evidence that the deviation from the self-regarding prediction in the ultimatum game did generalize for populations all over the globe.

However, in 1996 a surprising finding broke the consensus when anthropologist Joe Henrich (Henrich, 2000) found that the Machiguenga, a slash-and-burn horticulturalist society living in the southeastern Peruvian Amazon, behaved in a way that was closer to the game-theoretic prediction. This "Machiguenga outlier" sparked the question of whether the behavior commonly observed in the ultimatum game was an artifact of the game being played by members of societies advanced in their economic development and propelled researchers to think about what economic and

cultural circumstances made it so that the Machiguenga found the modal offer of 15% a fair offer.

The answer to these questions came when a group of 12 anthropologists, including Henrich, adapted the ultimatum game, the dictator game, and the public goods game[2] so that these were not reliant on the administration through a computer and could thus be implemented in the field among non-literate subjects (Henrich et al., 2005). They proceeded to gather evidence from 15 small-scale societies exhibiting a wide variety of economic and cultural conditions.

In line with previous research, the predictions from the self-regarding model were not borne out in any of these societies, though there was wide variation in the results. The mean offers ranged from 26% to 57%, with the Machiguenga having the lowest mean offer and the Lamalera, a whale hunting people from near Indonesia, having the highest one. Indeed, the wide variation in how these societies approach the ultimatum game is quite interesting. The Hadza, a group of small-scale foragers from Tanzania, made low offers at the same time that they had a high rejection rate, while the Aché, from Paraguay, made consistently high offers with no rejections. The authors propose that this variation reflects their differing patterns of everyday life. Both groups share between members the meat that is obtained by hunters, though their levels of cooperation and expectations vary significantly. The Aché distribute their prey equally among all other households, and there is no consistent relationship between how much meat a hunter brings in and how much his family receives. Indeed successful hunters often leave their prey outside the camp to be discovered by others to avoid being considered boastful by their peers. By contrast, Hadza hunters sometimes wait until nightfall so they can sneak meat into their shelter, and when meat is shared between the group it is not done so without complaint and without some looking for opportunities not to share.

The authors reach the conclusion that increased sociality is dependent on the extent of the market integration in each society, that is, whether its people buy and sell wares and goods between one another and work for a wage. They find that increased cooperation in production is also associated with increased sociality, which might explain why the whale hunters of Lamalera feature such high levels of sociality, since it is necessary to sustain high levels of cooperation between multiple non-kin members to bring such an animal down. Taken these two aspects together, market integration and the payoffs to cooperation account for 66% of the variation in the outcomes in the ultimatum game.

More amusingly, Carter and Irons (1991) find that economists play closer to the standard self-regarding prediction than non-economists. But there does not seem be a difference between freshman and senior economists. Economists, it seems, are just different from everyone else!

[2] The public goods game is one where the subjects must decide how many tokens to contribute to a public good whose payoff will be equally distributed amongst all subjects and that is higher than the initial endowment. The standard prediction is that each subject will free ride. Experimental evidence shows that this prediction is only true if there is no opportunity for other subjects to punish the free riders.

The dictator game

The dictator game is a variant of the ultimatum game where the responder is forced to accept the proposer's offer regardless of the amount proposed. If the proposer has self--regarding preferences we would predict for him to propose the lowest denominator he is allowed to since there is nothing to be gained by offering a higher share of the endowment.

Camerer (2011a) offers a summary of the results from multiple experiments that have employed the dictator game. The mean offer across these studies is roughly 20% of the initial endowment, and about 60% of the subjects in these studies offered a positive amount of the endowment.

That the proposer in the dictator game makes a mean offer that is higher than the minimum required, though lower than the mean offer in the ultimatum game, provides us with knowledge about the motives behind the offers made in the ultimatum game as well as the nature of those made in the dictator game itself. Given that the only meaningful difference between the dictator and ultimatum game is that in the first the ability of the responder to reject the offer made is removed, we can infer from the lower mean offer that strategic concern drives at least a portion of the offer made in the ultimatum game. That is, in the ultimatum game, the proposer offers more than he would otherwise have offered due to a fear of his offer being rejected.

However, that the mean offer in the dictator game is not the minimum required tells us that this strategic concern does not entirely drive the offer in the ultimatum game. Given that the proposer is made worse off by offering more than minimum, and since the responder is a passive actor in this interaction, we can interpret the offer made in the dictator game as being driven by an aversion to inequality, or altruism.

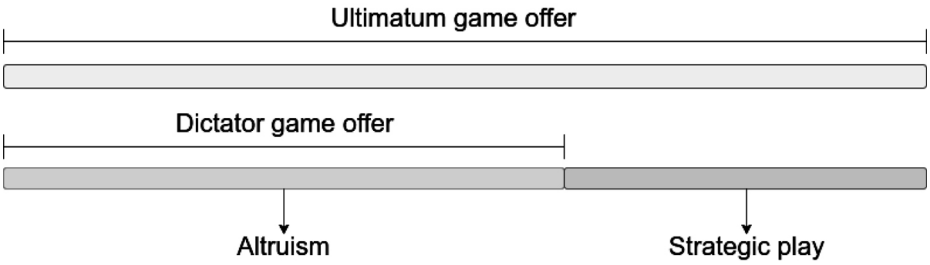


Figure 1: The two components that drive the ultimatum game offer

List (2007) pushes back against the standard interpretation that positive offers in the dictator game reflect altruism and/or inequity aversion from the part of the proposer. For example, lower offers are seen when anonymity between proposer and responder is added, indicating that a concern about how one is seen by their peers is a driving force for the positive offers seen in the dictator game. List (2004) also find that, in the public goods game, the more anonymous decisions were amongst subjects the less the subjects opted to give in the one-shot version of the game.

Dana, Cain, and Dawes (2006) consider a variant of the dictator game where the proposer is given an initial endowment of $10 and, after having made his choice, is offered the option of exiting the game with $9. The exit option leaves the responder with nothing but ensures that he never knows that the game has been played. Even though proposers could get a higher payoff by engaging the receiver in a dictator game and not offering anything, 28% of the proposers opted for the exit option, perhaps because they didn't want to appear unfair to the receiver were they to enter the dictator game. In their second experiment the receiver never knows whether the money offered to them comes from the proposer or from the experimenters, thus allowing the authors to determine with more clarity whether appearing to be fair is indeed a concern for proposers. They find that 9 out of 24 proposers exited, which does imply a significant minority of proposers is concerned about not appearing self-regarding to the receivers.

In Dana, Weber, and Kuang (2007) a variant of the dictator game is played. Proposers are sorted into two different treatments, the baseline and the hidden payoff treatment. In the baseline treatment, the proposer can choose one of two actions, A and B, with respective payoffs (6,1) and (5,5) for the proposer and responder respectively. In the hidden payoff treatment the payoff for the responder is uncertain, so proposers must choose between actions A and B where the payoffs are shown to them as (6, ?) and (5, ?). All subjects are told that the payoffs from A and B are equally likely to be either (*i*)(6,1) and (5,5), or (*ii*)(6,5) and (5,1). The proposer can, at no cost to himself, choose to reveal the payoffs by clicking a button on the computer screen, and the responder is not made aware of that this choice has been made. The prediction is that if altruism is a better motivator for the proposer's actions the proposer will choose B in (*i*) and A in (*ii*).

By comparing the proportion of proposers that chose option B in the baseline treatment with the proportion in the hidden payoff treatment that chose to reveal the payoffs the authors are able to infer whether inequity aversion is an important motivation behind the offers.

They find 14 out of 19, or 74%, of proposers in the baseline treatment chose the more generous option B. However, in the hidden payoff treatment, 56% did not choose to click the button to reveal the payoffs, a difference in proportion that is statistically significant. These differences imply that the appearance of being fair is an important determinant in the offers made in the dictator game and that inequity aversion does not provide a full explanation. It is possible then that at least part of the positive offers in dictator games are made not because proposers are altruists but because they are reluctant altruists. They want to appear to be altruists to everyone else but they would much rather keep the money to themselves.

2.2. Gift Exchange and Trust Games

The gift exchange game

The gift exchange game was introduced by Fehr, Kirchsteiger, and Riedl (1993) in an attempt to empirically investigate whether the notion of fairness held by agents impeded the formation of a market clearing equilibrium in labor markets, a topic first broached by Akerlof (1982).

In the gift exchange game, two players are each assigned one of two roles: a firm or a worker. The firm offers a wage w to the worker, which the worker can then reject, in which case both earn nothing, or accept, in which case the worker must now expend an effort level, e, of his choice.

The standard prediction in such a setting can be discerned using a neoclassical model. Suppose the firm decides to offer the same wage to all its workers, $\omega = \varpi$. Workers have a utility function, $u(\omega,e)$, where ω is the wage rate and e is the effort level they expend. The firm dictates that workers provide a minimum effort level in exchange for their wage, $e_{\min}$. Workers, mindful of the firm's work rules, should choose their effort such that it maximizes:

$$u(\omega,e)$$

subject to the constraints

$$\omega = \varpi$$

and

$$e_{\min}.$$

This maximization problem yields the prediction that workers will choose the lowest effort level possible, $e_{\min}$. The firm, aware of this, will set ϖ as low as possible in an effort to maximize profits.

In his paper, George Akerlof is motivated to explore the effects of fairness in the formation of involuntary unemployment due to the curious results from a study of social relations among workers at a utility company in the eastern United States (Homans, 1954).

In this study, a group of women were found to be exceeding the minimum work requirements set by the firm by a considerable margin, a behavior that the neoclassical model above does not predict. Akerlof envisions this seemingly perplexing behavior as the result of the firm and the workers modelling their relationship as a "gift" exchange mediated by endogenous social norms. The workers offer a "gift" to the firm in the form of additional effort level, and in exchange the firm offers a "gift" in the form of a wage that the workers consider fair and that is in excess of what they could receive were they to leave their jobs. Thus, a labor market equilibrium is created where workers work harder because they are paid above opportunity cost. This wage level is higher than the market clearing one, ensuring that unemployment is present in equilibrium.

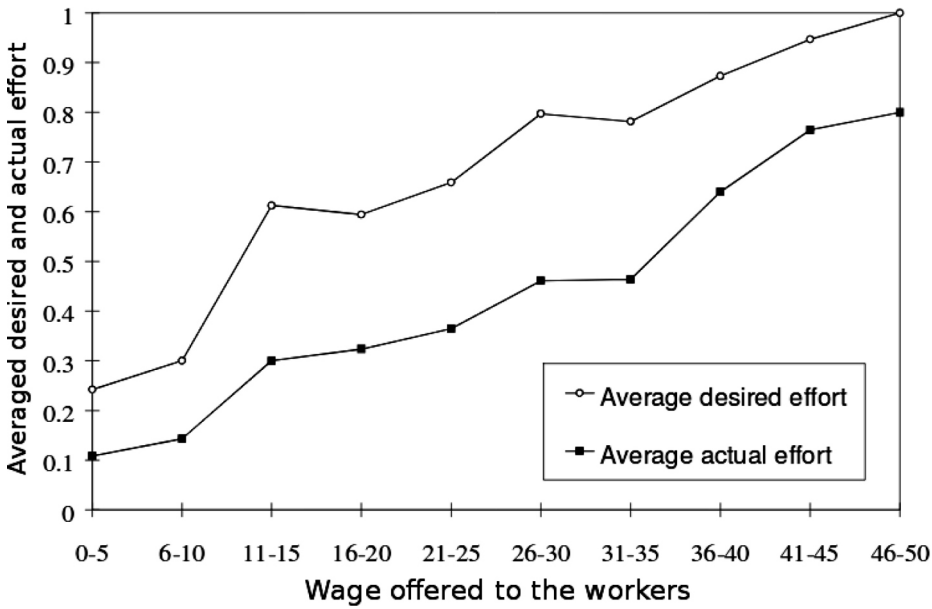
The gift exchange game permits us to study the level of intrinsic reciprocity in social relations such as the one described previously. This reciprocity falls in the category of other-regarding preferences.

Consider the following experiment from Fehr, Gächter, and Kirchsteiger (1997) Subjects were assigned into one of two roles: a principal or an agent. Identities were kept anonymous, so no reputation building was possible. Principals make a job offer to the group of agents, meaning that principals stand in for employers and agents for workers. Agents are given the option to accept or reject the offer, and in an effort to spur competition there are more agents than principals. The job offer consists of an incomplete contract, (w_b, e_n), that specifies a binding wage level, w_b, and a non-binding effort level, e_n. The choice of the effort level is represented by the choice of a number in which the higher the number the higher the effort is, and the higher the costs borne

by the agent are. Nothing in the experiment impedes agents from choosing an effort level that is lower than the proposed effort level in the contract as there is no punishment for doing so.

The expected behavior for both workers and firms are as noted earlier: agents will choose the lowest possible effort level and principals, knowing this, will offer the lowest possible wage level. However, if the principal believes there are sufficiently many reciprocal agents, he has an incentive to offer higher wages in an attempt to induce higher effort levels from the agents in reciprocity. Additionally, agents may induce reciprocity by the firms by offering a higher effort level than the one initially proposed.

Figure 2: Relation of desired and actual effort to the rent offered to workers



Source: Fehr and Falk (2002).

The experimental results are depicted in Figure 2. Two conclusions follow:

1. Higher desired effort levels are associated with more generous offers to the workers, which suggests employers try to elicit reciprocal responses from the workers.
2. *On average*, the workers respond reciprocally to the employer's higher offers, though there is always a certain amount of shirking present.

The authors further add that *"there is also a substantial fraction of selfish workers who always choose the minimal effort or who rarely respond in a reciprocal manner."* The authors summarize

the evidence from multiple studies to suggest that the fraction of self-regarding agents lies between 40% and 60%.

While we will take up the issue of the external validity of laboratory experiments in a later section, it is worthwhile to point out some of the pushback against the main conclusions of the gift exchange game that have arisen from the results gathered from the use of field studies to study reciprocity.

Gneezy and List (2006) hired students to a data-entry job where they would enter books into a library information system. Each student performed the task alone, and were offered $12 for the job. In the experimental treatment, after the training phase, a portion of the students were informed they would receive $20 per hour, with no explanation for the increase in pay. Students in the control condition were payed the previously agreed upon $12.

The results seemed to cast doubt over the idea that offering a wage premium is an effective measure to elicit higher worker performance. In the first 90 minutes, those workers in the treatment condition produced around 25% more than their control peers. Although this percentage difference in effort is noteworthy, the increase in effort vanished as the experiment continued and effort levels for both treatment and control conditions were found to not be significantly different. The authors interpret their results as showing that while higher wages are reciprocated by greater effort on the part of the workers, this higher effort is not persistent and thus we need be careful to extrapolate from the single round interactions featured in the gift exchange game to how these relationships actually develop in the real world.

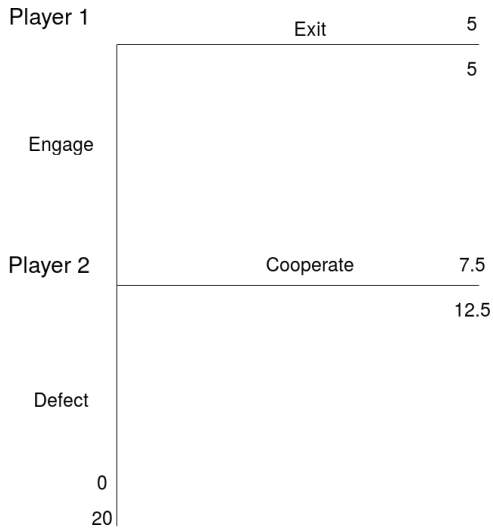
A problem with Gneezy and List (2006) is that of a small sample size, which limits their ability to detect statistical significance if the effect of a wage premium on effort is modest or small, a point mentioned by Fehr, Goette, and Zehnder (2009). Indeed, Cohn, Fehr, and Goette (2008) use a larger sample size, which allows them to have enough power to detect a statistically significant increase in effort from the increased wage, not replicating Gneezy and List (2006). Fehr, Goette, and Zehnder (2009) surveys the literature and concludes the positive relationship between wage and effort to be robust and well replicated.

The trust game

The *trust game* is played between an *investor* and a *responder*. Each player is endowed with a fixed amount of tokens, x. The investor must decide an amount $i \leq x$ to send to, or invest with, the responder. Before the amount chosen is delivered to the responder, the experimenter multiplies it by a multiplier m, meant to capture market return, and passes it on to the responder. The responder must then return an amount $r \leq mi$ back to the investor.

If both subjects have self-regarding preferences then the responder will never send any money back to the investor. The investor, correctly anticipating the responder's behavior, will decide not to invest any amount i.

Figure 3: The trust game



We can use a concrete example to prove this prediction. Let us assume a game with two players, Player 1 and Player 2. Both are endowed with \$5. Player 1 can decide between keeping his endowment, in which case the game is ended and both players walk off with a payoff of \$5, or he may pass the entire endowment to Player 2. If the latter, the endowment is tripled by the experimenter and it is then up to Player 2 to decide whether to keep the additional \$15 for himself, or return, for example, \$7.5 to Player 1. The payoffs are, respectively, (\$0, \$20) and (\$7.5, \$12.5).

The subgame perfect Nash Equilibrium of the trust game for the self-regarding preferences model can be determined using backward induction. In the second stage of the game, Player 2 maximizes his payoff by defecting and walking off with the full amount. Predicting this, Player 1 will not send his endowment to Player 2. Thus, the predicted payoff will be (\$5, \$5), that is, both players walk out with their initial endowment, having not cooperated. Traditionally, trust game experimenters allow for both players to choose how much they intend to send to the other player, but this does not change what the subgame perfect equilibrium is.

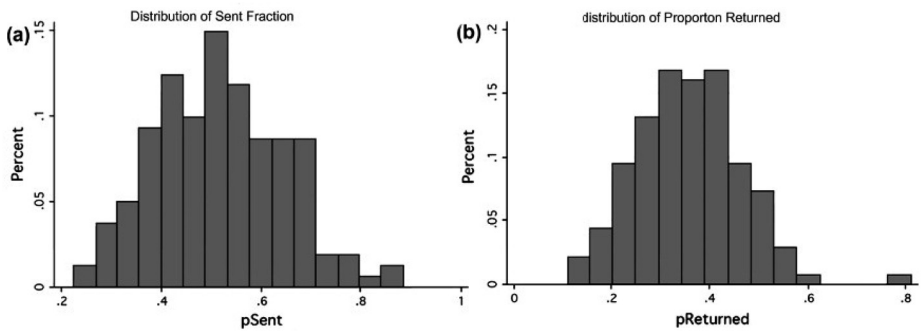
The share of the endowment the investor decides to invest is said to capture trust and the share sent by the responder, trustworthiness. Both are forms of other-regarding preferences.

If both players have self-regarding preferences, then the subgame perfect equilibrium will be met. However, if both players have other-regarding preferences we should see a positive amount sent by the investor, $i > 0$, and a positive amount returned by the responder, $r > 0$. Figure 3 shows the distribution of offers made by both the investor and responder in

a meta-analysis of 161 studies involving approximately 24,000 participants (Johnson and Mislin, 2011).

The meta-analysis found that the mean offer made by the investor across all studies is .502, around half of the initial endowment, while the mean amount returned by the responder is .372. Due to aggregation of multiple experiments from multiple parts of the world the authors were able to study the differences in how people from around the world play the trust game. They find Africa sends and receives the lowest amount of all continents, with North America and Europe featuring the highest amounts both sent and received. They find further that older people send larger amounts, students send significantly lower amounts than non-students, and that amounts sent are larger if the subject believes he is playing with another human player.

Figure 4: Distribution of percentages sent by investors (left) and responders (right)



Source: Johnson and Mislin (2011).

–

3. On the Validity of Using Experiments in Economics

> "One possible way of figuring out economic laws ... is by controlled experiments... Economists [unfortunately]... cannot perform the controlled experiments of chemists or biologists because they cannot easily control other important factors. Like astronomers or meteorologists, they generally must be content largely to observe."
>
> Samuelson and Nordhaus (1985)

Laboratory experiments are a widely used tool in the physical and life sciences. In contrast, the social sciences have traditionally been considered nonexperimental, that is, the data upon which social scientists base their theories are collected not through experimentation, but observation. This is due to the obvious constraints of this class of scientific disciplines (e.g. historians are not able to recreate in a lab the Napoleonic Wars). This is not to say the

social sciences have not or do not use laboratory experiments when it is feasible to do so. Psychology, for instance, has used laboratory experiments, or the experimental method more broadly, successfully for more than two centuries now (e.g. Ebbinghaus, 1885). In economics, however, the use of laboratory experiments is a recent development.

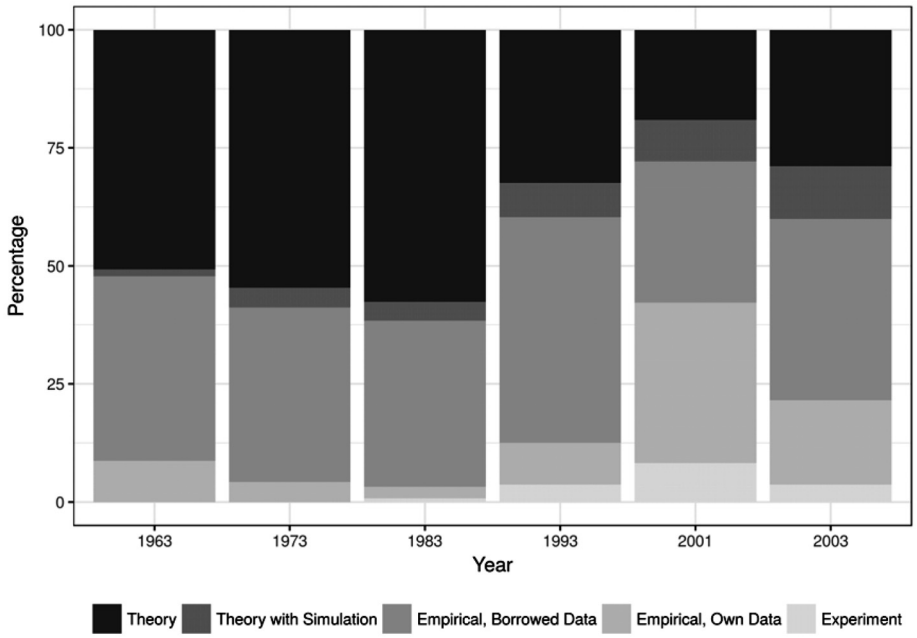
Although there was some proto-experimental work done in the 1930s on the topic of consumer demand theory (Moscati, 2007), it is generally agreed that, as an institutional and intellectual programme, experimental economics took form in the late 1940s following the publication of John von Neumann and Oskar Morgenstern seminal *Theory of Games and Economic Behavior* in 1944 (Guala, 2008). Since then, the growth of published papers using laboratory experiments has been remarkable.

In three of the most prestigious economics journals — *American Economic Review*, *Econometrica*, and *Quarterly Journal of Economics* — the fraction of experimental papers published in proportion to all published papers was between 0.84 and 1.58 in the 1980s, jumping to 3.8 and 4.15 between 2000 and 2008 (Falk and Heckman 2009). The first specialty journal, aptly named *Experimental Economics*, was founded in 1998. Moreover, 6 of the Nobel economics prizes awarded since 1969 have been to economists who can be described as working in experimental economics, including heavyweights such as Elinor Ostrom, Daniel Kahneman, and, more recently, Richard Thaler. The rise in prominence of experimental economics has been such that, despite the above quote from Samuelson and Nordhaus, in the 1992 revision of their famous textbook they saw the need to further add that experimental economics is "an exciting new development" (Samuelson and Nordhaus, 1992, p. 5).

There is however an elephant in the room. Despite the evident growth in the use of laboratory experiments in Economics, there have been challenges regarding whether the results coming out from these experiments allow researchers to say anything about economic behavior outside the lab. The issue of external validity, the ability of experiments to provide findings that are likely to allow for reliable inferences outside the laboratory, is a pernicious problem for the social sciences that does not exist to the same extent in the physical sciences.

By way of illumination, a Physics student performing a careful experiment to determine the value for Earth's gravity need not concern himself as to whether his results will generalize to outside the lab. The same is not necessarily true in the social sciences, in general, and for most of the experiments we have surveyed above, in particular. While it does seem that subjects offer around 20% of their endowment in the dictator game, we do not regularly see strangers in the streets spontaneously offering a fifth of the contents in their wallets to strangers passing by.

Figure 5: Methodology of articles in top economics journals, as percent of total



Source: Mearsheimer and Walt (2013).

If we are to use the extensive evidence surveyed previously to argue for the existence of other-regarding preferences we must first establish that it offers us reliable evidence that extends beyond the artificial conditions of the laboratory. Indeed, it is this artificiality that, while imbuing laboratory experiments with their unique methodological strength, also gives it the weakness that has been an influential source of skepticism about their use as a tool in the economist's tool box.

Laboratory experiments are often contrasted with field experiments. In the debate between the two, field experiments are often touted as possessing more 'realistic' conditions, even if they are perhaps less tightly controlled.

List (2006) provides an interesting example of this idea. In his paper, a gift exchange game is played where buyers make price offers to sellers and, in return, sellers select the quality of the good they will exchange with the buyer. The experiment was run in a standard laboratory setting and used experienced sports-card traders as subjects. The results mirrored the typical findings for this type of game: in the presence of higher offered prices, sellers tended to offer higher quality goods in return, even though they were not obligated to do so by the rules of the experiment. Thus this laboratory experiment points toward the existence of other-regarding preferences in the gift exchange game.

The experiment was then carried out by making a single change in which the goods exchanged where actual baseball cards whose market value was influenced by differences in their physical condition that experienced card sellers are more likely to detect than unexperienced sellers. In this experiment, other-regarding preferences were also observed. Higher quality cards were offered to buyers who offered higher prices.

The two experiments are therefore concordant in their conclusion of the existence of other-regarding preferences. However, List did not stop there. He wanted to know whether his results would also be observed in the card sellers natural environment, a sports-card show. Dealers in this field study were unaware that their behavior was being observed and studied. Confederates were instructed to approach sellers and offer different prices in return for sports-cards of varying quality, mirroring the methodology of the previous experiments.

In this field study, where the dealers did not believe that the consumers could grade the cards appropriately or there was little possibility of future interaction, little statistical relationship between price and quality was observed. Only when there was concern for one's reputational standing, when sellers expected future interactions to happen or buyers could verify the quality of the cards by using a third-party, was high offered price met with high quality offered. Thus, the other-regarding preferences routinely observed in the lab were attenuated or not observed in the field condition.[3]

The dichotomy of results between the laboratory and field conditions should certainly make us pause before claiming with certainty that other-regarding preferences matter in an economic context. It is tempting to think that field experiments reflect a more realistic condition and should thus be held in higher regard when making inferences about the inexistence of other-regarding preferences. Thus, that card dealers do not seem to match quality with price in a sports-card show seems to suggest that we should not change our priors with regards to the existence of other-regarding preferences. It is, however, worth thinking more deeply about this notion that field studies offer us a more realistic picture than laboratory experiments.

Camerer (2011b) makes the helpful distinction between the *policy view* and the *scientific view*. In the policy view, the generalizability of experimental findings to the outside world, that is, their *external validity*, is of paramount importance. This is because in choosing what policy to apply, evidence that has been collected in the same domain as the policy has obvious advantages in the inferences one can draw from it with respect to the effects of policy in question. Field experiments should be given more weight in this view than laboratory experiments do. For the scientific view, however, both laboratory and field studies constitute ways to enhance our understanding of human behavior and should therefore hold equal weight in the inference process. Provided the evidence was properly gathered and is valid and contextually relevant, there is no hierarchical relationship between the two methodologies. Both constitute tools to be used in the accumulation of knowledge. As Camerer puts it: "in this view, since the goal is to understand general principles, whether the 'lab generalizes to the field'... is distracting, difficult to know..., and is no more useful than asking whether 'the field generalizes to the lab'."

[3] Camerer (2011b) provides a critical reply to List (2006) and reanalyzes the data with different statistical techniques. He notes that "...the conclusion that the lab and field show different reciprocity is suggestive but is just not robustly significant."

To illustrate this point, consider the following formalization from Falk and Heckman (2009). Suppose a variable of interest Y can be fully explained by the variables X_1, ..., X_N and that the functional relationship between them is given by $Y = f(X_1, ..., X_N)$, known as the all-causes model. Suppose we are interested in examining the causal effect of X_1 on Y, which requires us to hold all other variables $\widehat{X} = (X_2, X_3, ..., X_N)$ fixed.

In a laboratory experiment the researcher estimates a model of the form $Y = f(X_1, X^L)$ where $X^L \neq \widehat{X}$. Following the same logic, in a field experiment one estimates a model of the form $Y = f(X_1, X^{F_1})$ where $X^{F_1} \neq \widehat{X}$ and typically $X^{F_1} \neq X^L$. The claim is usually made that does not satisfy external validity but $f(X_1, X^{F_1})$ does.

We can write the field study in List (2006) as $Y = f(X_1, X^{F_1})$ where Y is the quality of the cards, X_1 is the price offered, and X^{F_1} are the remaining variables. Suppose we were to repeat that same experiment using a different subject pool, possibly stamp collectors. This gives rise to yet another estimated relation $Y = f(X_1, X^{F_2})$ where X^{F_2} reflects the set of variables and characteristics in this new experiment, including the new subject pool. How are we to adjudicate between $f(X_1, X^L)$ and $f(X_1, X^{F_1})$ if we want to predict the causal relationship between X_1 and Y in the new relationship $Y = f(X_1, X^{F_2})$?

One might be tempted to reply that the field experiment should hold more weight than the laboratory experiment in this particular inference since the new situation also involves a field experiment. There is not, however, an explicit reason for why this should be so. While both situations are field experiments, there is no a priori reason to expect the behavior from sports-card dealers to generalize toward that of stamp collectors. As Camerer (2011b) puts it (using different notation), "if the litmus test of 'external validity' is accurate extrapolation to X^{F_2}, is the lab X^L necessarily less externally valid than the field setting X^{F_1}? How should this even be judged?"

It is best then, under the scientific view, to treat laboratory and field data as complementary. Both have their strengths, and the usefulness of one versus the other is ultimately a matter of the underlying research question. Laboratory experiments, due to their tight control, are more prone to be replicable, whereas in field experiments replicability can be challenging and is more often than not impossible. The smaller cost of laboratory experiments, as well as the easiness of archiving and reproducing instructions, software, recruiting protocols, databases, and statistical tools, also make laboratory experiments easier to replicate.

The two methods differ in what variables $\widehat{X}$ they are able to provide a larger variation for. Laboratory experiments can explore the parameter space for values that can be hard or rare to find in the field. For example, in Andersen et al. (2011) the authors raise the stakes in the ultimatum game by a factor of 1,000 such that the highest stakes in play equal 16,000 hours of work, an amount that would be nigh impossible to find in the field. Field experiments do a better job at collecting evidence for different subject pools with different demographics and social characteristics, as we saw in the multidisciplinary work of Henrich et al. (2005) who employed the ultimatum, dictator, and public goods game in 15 small-scale societies.

Having argued for the usefulness of laboratory experiments, there still remain some concerns that need be addressed before we can be safe in using the experimental evidence to argue for the existence of other-regarding preferences. While we have seen that there is no a priori reason for laboratory experiments to not be used in the making of inferences

about economic behavior in lieu of field experiments, this is only so if laboratory experiments are a valid tool for the study of this particular subject. That is, while by themselves laboratory experiments are a valid tool in the economist's tool box, there might be flaws in the experimental process that invalid the use of the results gathered to infer the existence of other-regarding preferences. A powerful critique of the experimental method is offered in Levitt and List (2007) where the authors raise legitimate concerns that put into question the external validity of the observed results.

By the order in which these objections will be tackled, they are:

1. There is unprecedented experimental scrutiny in laboratory experiments. This may give rise to *experimenter demand effects* where subjects, perceiving that they are being observed by the experimenter, may behave in ways that they believe the experimenters desire, or may also behave in ways that end up not revealing their true preferences, e.g., being observed may lead to more prosocial behavior than the subject actually desires.
2. Human behavior is context-dependent and it is not clear that experiments can either capture or control this.
3. There is the possibility of self-selection bias where experiments might be being run with an homogeneous sample of students who might be more prosocial, more educated, and have a higher need for approval than the average human population.

Levitt and List additionally question the common use of small stakes in laboratory experiments arguing that it might not capture the richness of human behavior. We choose not to deal with this objection because the previous section already dealt with the effects of varying stake size in the experiments. Indeed, the issue of stake size has long been a topic of interest in behavioral economics (Camerer et al., 1999) and is no longer a novel nor potent objection.[4]

Experimenter demand and audience effects

Subjects in laboratory experiments know that their behavior is being recorded and will be under intense scrutiny. This intense obtrusiveness might lead to subjects matching their behavior with what they perceive to be the experimenter's desired behavior (experimenter demand effect), or may lead subjects to behave in more prosocial ways because they believe that self-regarding behavior may be frowned upon by the experimenter (audience effect).

Let us assume for a moment that subjects hold an accurate view of what the experimenters expect and favor a particular outcome, which is not obvious and is something that experimenters are aware of when designing experimental procedures and thus work to circumvent. For there to be a demand effect it is necessary that the subject be willing to sacrifice his earnings by behaving in the way that the experimenter desires. Even if the subject is willing

[4] As Colin Camerer writes in his 2003 textbook, five years before Levitt and List's paper: "If I had a dollar for every time an economist claimed that raising the stakes would drive ultimatum behavior towards self-interest, I'd have a private jet on standby all day (Camerer 2011a, p. 60)."

to do so, that willingness is but a component of his overall preference bundle, meaning that it should be possible to devise a situation that stresses his initial desire and makes him more reluctant to sacrifice his earnings. That increasing the stakes involved in the experiments typically has little effect suggests that demand effects are not a strong concern.

This still leaves us with the possibility that the intense experimental scrutiny leads to subjects behaving more prosocially. Barmettler, Fehr, and Zehnder (2012) employ a novel experimental procedure that allows the manipulation of experimenter-subject anonymity and employ it in three experimental games: the ultimatum, the dictator, and the trust game. In none of these, for any player in any role, is there a statistical significant difference in the choices made between the treatment condition where anonymity between subject and experimenter is maintained and the one where it isn't.

Experimenter demand and audience effects do not seem to be strong objections against the experimental evidence. Indeed, if experimental scrutiny made such a noticeable effect on the emergence of other-regarding preferences that would itself be evidence for the existence of other-regarding preferences for we are regularly being directly observed in many of our real-world interactions. More distopically, if scrutiny was a powerful enough force to flip a person's switch towards behaving more prosocially, it is likely we would be seeing intense institutional efforts to promote that scrutiny.

Context dependence

Human behavior is embedded in a rich, complex, and tangled web of social norms, frames, and the lessons learned from past experience. Levitt and List argue that different experimental procedures, such as writing the instructions the subjects read in different ways, may lead to differing results. Defection rates in the prisoner dilemma game vary depending on whether subjects are playing a "Community" or "Wall Street" game (Ward, 1997), framing the allocation of funds in a public goods game as a "contribution" or "allocation" matters, as does whether the game is framed as a positive externality or a negative one (Andreoni, 1995).

However, even though Levitt and List make a persuasive case for the importance of the context, they also argue that this context is "is not completely controlled by the experimenter." This collides with their previous argument because if it is possible to elicit different results by varying some of the experiment's parameters then it follows that it is also possible to control and account for that context. Indeed, that variation is highly desirable in the scientific view since it allows for the accumulation of knowledge about the boundary conditions of human behavior. Thus the very idea that experimental context might influence the results is an argument for more experiments to be run.

Moreover, to the extent that there is a subset of context that is not liable of being captured by any methodological variation, then it is also unlikely that uncaptured context is possible to control for in a field experiment. The problem of uncontrolled variables that cannot be measured and controlled for is not a problem of laboratory experiments per se, for field studies also run into it, but is instead a problem that every empiricist must wrestle and contend with.

Self-selection bias

The empirical evidence we have surveyed is largely based on laboratory experiments using self-selected students. This homogeneous sample might constitute a problem if students are found to behave in systematically different ways than the rest of the population. If the students who self-select into experiments behave more prosocially then the results provide a biased estimate of the extent to which there are other-regarding preferences in the population.

Falk, Meier, and Zehnder (2013) provide an interesting study of this question where they are able to distinguish whether students who self-select into laboratory experiments are any different from those students who don't, as well as whether these have different social preferences from the rest of the population.

They analyze the decisions of a sample of 16,666 undergraduates at the University of Zurich for which they know who participated in experiments and how many times. To measure the extent of their prosocial inclinations they use data collected from a naturally occurring repeated decision where each student must decide whether or not he or she wants to contribute a pre-determinate amount to two social funds which provide charitable services. The authors conclude that participating students do not have different social preferences than their nonparticipating colleagues. If there is a bias then it is because students are different from everyone else.

To see if such a difference exists the authors employ two identical trust games using distinct subject pools so that the only difference in prosocial behavior comes from differences between the two subject pools employed and not changes in experimental design. One group was recruited from the student pool at the University of Zurich and the other from a representative sample of the population of the city of Zurich. In total 1,296 participants were recruited (295 students and 1,001 from the general population)

They find that the non-student subject pool behaves more prosocially than the student subject pool which implies that by regularly employing students in their experiments researchers might be downwardly biasing their inferences about the existence of other-regarding preferences, i.e., to the extent that this is an issue is more one of magnitude than direction. The finding that students exhibit less prosocial behavior than the rest of the population is consistent with a range of other evidence, e.g., CEOs tend to be more prosocial than students (Fehr and List, 2004).

Indeed, concerns about the homogeneity of the subject pools used in laboratory experiments being problematic has been dealt with powerfully with the experiments performed by Henrich et al. (2005). Even though they report wide variation in the extent to which people in different societies have other-regarding preferences, the fact that none of the societies studied confirmed the self-regarding predictions implies that other-regarding preferences may well be a general feature of human nature.

4. Modelling Other-Regarding Preferences

The evidence surveyed previously establishes that agents often have preoccupations not only about what happens to themselves but also with what happens to other people. Economic models often do not include these social preferences in their structure, possibly limiting the set of behaviors they are able to explore. At worst, by not taking into account the existence of other-regarding preferences these models may reach incorrect conclusions about the economic behavior of agents.

There is thus the need for a theoretical model that takes into account the empirical findings on other-regarding preferences. The Fehr-Schmidt model of inequity aversion developed in Fehr and Schmidt (1999) is a proposal for such a model. A description of this model and an application to the ultimatum game follows.

4.1. The Fehr-Schmidt Model of Inequality Aversion

Consider n individuals, each with a respective monetary payoff Y_1, Y_2, ..., Yn. The payoffs of all individuals but the individual i is denoted by the vector $\vec{Y}$-i. For any i, the Fehr-Schmidt utility function, henceforth FS utility function, is defined as:

$$U_i\,(y_i, \vec{y}_{-i}; \alpha_i, \beta_i) = y_i - \frac{\alpha i}{n-1}\Sigma_{j\neq i}\{y_j - y_i, 0\} - \frac{\beta i}{n-1}\Sigma_{j\neq i}\{y_i - y_j, 0\},$$

where $\alpha_i \geq 0$ and $0 \leq \beta_i < 1$. It is easy to see that the FS utility function describes an individual with self-regarding preferences if $\alpha_i = \beta_i = 0$, in which case $U_i\,(y_i, \vec{y}_{-i}; \alpha_i, \beta_i) \rightarrow U_i\,(y_i)$.

The FS utility function models the individual as comparing his situation with the situation of the individuals around him. That is, the individual exhibits self-centered inequity aversion, where some people are better off than him and he his better off than other people. This is captured by the second and third terms in the function, respectively.

The second term in the FS utility function measures the utility loss from disadvantageous inequality, more colloquially called envy, while the third term measures the loss from advantageous inequality, or altruism. It is assumed that $\alpha_i \geq \beta_i$. This means that, for the same magnitude, an individual loses more utility from another individual being better of than him than in the contrary situation. Envy is a more psychologically salient condition than altruism.

While there is no upper bound on α_i an individual with $\beta_i \geq 1$ reduces his advantage over other individuals he is increasing his utility by more than he is reducing his advantage. This seems implausible as it would require extremely high levels of altruism from the individual. Eckel and Gintis (2010) report the magnitudes of α_i and β_i from various studies. The evidence indeed suggests that for most individuals $\beta_i < 1$ and $\beta_i < \alpha_i$.

We will now see how the Fehr-Schmidt model can be used to understand the behavior seen in the ultimatum game that was previously unexplainable using models with just self--regarding preferences.

The ultimatum Game

A proposer and a responder play the ultimatum game. They are, respectively, Player 1 and Player 2. Of the full endowment researchers award to Player 1, the share of it that is proposed is denoted by S.

The responder accepts all offers $S \geq 0.5$. There is a critical share, $S_c < 0.5$ such that the responder rejects all offers below it and accepts all offers $S \geq S_c$.

If $S > 0.5$, taking into account the FS utility function for the case with only 2 players, we have the following for the responder:

$$U_2 = S - \beta_2 [S - (1 - S)],$$

which is positive because $\beta \in [0,1]$, hence the responder will accept.

Now suppose $S < 0.5$. In this case we have

$$U_2 = S - \alpha_2 [(1 - S) - S] = S(1 + 2\alpha_2) - \alpha_2.$$

For this to be positive we need S such that

$$S \geq \frac{\alpha_2}{1 + 2\alpha_2}.$$

Taking $\alpha_2 \to \infty$ reveals that the critical threshold, S_c, is 0.5.

The equilibrium share offered by the proposer is given by:

$$S = \begin{cases} S_c & \text{if } \beta_1 < 0.5 \\ 0.5 & \text{if } \beta_1 > 0.5 \\ S \in [S_c, 0.5] & \text{if } \beta_1 = 0.5 \end{cases}$$

From the previous lemma we know that the responder will accept any share $S_c \leq S \leq 0.5$. Let us consider such a share. From the FS utility function we have, for the proposer, $U_1 = (1 - S) - \beta_1 [(1 - S) - S]$. Taking the first derivative with respect to S leaves us with $2\beta_1 - 1$. Thus, if $\beta_1 < 0.5$, we have $\frac{\partial U_1}{\partial S} < 0$ so the proposer should offer the minimum possible that the responder will accept, i.e., S_c.[5]

If $\beta_1 = 0.5$ we have $\frac{\partial U_1}{\partial S} = 0$ so any feasible share between S_c and 0.5 may be offered and will be accepted.

For values of β_1 higher than 0.5, $\frac{\partial U_1}{\partial S} > 0$, so we have $S_c = 0.5$.

Given what we know about the results usually observed from playing the ultimatum game, we can see that the Fehr-Schmidt model matches the experimental results reason-

[5] The question of how the proposer comes to know S_c is an important one though we can assume that the proposer does know it. Henrich and Henrich (2007) provides a valuable overview to how humans come to learn how to successfully cooperate with their peers.

ably well. Fehr and Schmidt note that their model "yields too extreme predictions in... the dictator game", where proposers only offer high offers ($S = 0.5$) or very low offers ($S = 0$), a prediction rejected by the data.

They note that their model can easily be modified so that the assumption that inequity aversion is linear can be dropped and substituted by a utility function that is concave in the amount of advantageous inequality. With this new assumption the results of the dictator game can also be accommodated.

5. Contract Design Under Other-Regarding Preferences

The results from the gift exchange game surveyed previously indicate that reciprocity influences the relationship between principal and worker in that the principal can elicit more effort from the worker by offering a higher wage even though there are no guarantees that the worker will reciprocate by exerting a higher effort. This leads us to question how these deviations from the self-regarding model influence how best to structure the incentives that mediate the relationship between the two parties.

In this section we will apply what we have learned about other-regarding preferences to understand how they affect the optimal choice between competing types of contracts under the existence of moral hazard. More colloquially, we are interested in knowing whether it is best for the principal to reach for the carrot rather than the stick.

Consider a principal who contracts an agent to work for him. The agent can expend effort $e \in [\underline{e}, \bar{e}]$ at a cost $c(e)$ such that $c' > 0$ and $c'' > 0$, where $c(e)$ denotes the effort cost function. The principal wants the agent to expend $e_{\min}$, which he introduces in the contract as the contracted effort level. However, $e_{\min}$ is non-binding since the agent's effort is not verifiable. Because the principal might want to provide evidence to a third party in the case where the worker is expending less effort than the contracted one the principal might invest in a verification technology that costs k. This technology is able to provide evidence of shirking with a probability $p \in [0,1]$ which the principal can then use to impose a fine f on the agent.

In Fehr, Klein, and Schmidt (2007) a cohort of individuals are randomly selected into the roles of principal and agent and are then matched into principal-agent pairs. The game is played for 10 periods where in each one a new pair is created so that no agent interacts more than once with the same principal. In the first stage of the game the principal chooses the type of contract and whether to invest in a verification technology; he proceeds to make an offer to the agent he is matched with. In the second stage the agent decides whether to accept the contract and, if he does, how much effort to expend. In the third and final stage, if the principal has invested in the verification technology, evidence of the agent's effort level is obtained with probability p and the terms of the contract are enforced.

The three types of contract the principal has at his disposal to offer are:

- Incentive contract: The contract specifies the wage w, the contracted effort level e_c, and the maximum fine $\overline{f}$ to be imposed on the agent if evidence of shirking is discovered. It is assumed the principal has invested in the verification technology.

- Trust contract: The contract specifies the wage w and the contracted effort level e_c but there is no investment in the verification technology. Therefore, e_c is non-verifiable and w is not contingent on the actual effort of the agent.
- Bonus contract: This contract is similar to the trust contract except where if the agent expends an effort level superior to e_c the principal might choose to reward the agent by offering a bonus b, which is not enforceable.

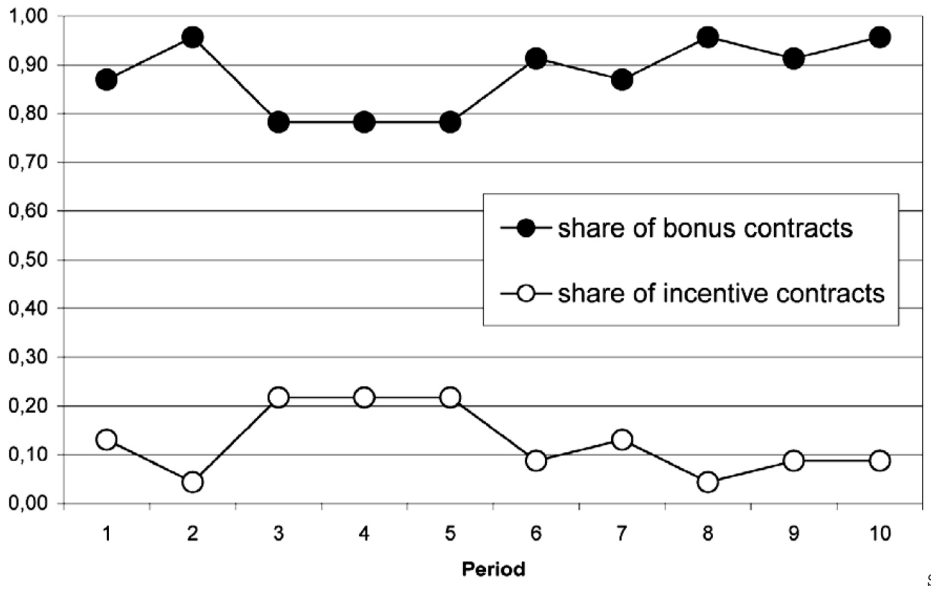
An analysis of what type of contract the principal would find preferable to offer under the assumption of self-regarding agents and principals is similar to the one made for the gift exchange game. Given the non-enforceability of e_c in both the trust and bonus contracts, and the added non-enforceability of b in the bonus contract, it is trivial to conclude that self-regarding agents will exert the lowest possible effort level.

However, the principal is able to induce a positive effort level from the risk-neutral agent if the verification technology is potent enough so that $\overline{p}f \geq c(e^*) - c(\underline{e})$. Therefore, under the assumption of self-regarding preferences, the principal will choose the incentive contract over the other two available alternatives: the trust and bonus contracts. This is a testable prediction.

Fehr, Klein, and Schmidt (2007) features two different experiments, one where principals can choose between a trust and an incentive contract and another where principals choose between a bonus and an incentive contract. In the trust-incentive experiment the self-regarding preferences are largely confirmed. Most principals choose to offer agents an incentive contract and their share, over the 10 rounds played, increases over time. While there is experimentation on the part of most principals by offering a trust contract at least once, the differences in payoffs from both contracts were such that principals preferred the incentive contract. Once principals learned how to create an appropriate incentive contract with the right mix of w, e_c, and f, they made up the large majority of contracts offered.

It is in the bonus-incentive experiment that deviations from the self-regarding prediction are observed. Recall that the bonus b in the bonus contract is non-enforceable. If the principal, in his self-regarding rationale, decides not to offer a bonus, then the bonus contract becomes a trust one. Since agents are aware that the bonus is non-binding and therefore not likely to be realized, they should equate the bonus contract with the trust one and act accordingly. This would lead us to predict that in the bonus-incentive experiment we should observe results similar to those from the trust-incentive one. This prediction is, however, not confirmed by the experimental results.

Figure 6: Share of bonus and incentive contracts in the bonus-incentive experiment



Source: Fehr, Klein, and Schmidt (2007).

The overwhelming majority of contracts in the bonus-incentive experiment are bonus contracts, with the incentive contract seldom being offered. This choice is driven by the ability of bonus contracts to elicit a higher effort level from the agents which increases the payoff the principals get in comparison with the possible payoffs from the other contracts. Part of this larger surplus is then allocated by the principals as a bonus to the agents. The average income gained by agents in the bonus contract is approximately 23% higher than the one earned in the incentive contract. It turns out therefore that that the use of a bonus contract is beneficial to both parties.

Since the self-regarding model does not explain this set of choices, we will suppose an alternative where both the agent and the principal have other-regarding preferences of the Fehr-Schmidt form. This analysis follows closely Dhami (2016).

For simplicity we will assume that the agent's output, $v(e)$, is equal to e while the effort cost function is $c(e) = \frac{1}{2} e^2$, where $e \in [0,1]$. Under these specifications, the utility of a self--regarding agent and principal is given by, respectively

$$\begin{cases} u = w - \frac{1}{2} e^2 - C_A \\ \pi = e - w - C_p \end{cases}$$

where C_A and C_p are whatever other individual costs the agent and principal, respectively, incur by taking part in the relationship, such as the cost k of the verification technology for the principal. The respective FS preferences are given by

$$U(u, \pi) = u - \alpha_A \max\{\pi - u, 0\} - \beta_A \max\{u - \pi, 0\}$$

$$\Pi(u, \pi) = \pi - \alpha_p \max\{u - \pi, 0\} - \beta_p \max\{\pi - u, 0\}$$

We make the additional assumption that α and β for both the agent and the principal are higher than 0.5, the reasoning being that under this assumption both parties will behave in ways such that their monetary payoffs are equalized. An indication of why this is so is provided in the Appendix. The parameter estimates gathered in Eckel and Gintis (2010) show more support for the assumption that $\alpha > 0.5$ than for $\beta > 0.5$, but the assumption affords us simplicity. We further assume that the value of the outside option of the agent is zero.

To show why the principal chooses to offer a bonus contract over the incentive contract we need to demonstrate that the former dominates the latter. We start by describing the expected profits for the principal under the incentive contract. The incentive compatibility constraint of the self-regarding agent is[6]

$$(1 - p)\, w + p\,(w - f) \leq w - \frac{1}{2} e_c^2,$$

where we assume $C_A = 0$. This gives us $e_c \leq \sqrt{2pf}$ as the set of effort levels that are incentive compatible. To maximize profits, the self-interested principal sets a contract (w, e_c) that maximizes expected profits

$$E(\pi) = (1 - p)(e - w - k) + p(e - w - k + df) = e - w - k + pdf,$$

where d is a binary variable dependent on whether $e < e_c$. If e_c satisfies the constraint that $e_c \leq \sqrt{2pf}$ and the constraint that $w - \frac{1}{2} e^2$ then $d = 0$. In light of these constraints we have that in an incentive contract in which the principal intends to maximize his profits the optimal contracted effort level is

$$e_I = min\{1, \sqrt{2pf}\}.$$

Recall that $e \in [0,1]$. Equation tells us the principal will choose to contract the minimum effort level that is also incentive compatible for the agent, which will depend on whether $\sqrt{2pf}$ is higher or lower than 1.

[6] Under the incentive contract, both agent and principal act as self-regarding given that there is no opportunity for one to exhibit reciprocity towards the other.

- If $\sqrt{2pf} \geq 1$ then we have $e_I = 1$ and $w = \frac{1}{2}$. In this case the expected profit of the firm is given by $E(\pi) = \frac{1}{2} - k$.
- If $\sqrt{2pf} < 1$ then $e_I = \sqrt{2pf}$. Therefore $w = pf$ and $E(\pi) = \sqrt{2pf} - pf - k$.

We now have the expected profits of the incentive contract which we can compare with the expected profit from a bonus contract offered by an other-regarding principal. The bonus contract is as described previously. In Stage 3, given that the experimental evidence suggests that $e > e_c$, the principal awards a bonus b. Because he has other-regarding preferences, this bonus will be such that the payoffs of both parties are equaled.

Thus, we have

$$e - w - b = w + b - \frac{1}{2} e^2,$$

which we solve for b to get

$$b = \frac{1}{2} e - w + \frac{1}{4} e^2 = b(e, w).$$

The other-regarding agent will chose an optimal effort choice such that his monetary payoff is equal to that of the principal, that is,

$$w + b(e, w) - \frac{1}{2} e^2 = e - w - b(e, w).$$

Because the bonus is chosen so that the payoffs are equal, then the previous equation is satisfied for any value of e. Taking the first derivative of in order to e gets us the result that the payoff is maximized at $e_F^b = 1$.

The other-regarding principal's expected payoff is $E(\pi_B) = e - w - b(e, w)$, which when $e = 1$ yields

$$\begin{aligned} E(\pi_B) &= e - w - (\frac{1}{2} e - w + \frac{1}{4} e^2) \\ &= 1 - w - \frac{1}{2} + w - \frac{1}{4} \\ &= \frac{1}{4} \end{aligned}$$

So given these options, which contract should the principal prefer? Under the self-regarding assumption we would expect the incentive contract to dominate over all others. However, taking into account that principals and agents might have other-regarding preferences, we conclude that the answer depends on a number of parameters.

Suppose that $\sqrt{2pf} = 1$, in which case $E(\pi_I) = \frac{1}{2} - k$ as shown earlier. For the incentive contract to dominate over the bonus contract it would be needed that $E(\pi_I) = E(\pi_B)$, that is, $\frac{1}{2} - k > \frac{1}{4}$. This is only true if $k < \frac{1}{4}$. For the case where $\sqrt{2pf} < 1$ we have that $E(\pi_I) = \sqrt{2pf} - pf - k$, which means that the incentive contract dominates over the bonus contract only if $\sqrt{2pf} - pf - k > \frac{1}{4}$.

What these two situations show is that rather than the incentive contract always dominating over the bonus contract, it instead only does so when the deterrence parameters are high enough so that the principal has reliable access to evidence of shirking and the monitoring technology isn't too costly. If this isn't the case, because both parties have other-regarding preferences and go above and beyond their self-regarding behavior, the bonus contract engenders a relationship that is more beneficial to both the agent and the principal than the one created by the incentive contract.

Fehr and Schmidt (2007) extend the results from Fehr, Klein, and Schmidt (2007) by combining the fine from the incentive contract with bonus contract, creating a contract that features both the carrot and the stick. It was thought that given the combination of both incentives that this new contract would dominate over the others but that was not the case as more than two thirds of all contract offers were bonus contracts. The authors advance two possible explanations for their results. One is that the introduction of a fine might be seen by the agent as being in bad faith, leading them to reciprocate by expending a lower effort level. They also offer the hypothesis that since agents do not know the principal's trustworthiness they infer from the introduction of the fine that the principal will make a lower bonus offer. Indeed, from the experimental data, principals who offer the combined contract do offer significantly lower bonus payments.

This illustrative example should be interpreted not as proving that every relationship between worker and firm will be such that other-regarding preferences are a strong determinant of the choice between competing types of contracts. Instead, the attempt has been to suggest that to the degree that other-regarding preferences are an important determinant of that choice, assuming self-regarding preferences will severely limit our ability to model and understand such a relationship.

6. Conclusion

It is our purpose as social scientists to venture farther into what we lack knowledge of and map out the intricacies that make up human behavior. We must observe the world around us, tease out its regularities, build up theories to explain them and test them against new observations. It is in the testing of those theories and the failure to explain behavior that we know our job is far from being over.

It has been argued throughout this study that sufficient experimental evidence has been accumulated to make us more doubtful about the assumption of agents possessing self--regarding preferences as sufficient to explain the full set of human behavior. This failure to explain documented behavior has motivated the introduction of the concept of other--regarding preferences, where agents are said to not only be preoccupied with themselves but also with those around them. We have introduced a new model that takes into account other-regarding preferences and argued that this type of preferences allows us to explain what was previously an unexplainable behavior.

It is worth pointing out that this by no way means the self-regarding assumption is incorrect. That it has been continually used with success for many years shows that even though it is not a full description of how agents behave, it still is a useful modelling assumption of great explicability. Indeed, an issue that was not dealt with in this study is how to mediate between the two assumptions. Under what circumstances is one well served by the self--regarding assumption and under which should we introduce other-regarding preferences? The literature has thus far scantly addressed this issue and some guiding principles will need to emerge before more economists use these new preferences productively.

Regarding the experimental evidence used throughout this study, it is worth noting recent developments on the topic of replicability. Poor experimental procedure, ineffectual use of statistical tools, along with unwarranted confidence put on the results from the combination of the previous two being true, has led to the accumulation of false or irrelevant results. Ioannidis (2005) provides a good introduction to this problem. Ioannidis, Stanley, and Doucouliagos (2017) deals directly with Economics, where it is concluded that "nearly 80% of the reported effects [in the empirical economics literature surveyed in the paper] are exaggerated." Given the reliance in many of the experiments surveyed in this study on small sample sizes, and the resulting low statistical power, it would not be surprising if some of the conclusions they reach are not correct. While it was argued that laboratory evidence has a role in the study of economic behavior, this is only so if that evidence is properly gathered.

It is our hope that this study motivates the use of other-regarding preferences in the examination of economic behavior which either has not been sufficiently examined, or has only been so through the use of the self-regarding model. Economics would only gain by expanding the lens through which it studies behavior.

Appendix

Contract Design under Other-Regarding Preferences

Consider the agent's problem of choosing an effort level. We intend to show that an other-regarding agent with $\alpha_A \geq \beta_A > 0.5$ will choose e such that his monetary payoff is equal to the principal's.

- $x_A > x_p$.

If the agent's monetary payoff is higher than the principal's then his utility is $U(x_A, x_p) = x_A - \beta_A(x_A - x_p)$. Suppose the agent transfers $\epsilon > 0$, where ϵ is an infinitesimal, to the principal. Then,

$$\begin{aligned} U(x_A, x_p) &= x_A - \epsilon - \beta[(x_A - \epsilon) - (x_p + \epsilon)] \\ &= x_A - \beta\,(x_A - x_p) + \epsilon\,(2\beta_A - 1). \end{aligned}$$

which implies a positive change since $\beta_A > 0.5$. The agent is therefore made better off by transferring resources to the principal.

- $x_p > x_A$.

Now consider the case where the agent's monetary payoff his lower than the principal's. The agent can punish the principal and reduce his payoff by a unit at cost $\Upsilon < 1$. The agent's utility if he does so is

$$\begin{aligned} U(x_A, x_p) &= (x_A - \Upsilon) - \alpha_A\,[(x_p - 1) - (x_A - \Upsilon)] \\ &= x_A - \alpha_A\,(x_p - x_A) + \alpha_A\,(1 - \Upsilon) - \Upsilon. \end{aligned}$$

For the change to be positive, we need that $\alpha_A(1 - \Upsilon) - \Upsilon \geq 0$. This means that $\Upsilon \leq \frac{\alpha_A}{1 + \alpha_A} < 1$.

A similar argument can be made for how the principal reacts to inequity, leading us to conclude that if both have $\alpha \geq \beta > 0.5$ they will work towards equaling their monetary payoffs.

References

Akerlof, G. A. (1982) Labor contracts as partial gift exchange, Quarterly Journal of Economics, 97 (4), 543-569.

Andersen, S.; Ertaç, S.; Gneezy, U.; Hoffman, M.; List, J. A. (2011) Stakes matter in ultimatum games, American Economic Review, 101 (7), 3427-3439.

Andreoni, J. (1995) Cooperation in public-goods experiments: Kindness or confusion? American Economic Review, 85 (4), 891-904.

Barmettler, F.; Fehr, E.; Zehnder, C. (2012) Big experimenter is watching you! Anonymity and prosocial behavior in the laboratory, Games and Economic Behavior, 75 (1), 17-34.

Camerer, C. (2011a) Behavioral Game Theory: Experiments in Strategic Interaction, Princeton, Princeton University Press.

Camerer, C. (2011b) The promise and success of lab-field generalizability in experimental economics: A critical reply to Levitt and List, in G. R. Fréchette and A. Schotter (eds.), Handbook of Experimental Economic Methodology, Oxford, Oxford University Press.

Camerer, C., Hogarth, R. M.; Budescu, D. V.; Eckel, C. (1999) The effects of financial incentives in experiments: A review and capital-labor-production framework, Journal of risk and uncertainty, 19 (1-3), 7-42.

Cameron, L. A. (1999) Raising the stakes in the ultimatum game: Experimental evidence from Indonesia, Economic Inquiry, 37 (1), 47-59.

Carter, J. R.; Irons, M. D. (1991) Are economists different, and if so, why? Journal of Economic Perspectives, 5 (2), 171-177.

Cohn, A.; Fehr, E.; Goette, L. (2008) Fairness and effort-evidence from a field experiment, IFW Discussion Paper No. 20, University of Zurich.

Dana, J.; Cain, D. M.; Dawes, R. M. (2006) What you don't know won't hurt me: Costly (but quiet) exit in dictator games, Organizational Behavior and Human Decision Processes, 100 (2), 193-201.

Dana, J.; Weber, R. A.; Kuang, J. X. (2007) Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness, Economic Theory, 33 (1), 67-80.

Dhami, S. (2016) The Foundations of Behavioral Economic Analysis, Oxford, Oxford University Press.

Ebbinghaus, H. (1885) Memory: A Contribution to Experimental Psychology, New York, Dover.

Eckel, C.; Gintis, H. (2010) Blaming the messenger: Notes on the current state of experimental economics, Journal of Economic Behavior and Organization, 73 (1), 109-119.

Edgeworth, F. Y. (1881) Mathematical Psychics: An essay on the Application of Mathematics to the Moral sciences, London, Kegan Paul.

Falk, A.; Heckman, J. J. (2009) Lab experiments are a major source of knowledge in the social sciences, Science, 326 (5952), 535-538.

Falk, A.; Meier, S.; Zehnder, C. (2013) Do lab experiments misrepresent social preferences? The case of self-selected student samples, Journal of the European Economic Association, 11 (4), 839-852.

Fehr, E.; Falk, A. (2002) Psychological foundations of incentives, European Economic Review, 46 (4--5), 687–724.

Fehr, E.; Gächter, S.; Kirchsteiger, G. (1997) Reciprocity as a contract enforcement device: Experimental evidence, Econometrica, 65 (4), 833-860.

Fehr, E.; Goette, L.; Zehnder, C. (2009) A Behavioral Account of the Labor Market: The Role of Fairness Concerns, Annual Review of Economics, 1 (1), 355-384.

Fehr, E.; Kirchsteiger, G.; Riedl, A. (1993) Does fairness prevent market clearing? An experimental investigation, Quarterly Journal of Economics, 108 (2), 437-459.

Fehr, E.; Klein, A.; Schmidt, K. M. (2007) Fairness and contract design, Econometrica, 75 (1), 121--154.

Fehr, E.; List, J. A. (2004) The hidden costs and returns of incentives: Trust and trustworthiness among CEOs, Journal of the European Economic Association, 2 (5), 743-771.

Fehr, E.; Schmidt, K. (1999) A Theory of Fairness, Competition and Cooperation. Quarterly journal of Economics, 114 (3), 817-868.

Fehr, E.; Schmidt, K. M. (2007) Adding a stick to the carrot? The interaction of bonuses and fines, American Economic Review, 97 (2), 177-181.

Gneezy, U.; List, J. A. (2006) Putting behavioral economics to work: resting for gift exchange in labor markets using field experiments, Econometrica, 74 (5), 365-384.

Guala, F. (2008) History of experimental economics, in The New Palgrave Dictionary of Economics, London, Palgrave Macmillan, 1-8.

Hausman, D. M. (2011). Preference, Value, Choice, and Welfare, Cambridge, Cambridge University Press.

Henrich, J. (2000). Does culture matter in economic behavior? Ultimatum Game Bargaining among the Machiguenga of the Peruvian Amazon, American Economic Review, 90 (4), 973-979.

Henrich, J.; Boyd, R.; Bowles, S.; Camerer, C.; Fehr, E.; Gintis, H.; McElreath, R.; Alvard, M.; Barr, A.; Ensminger, J.; Henrich, N. S. (2005), "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies, Behavioral and Brain Sciences, 28 (6), 795-815.

Homans, G. C. (1954). The cash posters: A study of a group of working girls, American Sociological Review, 19 (6), 724-733.

Ioannidis, J.; Stanley, T. D.; Doucouliagos, H. (2017) The power of bias in economics research, Economic Journal, 127, F236-265.

Ioannidis, J. P. A. (2005) Why most published research findings are false, PloS Medicine, 2 (8), e124.

Johnson, N. D.; Mislin, A. A. (2011) Trust games: A meta-analysis, Journal of Economic Psychology, 32 (5), 865-889.

Levitt, S. D.; List, J. A. (2007) What do laboratory experiments measuring social preferences reveal about the real world? Journal of Economic Perspectives, 21 (2), 153-174.

List, J. A. (2004) Young, selfish and male: Field evidence of social preferences Economic Journal, 114, 121-149.

List, J. A. (2006) The behavioralist meets the market: Measuring social preferences and reputation effects in actual transactions, Journal of Political Economy, 114 (1), 1-37.

List, J. A. (2007) On the interpretation of giving in dictator games, Journal of Political economy, 115 (3), 482-493.

Mearsheimer, J. J.; Walt, S. M. (2013) Leaving theory behind: Why simplistic hypothesis testing is bad for international relations, European Journal of International Relations, 19 (3), 427-457.

Moscati, I. (2007) Early experiments in consumer demand theory: 1930-1970. History of political economy, 39 (3), 359-401.

Roth, A. E.; Prasnikar, V.; Okuno-Fujiwara, M.; Zamir, S. (1991) Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An experimental study, American Economic Review, 81 (5), 1068-1095.

Samuelson, P.; Nordhaus, W. D. (1985) Economics, New York, McGraw-Hill.

Samuelson, P.; Nordhaus, W. D. (1992) Economics, New York, McGraw-Hill.

Slonim, R.; Roth, A. E. (1998) Learning in high stakes ultimatum games: An experiment in the Slovak republic, Econometrica, 66 (3), 569-596.

Smith, A. (1822). The Theory of Moral Sentiments (Vol. 1), London, J. Richardson.

Ward, A. (1997) Naive realism in everyday life: Implications for social conflict and misunderstanding, in E. Reed, E. Turiel, and T. Brown (eds.), Values and Knowledge, Mahwah, Lawrence Erlbaum, 103-135.

António Afonso
ISEG - UL, Universidade de Lisboa
REM - Research in Economics and Mathematics
UECE - Research Unit on Complexity and Economics
aafonso@iseg.utl.pt
André Albuquerque
ISEG - UL, Universidade de Lisboa
andre.malbuq@gmail.com

DOI: https://doi.org/10.14195/2183-203X_46_3

Sovereign Credit Rating Mismatches

António Afonso
André Albuquerque

Received for publication: March 9, 2018
Revision accepted for publication: April 16, 2018

ABSTRACT
We study the factors behind ratings mismatches in sovereign credit ratings from different agencies, for the period 1980-2015. Using random effects ordered and simple probit approaches, we find that structural balances and the existence of a default in the last ten years were the least significant variables. In addition, the level of net debt, budget balances, GDP per capita and the existence of a default in the last five years were found to be the most relevant variables for rating mismatches across agencies. For speculative-grade ratings, a default in the last two or five years decreases the rating difference between S&P and Fitch. For the positive rating difference between S&P and Moody's, and for investment-grade ratings, an increase in external debt leads to a smaller rating gap between the two agencies.
Keywords: Sovereign ratings; split ratings; panel data; random effects ordered probit.

JEL Classification: C23; C25; E44; F34; G15; H63.

Acknowledgements: We thank an anonymous referee and participants at the 49th Money, Macro and Finance Research Group Annual Conference, London (2017) for very useful comments. The opinions expressed herein are those of the authors and do not necessarily reflect those of their employers. Any remaining errors are the authors' sole responsibility. UECE is supported by FCT – Fundação para a Ciência e a Tecnologia, Portugal.

1. Introduction

Credit rating agencies play a crucial role in reducing information asymmetries in the financial markets and provide a fundamental input to the financial institutions risk assessment required by regulators. In fact, capital requirements are calculated notably by applying to the institution financial assets a weighting factor depending on the associated credit rating. Sovereign credit ratings summarise in an ordinal qualitative scale a complex and thorough analysis of the ability a country has to service its debt. Since institutional investors nowadays are only allowed to acquire financial assets above a certain rating, countries willing to issue debt are in practice obliged to pay for a credit rating.

With the globalization of financial markets and the proliferation of credit ratings, rating agencies assigning different credit ratings to the same country became more frequent. Our contribution is twofold: first, we set up the possible pairs of rating mismatches across the three main Nationally Recognized Statistical Rating Organizations (NRSROs) for 105 countries, highlighting persistent split ratings. Second, we analyse the rating differences between S&P, Moody's and Fitch in the light of a random-effects probit framework and using as explanatory variables a set of economic variables found in the literature as important determinants of sovereign ratings.

Our ordered probit results found, for every dataset used, that the level of net debt, budget balances, GDP per capita, and a default in the last five years contribute in more than 20% of the regressions to the overall rating differences. On the other hand, the structural balance did not significantly contribute to the rating differences here considered. The structural balance and the default in the last ten years were the least significant across all our regressions, In addition, for speculative-grade ratings, we find that a default in the last two or five years decreases the rating difference between S&P and Fitch. For the positive rating difference between S&P and Moody's for investment-grade ratings, an increase in external debt leads to a smaller rating gap between the two agencies.

From a policy perspective, the economic implications of our results imply that sovereigns and fiscal policy makers might learn which determinants matter most for each rating agency, allowing for a better ex-ante fine-tuning of the rating process.

The remainder of the paper is organized as follows: section two provides the literature review; section three explains the methodology; section four discusses the results of the analysis; and section five is a conclusion.

2. Literature Review

Amstad and Packer (2015) define sovereign ratings as "opinions about the creditworthiness of sovereign borrowers that indicate the relative likelihood of default on their outstanding debt obligations". These ratings, like the ratings about other types of credit, try to assess both the ability and willingness of the borrower to pay. To accomplish this, qualitative factors, like institutional strength and the rule of law, and quantitative factors, like measures of fiscal and economic strength, the monetary regime, foreign exchange reserves, are analysed to rate a sovereign issuer. Kiff et al. (2012) state that ratings are not only about credit risk

but also convey information about credit stability (changes in credit risk), and the assessments represented by ratings are medium-term outlooks that should not change due to the impact of cyclical components. Rating agencies minimize rating volatility by assessing through the cycle: a rating should be changed only to reflect a shift in fundamental factors (and consequently a change in basic creditworthiness), and not as a response to a recession or a global liquidity shortage, for example.

Bhatia (2002) affirms that the widespread use by investors of the credit ratings attributed by Standard & Poor's (S&P), Moody's Investors Service (Moody's) and Fitch Ratings (Fitch) reflects their utility for the market. This usefulness results from the simplicity and comparability of the rating systems used by those rating agencies, condensing detailed analysis into brief indicators, and from the "perceived analytical strength and independence of the agencies themselves."

A sovereign credit rating normally serves as the "ceiling" of the ratings within its territory, since the sovereign bond yields are considered riskless and therefore used as a benchmark against which returns on domestic investments are compared. In parallel, each sovereign creditworthiness is compared with the most trustworthy issuers (rated with an 'AAA' rating), and among those is the German government, whose bonds are regarded as one of the global risk-free benchmarks. Given the increasing integration of the capital markets, the growing issuance of bonded debt and the regulatory role of sovereign ratings on investors risk management, changes in sovereign ratings can have profound implications.

Both the Asian crisis in 1997 and the global financial crisis of 2007-08 highlighted flaws in the rating systems. In the first case, a rating approach based only on macroeconomic fundamentals was the culprit, revealing the importance of contingent liabilities and the international liquidity position of the issuers (Bhatia, 2002). In the latter case, and according to Brunnermeier (2009), one of the deciding factors contributing to the latest financial crisis was the fact that structured debt products (collateralized debt obligations (CDO)), had always a tranche reaching the 'AAA' rating, even if the underlying default risk was not equivalent to the default risk associated with a 'AAA' bond rating. Fund managers were attracted to buying these structured products offering seemingly high expected returns with an acceptable level of risk, and when the quality of the securitized assets deteriorated (signalled by a spike in the default rate of the so-called subprime mortgages), every holder inevitably faced losses and eventually had to write-down a significant part of their mortgage-related securities.

In the wake of the global financial crisis and the European sovereign debt crisis, Amstad and Packer (2015) highlight the changes in the sovereign risk methodologies used by the major rating agencies. These rating methodologies explain which factors drive the evaluation of the likelihood of default. A common principle to these revisions is that agencies tried to adopt assessment systems more reliant on quantitative inputs, to make ratings more transparent and replicable.

For instance, the Moody's rating methodology (Moody's Investors Service, 2015) explains its sovereign credit risk assessment on the "interplay" of four key factors: economic strength, institutional strength, fiscal strength and susceptibility to event risk. In addition, each factor usually encompasses one or more indicator, like the average real GDP growth and volatility, nominal GDP, GDP per capita, inflation level and volatility, etc.

Al-Sakka and ap Gwilym (2010) associate the growing importance of credit rating agencies to the increasing number of issuers and debt products, and globalization, but also to the requirements applied to financial institutions and banks. The first ones are only allowed to trade debt securities rated by NRSRO, whereas the latter, stemming from the Basel II Accord, usually employ external credit ratings to assess their credit risks and to determine capital adequacy requirements.

The determinants of sovereign credit ratings are an object of study since the seminal work of Cantor and Packer (1996), a cross sectional OLS estimation which identified per capita income, GDP growth, inflation, external debt, level of economic development and default history as important determinants of sovereign ratings assigned by Moody's and S&P. This methodology was also used by Afonso (2003), which also included a logistic and an exponential transformation of the ratings, in addition to the linear transformation already used. Mulder and Monfort (2000) and Eliasson (2002) generalized the OLS approach to panel data, both using a linear transformation of the ratings.

On the other hand, and to overcome the limitation of OLS regressions with a linear transformation of the ratings, Bissoondoyal-Bheenick (2005) used an ordered probit model for a period of five years and 95 countries.[1]

Afonso et al. (2008) analysed the determinants of sovereign ratings from the three main agencies by using a linear regression framework (random effects estimation, pooled OLS estimation and fixed effects estimation) versus an ordered probit response framework.[2] In addition, Afonso et al. (2011) confirm that logistic and exponential transformations to ratings provide little improvement over the linear transformation, not finding evidence of the so-called "cliff effects" (when investors adjust their portfolio composition to select only investment grade securities). This work also highlights the difference between short- and long-term determinants, concluding that GDP per capita, GDP growth, government debt and budget balance have a short-term impact, whereas government effectiveness, external debt, foreign reserves and default history influence ratings in the long-run.

In addition, Amstad and Packer (2015) used several explanatory variables as proxies for fiscal, economic and institutional strength, monetary regime, external position and default history and concludes that a small set of factors can largely explain the rating scale. Finally, Vu et al. (2017) report that political risk can contribute to explain rating mismatches in a country sample during the period 1997-2011.

3. Methodology

To understand which factors may explain split sovereign ratings and if some of those factors are considered more relevant by certain agencies, we propose to analyse the collected dataset using a random-effects ordered probit model regression framework.

[1] An OLS regression with a linear transformation of the ratings assumes a constant distance between adjacent rating notches. However, ratings represent a qualitative ordinal assessment of a sovereign credit risk, thus the distance between two adjacent ratings may not be the same

[2] Instead of assuming a rigid shape of the ratings scale, this model estimates the threshold values between rating notches, defining the shape of the ratings curve.

The source of the information used to create the dependent variables were the rating changes for long-term sovereign foreign currency ratings obtained from Bloomberg for the three main credit rating agencies (Standard & Poor's, Moody's, and Fitch Ratings). For each country and for each year, we selected the last rating change of the year as that country's year rating. In addition, we filled the years without any rating change by extending the rating of the previous year and rating withdrawals by the rating agencies were ignored, since the rating given before the withdrawal keeps its relevance for the markets.

The qualitative rating given by the rating agencies were then converted into a numerical scale, from 0 to 21, where 21 corresponded to the 'AAA' from S&P and Fitch and 'Aaa' from Moody's and 0 corresponded to a (selective) default.

Our six dependent variables – $Diff\_UP_{it}^{SF}$, $Diff\_DW_{it}^{SF}$, $Diff\_UP_{it}^{MF}$, $Diff\_DW_{it}^{MF}$, $Diff\_UP_{it}^{SM}$ and $Diff\_DW_{it}^{SM}$ – represent the difference in ratings between the credit rating agencies considered in this work, as follows:

$Diff\_UP_{it}^{SF}$ – difference between the ratings given by S&P and Fitch, when S&P rating was higher or equal than Fitch's rating;
$Diff\_DW_{it}^{SF}$ – difference between the ratings given by S&P and Fitch, when S&P rating was lower or equal than Fitch's rating;
$Diff\_UP_{it}^{MF}$ – difference between the ratings given by Moody's and Fitch, when Moody's rating was higher or equal than Fitch's rating;
$Diff\_DW_{it}^{MF}$ – difference between the ratings given by Moody's and Fitch, when Moody's rating was lower or equal than Fitch's rating;
$Diff\_UP_{it}^{SM}$ – difference between the ratings given by S&P and Moody's, when S&P rating was higher or equal than Moody's rating;
$Diff\_DW_{it}^{SM}$ – difference between the ratings given by S&P and Moody's, when S&P rating was lower or equal than Moody's rating.

As an example, let R_{it}^{X} represent the rating from credit rating agency X for the country i in year t and consider the dependent variable $Diff\_UP_{it}^{SM}$, representing the difference between S&P and Moody's ratings: $Diff\_UP_{it}^{SM} = R_{it}^{S} - R_{it}^{M}$, when $R_{it}^{S} \geq R_{it}^{M}$. If $Diff\_UP_{it}^{SM} > 0$, then S&P considers country i, in time t, more capable of fulfilling its debt obligations than what is assessed by Moody's.

3.1. Explanatory Variables

In this paper we selected the explanatory variables according to the existing literature on the determinants of sovereign ratings (see Cantor and Packer, 1996, Afonso, 2003, and, for ordered response models, Afonso et al., 2008, and Afonso et al., 2011). Accordingly, the predictors that had better explanatory power for the rating scaled are the level of GDP per capita, real GDP growth, external debt, government debt and the government budget balance.

In addition to these predictors[3], this study also considered as explanatory variables the government structural balance, inflation and the default history of a country. The list of explanatory variables used in this work (the Appendix describes the data) is as follows:

Budget balance. Successive budget deficits may signal problems with the implemented policies;
Structural balance. Changes in the non-cyclical, or structural component, may be indicative of discretionary policy adjustments;
Gross debt. Summation of all liabilities that will require payments of interest and/or principal by the government, might signal rating deterioration;
Net debt. Net debt is calculated as gross debt minus the financial assets a government holds;
GDP growth rate. GDP per capita. A higher value strengthens the government ability to pay its debt;
Inflation. It helps governments by reducing the real stock of outstanding debt in domestic currency, but a consistent high value is associated with macroeconomic imbalances;
External debt. In addition, called foreign debt, represents the total debt a country (its government, corporations and citizens) owes to foreign creditors. It does not include contingent liabilities;
Four dummy variables for a default within the last year, last 2 years, last 5 years, and last 10 years. The definition of default by Beers and Mavalwalla (2016) here used is consistent with the literature on sovereign defaults. In fact, one considers that "a default has occurred when debt service is not paid on the due date, payments are not made within the time frame specified under a guarantee or, absent an outright payment default, creditors face material economic losses on the sovereign debt they hold."

3.2. Ordered Probit Regression Framework

We use a random effects ordered probit panel model, similar to what Afonso et al. (2011) used to identify the determinants of sovereign debt credit ratings and what Al-Sakka and ap Gwilym (2010) used to analyse the impact of split ratings on sovereign rating changes. According to Afonso et al. (2011), the ordered probit random-effects estimations consider the existence of an additional cross-country error term and therefore yield better results using panel data when compared with linear regression methods or fixed-effects probit estimations.

Our approach considers the discrete, ordinal nature of rating differences between credit rating agencies. The negative and positive rating differences for each pair of agencies was analysed separately, due to expected symmetrical reading if the dependent variable is positive

[3] Regarding government debt, we have analysed both gross and net government debt separately.

or negative by construction, comparable to what Al-Sakka and ap Gwilym (2010) expected with rating migrations.

Consider our ordered probit regression setting, when we are regressing $Diff\_UP_{it}^{SM}$ as the dependent variable. (In this case, all observations have the rating from S&P higher or equal than the rating from Moody's.) If the resulting coefficient of an explanatory variable, say, real GDP growth, is positive and significant, we conclude that an increase in real GDP growth will contribute to a bigger difference between S&P and Moody's ratings.[4] In a similar way, if the coefficient of the level of public debt is negative, we may conclude that an increase in the level of public debt, will contribute to a smaller difference between the ratings given by S&P and Moody's.[5] In practice, a positive coefficient has a symmetrical reading if it is related to a UP or a DW variable.

Our specification is defined as follows, and the value of our y_{it} dependent variable depends on whether we are considering the ordered probit or the simple probit approach:

$$y_{it} = \beta_1 \Delta GD_{it} + \beta_2 NGDP_{RPCH_{it}} + \beta_3 NGDPDPC_{it} + \beta_4 PCPIPCH_{it} + \beta_5 \Delta ED_{it} + \gamma Default\mathcal{Z}_{it} + \varepsilon_{it};\ \varepsilon_{it} \sim \mathcal{N}(0,1), \quad (1)$$

$$i = 1, \ldots C \text{ (countries)},\ t = 1, \ldots Y \text{ (years)},$$

where y_{it} is an ordinal variable equal to either $Diff\_UP_{it}^{AB}$ or $Diff\_DW_{it}^{AB}$.

In our ordered probit model, $Diff\_UP_{it}^{AB}$ ($Diff\_DW_{it}^{AB}$) = 1 or 2 if the rating from agency A is higher (lower) than the rating from agency B by one or more-than-one-notch, respectively, for sovereign i in year t, and 0 otherwise.

ΔGD_{it} may assume the variation value of the budget balance, gross debt, net debt or structural balance of country i in year t, depending on the chosen specification. $NGDP_{RPCH_{it}}$ – growth rate of GDP for country i in year t; $NGDPDPC_{it}$ – GDP per capita variation for country i in year t; $PCPIPCH_{it}$ – IPCH percentage change (inflation) for country i in year t; ΔED_{it} – external debt variation for country i in year t as percentage of GNI; and $Default\mathcal{Z}_{it}$ – dummy variable taking the value of 1 if country i in year t had defaulted in the last $\mathcal{Z}$ years, and 0 otherwise.

In the scope of the ordered probit framework, our six dependent variables were defined as to only having values of 1, 2 or 0, representing a rating gap of 1-notch, 2-or-more-notches or the inexistence of a rating gap, respectively. Equations 2 and 3 explain how the target variables were created:

$$Diff_{UP_{it}^{AB}} = \begin{cases} 1, \text{ if } |R_{it}^{\alpha} - R_{it}^{\beta}| = 1 \\ 2, \text{ if } |R_{it}^{\alpha} - R_{it}^{\beta}| \geq 2, \\ 0, \text{ otherwise} \end{cases} \quad \text{when } R_{it}^{\alpha} \geq R_{it}^{\beta}; \quad 2$$

[4] This could be interpreted as an increase in real GDP growth contributing to a higher S&P rating or a lower Moody's rating.

[5] In this case this could be interpreted as an increase in the level of public debt contributing to a lower S&P rating or a higher Moody's rating.

$$Diff_{DW_{it}^{AB}} = \begin{cases} 1, \textit{ if } |R_{it}^{\alpha} - R_{it}^{\beta}| = 1 \\ 2, \textit{ if } |R_{it}^{\alpha} - R_{it}^{\beta}| \geq 2, \\ 0, \textit{ otherwise} \end{cases} \quad \text{when } R_{it}^{\alpha} \leq R_{it}^{\beta}, \tag{3}$$

where A and B and α and $\beta \in \{SF, MF, SM\}$, and α (A) $\neq \beta$ (B).

Independently of the ordered or simple probit setup, when an observation has equivalent ratings from the considered rating agencies, the value of both $Diff\_UP_{it}^{AB}$ and $Diff\_DW_{it}^{AB}$ target variables is zero. Therefore, for each agency pair considered, both target variables use the same observations with no rating difference.[6]

Four different specifications of predicting variables were considered to overcome the correlation between some of the variables, using the four-abovementioned measures of fiscal developments: budget balance, structural balance, gross debt, and net debt. Within each specification, the four different default dummies were also combined.

4. **Empirical Analysis**

4.1. Data

Concerning the dependent variables, all the sovereign rating changes[7] were downloaded from Bloomberg and converted into the already mentioned numerical scale. Afterwards, we created six dependent variables, two variables for each rating agency pair, with the value of each variable reflecting the numerical rating difference between the ratings given by those specific agencies (comparable to what Livingston et al. (2008) did with the split rated issues).

The initial objective of this work was to study rating differences from 1970 onwards. However, and due to the inexistence of both macroeconomic values for many countries on those early years and ratings from at least two of the three selected agencies, our observations happened to comprehend only the period between 1980 and 2015. We only have observations with a rating from Fitch from 1994 onwards. Therefore, we have an unbalanced panel and by using first differences in the explanatory variables, one ensures stationarity. Naturally, the number of regressions reported varies according to the time span of the several variables and according to the existence of ratings for each pair of agencies for a specific country i in year t.

From 1990 and until 2000, we observe a bigger increase in the number or countries rated by at least two agencies, whereas from 2000 onwards the pace of this increase slowed, ending with 105 countries in our dataset with ratings from at least two of the main rating agencies.[8]

---

[6] An observation with $R_{it}^{\alpha} = R_{it}^{\beta}$ will make $Diff\_UP_{it}^{AB} = Diff\_DW_{it}^{AB} = 0$, so it has to be considered on the regressions of both target variables.

[7] We used the sovereign issuer ratings for foreign currency denominated debt.

[8] Countries in the sample: Angola, Albania, United Arab Emirates, Argentina, Armenia, Australia, Austria, Azerbaijan, Belgium, Bulgaria, Bahrain, Bosnia and Herzegovina, Belarus, Belize, Brazil, Barbados, Canada, Switzerland,

The distribution of the sovereign ratings on our dataset show that S&P is the agency which assigns more countries a rating of 'AA-' or above, and that the great majority of our observations are equal or above 'B-'. A higher degree of agreement on the top of the rating scale may explain the number of observations that had a rating of 'AAA' from all three agencies.

Our independent variables were obtained from datasets from the IMF (World Economic Outlook), World Bank (World Development Indicators), Bank of Canada (Database of Sovereign Defaults), and from the Quarterly External Debt Statistics dataset developed in collaboration between the World Bank and the IMF. Details on how those variables were created can be found in the Appendix.

4.2. Ordered Probit: Full Sample Analysis

We started by running the ordered probit regression with the full dataset. This dataset was composed by more than 850 observations for each dependent variable, comprised a period of at least 22 years (36 years only for the rating agency pair S&P and Moody's) and 69 or more countries. More than 65% of our observations for each of our target variables had no rating difference, whereas a rating difference of 1-notch was found at least in 19% of the observations. A rating difference of two or more notches can only be found 3.5%[9] of the times when analysing comparable ratings from S&P and Fitch; on the other hand, 9%[10] of the observations about the rating differences between S&P and Moody's have a 2-notch rating difference. This shows how S&P and Moody's disagree more when compared with the other rating agency pairs. Table 1 summarizes the full dataset.

Chile, China, Côte d'Ivoire, Cameroon, Colombia, Costa Rica, Czech Republic, Germany, Denmark, Dominican Republic, Ecuador, Egypt, Spain, Estonia, Ethiopia, Finland, Fiji, France, Gabon, United Kingdom of Great Britain and Northern Ireland, Georgia, Ghana, Greece, Guatemala, Honduras, Croatia, Hungary, Indonesia, India, Ireland, Iraq, Iceland, Israel, Italy, Jamaica, Jordan, Japan, Kazakhstan, Kenya, Korea (Republic of), Kuwait, Lebanon, Libya, Sri Lanka, Luxembourg, Morocco, Mexico, Mali, Mongolia, Mozambique, Malaysia, Namibia, Nigeria, Netherlands, Norway, New Zealand, Oman, Pakistan, Panama, Peru, Philippines, Papua New Guinea, Poland, Portugal, Paraguay, Qatar, Romania, Russian Federation, Rwanda, Saudi Arabia, Senegal, El Salvador, Serbia, Slovenia, Sweden, Seychelles, Thailand, Trinidad and Tobago, Tunisia, Turkey, Uganda, Ukraine, Uruguay, United States of America, Viet Nam, South Africa, Zambia.

[9] This value was obtained by calculating the average of the percentages of a rating difference of two or more notches between S&P and Fitch, when the first gave a higher rating than the latter ($Diff\_UP_{it}^{SF}$) and when the first gave a lower rating than the latter ($Diff\_DW_{it}^{SF}$).

[10] This value was obtained by calculating the average of the percentages of a rating difference of two or more notches between S&P and Moody's, when the first gave a higher rating than the latter ($Diff\_UP_{it}^{SM}$) and when the first gave a lower rating than the latter ($Diff\_DW_{it}^{SM}$).

Table 1: Summary of the full dataset, divided by the six target variables

	$Diff\_UP_{it}^{SF}$	$Diff\_DW_{it}^{SF}$	$Diff\_UP_{it}^{MF}$	$Diff\_DW_{it}^{MF}$	$Diff\_UP_{it}^{SM}$	$Diff\_DW_{it}^{SM}$
No. of countries	87	87	70	69	82	82
No. of years	22	22	22	22	36	36
First and last year	1994-2015	1994-2015	1994-2015	1994-2015	1980-2015	1980-2015
No. of observations	1149	1194	903	851	1103	1165
Observations with:						
Rating difference = 0	898 (78%)	898 (75%)	606 (67%)	606 (71%)	764 (69%)	764 (66%)
Rating difference = 1	221 (19%)	248 (21%)	223 (25%)	187 (22%)	247 (22%)	286 (25%)
Rating difference = 2	30 (3%)	48 (4%)	74 (8%)	58 (7%)	92 (8%)	115 (10%)
Observations with a value:						
GDP per capita	1149 (100%)	1194 (100%)	903 (100%)	851 (100%)	1103 (100%)	1165 (100%)
Real GDP growth rate	1148 (100%)	1194 (100%)	903 (100%)	851 (100%)	1103 (100%)	1165 (100%)
External debt	841 (73%)	897 (75%)	685 (76%)	648 (76%)	701 (64%)	808 (69%)
Gov. gross debt	1096 (95%)	1135 (95%)	865 (96%)	807 (95%)	1018 (92%)	1065 (91%)
Gov. net debt	1046 (91%)	1085 (91%)	822 (91%)	770 (90%)	954 (86%)	1004 (86%)
Budget balance	1112 (97%)	1153 (97%)	877 (97%)	824 (97%)	1057 (96%)	1104 (95%)
Structural balance	1064 (93%)	1100 (92%)	842 (93%)	774 (91%)	970 (88%)	1028 (88%)
Inflation	1147 (100%)	1191 (100%)	901 (100%)	848 (100%)	1100 (100%)	1160 (100%)
Default in the:						
Last year	312 (27%)	321 (27%)	164 (18%)	211 (25%)	268 (24%)	258 (22%)
Last two years	349 (30%)	363 (30%)	190 (21%)	247 (29%)	311 (28%)	297 (25%)
Last five years	419 (36%)	446 (37%)	248 (27%)	313 (37%)	379 (34%)	375 (32%)
Last ten years	522 (45%)	539 (45%)	331 (37%)	366 (43%)	448 (41%)	454 (39%)

Source: Rating agencies and own calculations.

Running the ordered probit regression for the full dataset, when the ratings from S&P are higher or equal to Fitch own ratings ($Diff\_UP_{it}^{SF}$ dependent variable), we get significant values for both budget balance and net debt variables. When budget balance increases, we expect the rating difference to decrease. For the net debt-predicting variable the opposite occurs: when its value increases, the rating difference increases as well (see Table 2).

Table 2: Summary of the regressions of the ordered probit full dataset

	Significant variables	Marginal Effect Rating difference = 1	Marginal Effect Rating difference = 2
$Diff\_UP_{it}^{SF}$	(-) Budget balance (4/4) (+) Net debt (4/4)	-0.001% 0.0004%	-0.00008% 0.00003%
$Diff\_DW_{it}^{SF}$	(-) GDP per capita (16/16) (-) External debt (16/16) (+) Default last 1Y (1/4) (+) Default last 2Y (1/4) (+) Default last 5Y (4/4)	-0.3% -0.1% 12.3% 11.5% 10.1%-10.5%	-0.03% -0.01% 1.9% 1.7% 1.3%-1.5%
$Diff\_UP_{it}^{MF}$	(-) GDP growth (9/16) (+) External debt (16/16)	-0.9%--1% 0.1%-0.2%	-0.2% 0.03%-0.04%
$Diff\_DW_{it}^{MF}$	(-) Gross debt (2/4) (+) Net debt (4/4) (+) Default last 2Y (3/4) (+) Default last 5Y (4/4)	-0.2% 0.0003% 10.8%-11.4% 11.3%-12.1%	-0.05%--0.06% 0.00007% 2.9%-3% 3%-3.2%
$Diff\_UP_{it}^{SM}$	(+) Default last Y (1/4) (+) Default last 2Y (4/4) (+) Default last 5Y (1/4) (+) Default last 10Y (1/4)	6.1% 8.1%-11.4% 12.9% 12.7%	2% 2.7%-3.5% 3.9% 3.6%
$Diff\_DW_{it}^{SM}$	(+) Budget balance (4/4) (+) Gross debt (4/4) (+) GDP growth (4/16) (-) GDP per capita (8/16)	0.005% 0.2% 0.8% -0.3%	0.002% 0.07% 0.2% -0.08%-0.09%

Notes: First parenthesis, coefficient signs; second parenthesis, number of significant regressions and total number of run regressions.

Regarding the $Diff\_DW_{it}^{SF}$ dependent variable (ratings from S&P being lower or equal to Fitch ratings), GDP per capita, external debt and the dummy default-in-the-last-5-years variables have statistically significant coefficients on all specifications. One can then conclude that if GDP per capita or external debt decrease the rating difference between those two rating agencies increases. The coefficients of the dummy default-in-the-last-5-years are also significant (and positive), showing that a default in the last five years increases the rating difference between S&P and Fitch in this case.

Analysing the rating difference between Moody's and Fitch, when the rating given by Moody's is higher than Fitch's rating ($Diff\_UP_{it}^{MF}$), we find significant values for two dependent variables, GDP growth (negative coefficient on two specifications) and external debt level (positive coefficients on all specifications). These results show that when GDP growth increases, the rating difference between these two agencies becomes smaller, whereas when the level of external debt increases, the gap between these two agencies increases.

When Moody's rating is lower than the rating from Fitch (Diff_DWitMF), we find that the dummy variable representing a default in the last five years has a positive coefficient in all specifications. For this reason, if a default in the last five years occurred, the rating difference in this setting between Moody's and Fitch increases as well.

The variables gross debt and net debt also have significant values of opposite signs: the gross debt contributes negatively for the rating difference, reducing the rating difference when its value increases, while the net debt has positive coefficients, so its increase is expected to positively influence the magnitude of the rating difference. We need to better understand the opposite signs of these two variables, since they should be correlated to a certain degree. The separate regressions of the investment and speculative ratings may shed some light into this topic.

The results from regressing our dependent variable Diff_UPitSM (when the S&P rating is higher than Moody's rating) display significant results only for the dummy default variables. The dummy default-in-the-last-2-years has positive coefficients on all specifications, meaning that if a country defaults in the last two years, the rating gap between S&P and Moody's will grow.

The results from regressing the last set of specifications, when the rating from S&P is lower than the rating from Moody's (Diff_UPitSM is the dependent variable), show that the budget balance, gross debt, GDP growth, and GDP per capita variables all contribute to the rating difference. Those first three variables have statistically significant and positive coefficients, meaning that when one of those variables increase, the rating difference between S&P and Moody's (Diff_UPitSM) will increase as well. The coefficient of the GDP per capita variable is negative, so when its value increases, the rating gap between S&P and Moody's becomes smaller. Overall, there does not seem to be a best proxy for government debt in the context of the empirical analysis.

4.3. Differentiation Between Investment and Speculative Ratings

We now report the ordered probit regression results when the observations were divided into two subsets, depending on the value of the average rating given by the rating agency pair. The observations with a numeric average rating of 12 or more (corresponding to 'BBB-' for S&P and Fitch or to 'Baa3' for Moody's) were grouped in the investment-grade subset, whereas those with a numeric rating less than 12 were grouped in the speculative--grade subset. In addition, the average rating is computed using the full sample since we want to divide the countries into "investment" and "speculative" categories. Therefore, some countries throughout the sample period may change from an "investment" category to a "speculative" category and vice-versa.

4.3.1. Investment-Grade Subset

When compared with the full dataset, the investment-grade dataset had observations for a smaller number of countries, between 49 and 57 different countries. The adopted criteria of considering only those observations with an investment-grade average rating reduced as expected the number of observations for each target variable (all target variables had less than 800 observations). It is important to note a higher percentage of observations with the same rating (when compared with the full dataset) from each rating agency in this setting, reflecting a greater coherence between the studied rating agencies when considering investment-grade sovereigns. This may be explained by Livingston et al. (2007) opaqueness idea that associates bond split ratings with the opaqueness of the issuer. In this case, investment-grade sovereign issuers disclose more detailed information, allowing rating agencies to better evaluate their ability to service debt and therefore rating agencies will agree more often about a country's rating in this context, leading to more observations with a rating difference of 0.

Our regression, when the S&P rating is higher than the rating from Fitch (Diff_UPitSF dependent variable), only yield significant results for one of the specifications (only one of the regressions show the budget balance variable as significant). This specification shows a positive correlation between government net debt and the observed rating difference, when the ratings from S&P and Fitch are investment-grade (see Table 3).

Table 3: Summary of the regressions of the ordered probit investment-grade subset

	Significant variables	Marginal Effect Rating difference = 1	Marginal Effect Rating difference = 2
$Diff\_UP_{it}^{SF}$	(-) Budget balance (1/4) (+) Net debt (4/4)	-0.0005% 0.0003%	-0.00002% 0.00001%
$Diff\_DW_{it}^{SF}$	(-) GDP per capita (15/16) (+) Default last 1Y (1/4)	-0.2% 12.6%	-0.005%--0.007% 0.52%
$Diff\_UP_{it}^{MF}$	(+) GDP per capita (12/16) (-) Inflation (16/16)	0.4% -2.0%--2.3%	0.05%-0.06% -0.2%--0.3%
$Diff\_DW_{it}^{MF}$	(-) Gross debt (1/4)	-0.1%	-0.02%
$Diff\_UP_{it}^{SM}$	(-) External debt (16/16)	-0.2%--0.3%	-0.04%
$Diff\_DW_{it}^{SM}$	(+) Budget balance (4/4) (+) Gross debt (4/4) (+) GDP growth (8/16) (-) Default last 1Y (4/4) (-) Default last 2Y (4/4)	0.004% 0.2% 1.1%-1.3% -10.9%--11.8% -8.4%--9.3%	0.0008% 0.04% 0.2%-0.3% -1.6%--1.8% -1.3%--1.5%

Note: See notes to Table 2.

When the rating from S&P is lower than the one from Fitch ($Diff\_DW_{it}^{SF}$), the obtained results for all specifications show a negative correlation between GDP per capita and the rating difference. This means that when GDP per capita increases, the rating difference is reduced. Only one of the regressions in this setting shows a significant and positive default dummy variable (in the last year).

The regressions of our dependent variable $Diff\_UP_{it}^{MF}$ (rating from Moody's higher than the one from Fitch, with the average classified as investment-grade) showed a positive and negative correlation between the rating difference and, respectively, GDP per capita and inflation. In this case, when GDP per capita increases, the rating gap increases, whereas with an inflation increase, the rating divergence between those two agencies will diminish.

While analysing the results when we regress the $Diff\_DW_{it}^{MF}$ (rating difference when the rating from Moody's is lower than the rating from Fitch), we only find one of the regressions showing a significant coefficient for the government gross debt predicting variable.

All the regressions of the $Diff\_UP_{it}^{SM}$ target variable (rating difference when the rating from S&P is higher than the rating from Moody's, and, on average, both ratings are investment-grade) show a significant negative correlation between external debt and the rating difference, leading to a smaller rating difference when the level of external debt rises.

The last dependent variable, $Diff\_DW_{it}^{SM}$, yield significant results when regressed against our predicting variables: both budget balance and government gross debt have significant positive coefficients,[11] meaning that an increase of those variables will lead to an increase in the rating difference between S&P and Moody's, when the rating of the first is lower than the rating of the latter.

The GDP growth-predicting variable also has significant positive coefficients on two of the four regressed specifications, showing an effect on the rating difference similar to the described effect of the budget balance and government gross debt on the rating gap. We also observe statistically significant and negative coefficients for two of the default dummy variables,[12] meaning that the existence of a default in the last year or two will contribute to a smaller rating difference between S&P and Moody's in this case.

4.3.2. Speculative-Grade Subset

Finally, the results from the ordered probit regression using the same specifications are analysed, this time using a subset of the full dataset composed only by observations with a speculative-grade average rating. This speculative-grade subset has observations for at least 38 countries[13] and comprises the period from 1992 to 2015. We have much less observations (between 238 and 435 observations) for the speculative-grade dataset when compared with the investment-grade and full datasets.

[11] With a significance level of 1% for all the relevant regressions.

[12] Default in the last year and in the last two years.

[13] For the $Diff\_DW_{it}^{MF}$ target variable; the remaining target variables include observations for more than 50 countries.

Moreover, we can observe that the same rating can only be found on 70% of the observations for the $Diff\_UP_{it}^{SF}$ target variable, reaching as low as 47% of the observations for the rating differences between Moody's and Fitch, when the rating from the first is lower than the rating from the latter. This fact reflects how opaque speculative-grade sovereigns are and how difficult is for credit rating agencies to assess the real capability of these sovereigns to service their debt. This lack of transparency leads to the information available to rating agencies having poor quality and increases the probability of a split rating (Al-Sakka and ap Gwilym, 2010).

The first regressions have the $Diff\_UP_{it}^{SF}$ as the dependent variable and produce significant results for the budget balance and government net debt variables (only one of the regressions with this target variable show the dummy default-in-the-last-5-years variable as significant). The budget balance coefficient is negative, leading to a smaller rating difference between S&P and Fitch when the budget balance grows. Government net debt has the opposite effect on the described rating difference: when it increases, the rating disparity between those two agencies increases as well (see Table 4).

Table 4: Summary of the regressions of the ordered probit speculative-grade subset

	Significant variables (Coefficient sign)	Marginal Effect Rating difference = 1	Marginal Effect Rating difference = 2
$Diff\_UP_{it}^{SF}$	(-) Budget balance (4/4) (+) Net debt (4/4) (-) Default last 5Y (1/4)	-0.002% 0.2% -17.3%	-0.0001% 0.01% -2.8%
$Diff\_DW_{it}^{SF}$	(+) Net debt (4/4) (-) GDP growth (15/16) (-) External debt (15/16) (-) Default last 10Y (3/4)	0.2% -1.2%--1.3% -0.1%--0.2% -11.7%--12.7%	0.04% -0.3%--1% -0.03%--0.07% -3.8%--5.9%
$Diff\_UP_{it}^{MF}$	(+) External debt (2/16) (-) Default last Y (1/4) (-) Default last 5Y (1/4)	0.2% -13.2% -20.4%	0.05% -3.1% -5.6%
$Diff\_DW_{it}^{MF}$	(-) Gross debt (4/4) (-) Inflation (4/4) (-) Default last 10Y (1/4)	-0.3% -0.3% -11%	-0.1%--0.2% -0.1%--0.2% -10.2%
$Diff\_UP_{it}^{SM}$	(-) Net debt (4/4)	-0.2%	-0.06%
$Diff\_DW_{it}^{SM}$	(+) Budget balance (3/4) (-) GDP per capita (12/16) (-) External debt (4/16)	0.007% -0.4%--0.5% -0.2%	0.001% -0.08%--0.1% -0.05%

Note: See notes to Table 2.

Concerning the obtained results when regressing the $Diff\_DW_{it}^{SF}$ variable, it is possible to observe that government net debt, GDP growth, external debt level and the dummy default-in-the-last-10-years variables all have an effect on the rating difference between S&P and Fitch, when the rating from the first is lower than the rating from the latter. The government net debt variable has a positive coefficient, increasing the rating difference when its value increases. The remaining significant variables (GDP growth, external debt level and the dummy default variable) have negative coefficients, so when their value increases (or becomes one, in the case of the dummy variable), the rating difference between S&P and Fitch shrinks.

Only one specification yields significant results when regressing the $Diff\_UP_{it}^{MF}$ variable (rating difference between Moody's and Fitch, with a higher rating from the first agency). External debt has positive and significant coefficients on two of the regressions, therefore when its value increases, the analysed rating difference increases as well. Two of the four dummy default variables (default in the last year and in the last five years) have significant negative coefficients, thus when a default happened in the last year or in the last five years, the rating difference would get smaller.

The regression of the $Diff\_DW_{it}^{MF}$ target variable against the different specifications of predicting variables highlights the effect of government gross debt and inflation on the rating difference between Moody's and Fitch, when the first is lower than the latter (the dummy default-in-the-last-10-years variable only yielded significant and negative results for one of the regressions). Both gross debt and inflation contribute negatively to the rating gap, therefore, the rating difference will shrink if one of those variables increases.

All the ordered probit regressions run with $Diff\_UP_{it}^{SM}$ as the dependent variable show that the government net debt contributes negatively to the rating difference, when the S&P rating is higher than the rating from Moody's. As a result, when the government net debt increases, the rating gap between S&P and Moody's shrinks.

The results from regressing the $Diff\_DW_{it}^{SM}$ target variable show a positive and a negative correlation between the rating difference (when the rating from S&P is lower than the one from Moody's) and, respectively, the budget balance on one hand, and GDP per capita and external debt on the other hand. For this reason, when the budget balance increases, the considered rating gap increases; whereas, when GDP per capita or external debt increase, the same rating gap decreases.

4.4. Simple Probit Estimations

As robustness exercise, we also estimated a simple probit model, with, for instance, $Diff\_UP_{it}^{AB}$ ($Diff\_DW_{it}^{AB}$) = 1 if the rating from agency *A* is higher (lower) than the rating from agency *B* by one or more notches, for sovereign *i* in year *t*, and 0 otherwise:

$$Diff_{UP_{it}^{AB}} = \begin{cases} 1, \textit{ if } |R_{it}^{\alpha} - R_{it}^{\beta}| \geq 1, \\ 0, \textit{ otherwise} \end{cases} \quad \text{when } R_{it}^{\alpha} \geq R_{it}^{\beta}, \tag{4}$$

$$Diff_{DW_{it}^{AB}} = \begin{cases} 1, \textit{ if } |R_{it}^{\alpha} - R_{it}^{\beta}| \geq 1, \\ 0, \textit{ otherwise} \end{cases} \quad \text{when } R_{it}^{\alpha} \leq R_{it}^{\beta}, \tag{5}$$

where A and B and α and $\beta \in \{SF, MF, SM\}$, and α (A) $\neq \beta$ (B).

In this context, our dependent variables have a value of 1 if there is a rating difference of 1-notch or higher and a value of 0 if the ratings from the considered pair of agencies are equivalent in our numerical rating scale. The main results (available on request) essentially go through. A summary comparison between those two sets of results is presented in Table 5.

5. Conclusion

By regressing the rating differences of the three main rating agencies with both an ordered and a simple probit random-effects model, we find some significant results, indicating the influence of some of our explanatory variables on those rating differences.

We used an ordered probit model, due to both the existence of rating differences above two notches and Al-Sakka and ap Gwilym (2010) approach to the split ratings topic. Nonetheless, and because of a lower percentage of rating differences higher than one notch, a simple probit model was also used to find if it improved on the results previously obtained.

For the rating differences between S&P and Fitch, when the assigned rating from the first was higher than the latter, we found that, independently of the dataset (full, investment- or speculative-grade), an increase in the budget balance would decrease the rating difference whereas an increase in net debt would increase that same difference. For the speculative--grade ratings, we also found that the existence of a default in the last two or five years would decrease the rating difference between S&P and Fitch.

When the rating from S&P is lower than the one from Fitch, we find different behaviours when comparing the results from the investment- and speculative-grade datasets: in the first case, GDP per capita contributes for a smaller rating gap, whereas a default in the last year and inflation contribute for a bigger rating difference. In the latter case, only net debt has an increasing effect on the rating difference; external debt, GDP growth and the existence of a default in the last year, two or ten years reduce the rating difference.

The results of our regressions when Moody's assigns a higher sovereign rating than Fitch are less precise. On the other hand, GDP per capita and inflation respectively influence an investment-grade rating difference in a positive and negative way, external debt and a default in the last year or five years respectively increase and decrease the analogous

speculative rating difference. When considering only the investment-grade regressions, our simple probit results also find the budget and structural balances and a default in the last five years as negatively correlated with the rating difference. On the other hand, for the speculative-grade results for Moody's and Fitch positive rating differences, the simple probit approach does not find external debt as significant, when compared with the ordered probit approach for the same dataset. It is also worth noting the fact that GDP growth only appears as significant for the ordered probit regressions with the full dataset.

In terms of rating differences when Moody's assigns a lower rating than Fitch, a higher level of government gross debt leads to a smaller rating difference for both investment- and speculative-grade datasets, with the exception of the simple probit regressions for speculative rating differences, which did not find gross debt significant. Our simple probit regressions with the investment-grade dataset also find that net debt positively affects the rating difference. Inflation is found to negatively influence a rating difference between Moody's and Fitch when the ratings are in the speculative category (irrespective of the chosen probit approach), and a default on the last ten years affect in the same negative way only the rating differences within the ordered probit results.

Looking at the results obtained for the positive rating difference between S&P and Moody's for the investment-grade dataset, we find that an increase in the level of external debt leads to a smaller rating gap between those two rating agencies. For the same dataset, we find that the simple probit approach also identifies GDP per capita as negatively correlated with the rating difference. For the speculative-grade dataset, both probit methods show net debt as negatively related with the rating difference. It's important to note that for this specific dependent variable, and contrary to what was seen on the regressions of the investment- and speculative- data subsets, only the regressions with the full dataset showed all four default dummy variables as significant and affecting positively the rating gap.

The last dependent variable represents the negative rating difference between S&P and Moody's (that is, a lower rating from S&P than from Moody's). Both of our probit regressions with the investment-grade dataset show a positive relation between budget balance, gross debt and GDP growth, the rating difference, and a negative relation between a default in the last year or two and the same rating difference. The simple probit results also point to the structural balance and a default in the last five years as contributing negatively to the rating difference. When considering the speculative-grade dataset, our results both show that an increase in GDP per capita leads to a smaller rating difference. Both budget balance and gross debt affect the rating difference positively, the former only for the ordered probit regressions and the latter only for the simple probit regressions. External debt also affects negatively the rating difference in our ordered probit regressions.

There are a few improvements and further questions that may be addressed in the future. One could also find a way of specifying which agency is responsible for the rating difference, or as an alternative, discover which factors, in a split rating situation, are correlated with a specific agency upgrade or downgrade.

Another question that can be further assessed is considering different periods, for instance the period before the 1997 Asian crisis, or periods before and after the 2008-2009 economic and financial crisis, possibly reflecting differences on how the rating agencies methodologies were applied in those specific periods.

Table 5: Comparison of the results obtained with the random-effects ordered and simple probit estimations for the full, investment-grade and speculative-grade datasets

	Full dataset		Investment-grade dataset		Speculative-grade dataset	
Significant variables	Ordered probit results	Simple probit results	Ordered probit results	Simple probit results	Ordered probit results	Simple probit results
$Diff\_UP_{it}^{SF}$	(-) Budget balance (4/4) (+) Net debt (4/4)	(-) Budget balance (4/4) (+) Net debt (4/4)	(-) Budget balance (1/4) (+) Net debt (4/4)	(+) Net debt (4/4)	(-) Budget balance (4/4) (+) Net debt (4/4) (-) Default last 5Y (1/4)	(-) Budget balance (3/4) (+) Net debt (4/4) (-) Default last 2Y (1/4) (-) Default last 5Y (1/4)
$Diff\_DW_{it}^{SF}$	(-) GDP per capita (16/16) (-) External debt (16/16) (+) Default last 1Y (1/4) (+) Default last 2Y (1/4) (+) Default last 5Y (4/4)	(-) GDP per capita (16/16) (+) Default last Y (1/4) (+) Default last 2Y (1/4) (+) Default last 5Y (4/4)	(-) GDP per capita (15/16) (+) Default last 1Y (1/4)	(-) GDP per capita (12/16) (+) Inflation (13/16) (+) Default last Y (1/4)	(+) Net debt (4/4) (-) External debt (15/16) (-) GDP growth (15/16) (-) Default last 10Y (3/4)	(+) Net debt (4/4) (-) GDP growth (15/16) (-) Default last Y (2/4) (-) Default last 2Y (2/4) (-) Default last 10Y (4/4)
$Diff\_UP_{it}^{MF}$	(-) GDP growth (9/16) (+) External debt (16/16)	(-) Structural balance (3/4) (+) External debt (16/16)	(+) GDP per capita (12/16) (-) Inflation (16/16)	(+) GDP per capita (2/16) (-) Inflation (13/16) (-) Budget balance (1/4) (-) Structural balance (4/4) (-) Default last 5Y (4/4)	(+) External debt (2/16) (-) Default last Y (1/4) (-) Default last 5Y (1/4)	(-) Default last Y (1/4) (-) Default last 5Y (1/4)
$Diff\_DW_{it}^{MF}$	(-) Gross debt (2/4) (+) Net debt (4/4) (+) Default last 2Y (3/4) (+) Default last 5Y (4/4)	(+) Net debt (4/4) (+) Default last 2Y (3/4) (+) Default last 5Y (4/4)	(-) Gross debt (1/4)	(-) Gross debt (1/4) (+) Net debt (4/4)	(-) Gross debt (4/4) (-) Inflation (4/4) (-) Default last 10Y (1/4)	(+) External debt (4/16) (-) Inflation (16/16)
$Diff\_UP_{it}^{SM}$	(+) Default last Y (1/4) (+) Default last 2Y (4/4) (+) Default last 5Y (1/4) (+) Default last 10Y (1/4)	(-) External debt (4/16) (+) Default last 5Y (1/4) (+) Default last 10Y (1/4)	(-) External debt (16/16)	(-) External debt (16/16) (-) GDP per capita (8/16)	(-) Net debt (4/4)	(-) Net debt (4/4)
$Diff\_DW_{it}^{SM}$	(+) Budget balance (+) Gross debt (+) GDP growth (-) GDP per capita	(+) Budget balance (4/4) (+) Gross debt (4/4) (+) GDP growth (4/16) (-) GDP per capita (12/16)	(+) Budget balance (4/4) (+) Gross debt (4/4) (+) GDP growth (8/16) (-) Default last 1Y (4/4) (-) Default last 2Y (4/4)	(+) Budget balance (4/4) (-) Structural balance (2/4) (+) Gross debt (4/4) (+) GDP growth (13/16) (-) Default last Y (4/4) (-) Default last 2Y (4/4) (-) Default last 5Y (4/4)	(-) External debt (4/16) (+) Budget balance (3/4) (-) GDP per capita (12/16)	(+) Gross debt (4/4) (-) GDP per capita (12/16)

Note: See notes to Table 2.

References

Afonso A. (2003) Understanding the determinants of sovereign debt ratings: Evidence for the two leading agencies, *Journal of Economics and Finance*, 27, 56-74.

Afonso, A.; Furceri, D.; Gomes, P. (2012) Sovereign credit ratings and financial markets linkages: Application to European data, *Journal of International Money and Finance*, 31, 606-638.

Afonso, A.; Gomes, P.; Rother, P. (2008) What "Hides" Behind Sovereign Debt Ratings?, in M. Catalão-Lopes, J. Soares, and J. Pina, (eds), *New Developments in Financial Modelling*, Cambridge: Cambridge Scholars Publishing, 314-343.

Afonso, A.; Gomes, P.; Rother, P. (2011) Short- and long-run determinants of sovereign debt credit ratings, *International Journal of Finance & Economics*, 16, 1-15.

Al-Sakka, R.; ap Gwilym, O. (2010) Split sovereign ratings and rating migrations in emerging economies, *Emerging Markets Review*, 11, 79-97.

Amstad, M.; Packer, F. (2015) Sovereign ratings of advanced and emerging economies after the crisis, *BIS Quarterly Review*, December.

Beers, D.; Mavalwalla, J. (2016) Database of Sovereign Defaults, 2016, Bank of Canada Technical Report, Vol. 101.

Bhatia, A. (2002) *Sovereign Credit Ratings Methodology*, International Monetary Fund.

Bissoondoyal-Bheenick, E. (2005) An analysis of the determinants of sovereign ratings, *Global Finance Journal*, 15, 251-280.

Brunnermeier, M. (2009) Deciphering the Liquidity and Credit Crunch 2007-2008, *Journal of Economic Perspectives*, 23, 77-100.

Cantor, R. (2012) Sovereign Risk Measurement Methodologies, Presentation at the 2012 Credit Risk Conference, *NYU Stern School of Business*.

Cantor, R.; Packer, F. (1996) Determinants and impact of sovereign credit ratings, *Economic Policy Review*, 2, 37-53.

Eliasson, A.-C. (2002) Sovereign credit ratings, *Research Notes 02-1*, Deutsche Bank Research.

Fitch Ratings (2014) Definitions of Ratings and Other Forms of Opinion, Technical report, Fitch Ratings.

Hill, P.; Brooks, R.; Faff, R. (2010) Variations in sovereign credit quality assessments across rating agencies, *Journal of Banking & Finance*, 34, 1327-1343.

Kiff, J.; Sylwia, N.; Schumacher. L. (2012) Are Rating Agencies Powerful? An Investigation into the Impact and Accuracy of Sovereign Ratings, Technical report, International Monetary Fund.

Levich, R.; Majnoni, G.; Reinhart, C. (2012) Ratings, Rating Agencies and the Global Financial System, New York University Salomon Center Series on Financial Markets and Institutions, Kluwer.

Livingston, M.; Naranjo, A.; Zhou, L. (2007) Asset Opaqueness and Split Bond Ratings, *Financial Management*, 36, 49-62.

Livingston, M.; Naranjo, A.; Zhou, L. (2008) Split bond ratings and rating migration", *Journal of Banking & Finance*, 32, 1613-1624.

Moody's Investors Service (2015) Sovereign Rating Methodology, Technical report, Moody's Investors Service.

Moody's Investors Service (2016) Rating Symbols and Definitions, Technical report, Moody's Investors Service.

Mulder, C.; Monfort, B. (2000) Using Credit Ratings for Capital Requirements on Lending to Emerging Market Economies: Possible Impact of a New Basel Accord, No. 0-69: International Monetary Fund.

S&P Global Ratings (2016) S&P Global Ratings Definitions, URL: https://www.standardandpoors.com/en_US/web/guest/article/-/view/sourceId/504352, Accessed on 15 September 2016.

Sylla, R. (2012) An historical primer on the business of credit rating, in R. Levich, G. Majnoni, and C. Reinhart, Ratings, Rating Agencies and the Global Financial System, New York University Salomon Center Series on Financial Markets and Institutions, Kluwer, 19-40.

Vu, H.; Alsakka, R.; ap Gwilym, O. (2017) What drives differences of opinion in sovereign ratings? The roles of information disclosure and political risk, *International Journal of Finance and Economics*, 22, 216-233.

Appendix: Data sources

Table A1: Summary of the explanatory variables

Predicting variables	Name, description	Variable original description and source	Description
BudgetBal_NGDP	Budget balance	GGR_NGDP: General government revenue (percentage of GDP) GGX_NGDP: General government total expenditure (percentage of GDP). Source: IMF (WEO)	Difference between government revenues and expenses (GGR_NGDP and GGX_NGDP).
GGSB_NPGDP	Structural balance	General government structural balance (percentage of potential GDP). Source: IMF (WEO)	
GGXWDG_NGDP	Gross debt	General government gross debt (percentage of GDP). Source: IMF (WEO)	
GGXWDN_NGDP	Net debt	General government net debt (percentage of GDP) Source: IMF (WEO)	
NGDP_RPCH	GDP growth rate	Gross domestic product, constant prices Source: IMF (WEO)	Annual percentages of constant price GDP, year-on-year changes.

NGDPDPC	GDP per capita	Gross domestic product per capita, current prices, expressed in current U.S. dollars per person. Source: IMF (WEO)	
PCPIPCH	Inflation	Inflation, average consumer prices Source: IMF (WEO)	Annual percentages of average consumer prices, year-on-year changes.
ED=ExtDebtPercGNI	External debt	GNI_USD: Gross National Income (current US$) Source: WB (WDI) ExtDebtStocksTotalUSD: External debt stocks, total (DOD, current US$) Source: WB (WDI) GrossExtDebtPosition: 0059_T1_Gross External Debt Position and External debt stocks, total (DOD, current US$) Source: JE (QEDS)	The WDI dataset has GNI values for the great majority of countries. The External Debt values existed on the WDI dataset. For OECD countries the QEDS dataset replaced the WDI dataset as the canonical source for external debt. The QEDS dataset has values from 2003 onwards, so we used the external debt values from the WDI dataset (ExtDebtStocksTotalUSD), and then we merged the values from the QEDS dataset when available (GrossExtDebtPosition). ExtDebtPercGNI was calculated using the combined values from WDI and QEDS dataset.
DefaultLastYear	Default in the last year	CRAG database has the values of debt defaulted by countries along the years, distributed by type of creditor (and the definition of 'default' used by the authors is consistent with much of the literature on sovereign defaults).	The debt value defaulted by country and year was processed and converted into a boolean variable named DefaultThisYear (1 if the country, in that year, had debt defaulted; 0 otherwise). Afterwards, the variables DefaultLastYear, DefaultLast2Years, DefaultLast5Years and DefaultLast10Years were created, assuming the value 1 if the value DefaultThisYear had the value 1 in the previous year/two years/five years/ten years, for the same country, and 0 otherwise.
DefaultLast2Years	In the last two years		
DefaultLast5Years	In the last five years		
DefaultLast10Years	In the last ten years		

Notes: The sources of information used in this work were the World Economic Outlook dataset (WEO) from the International Monetary Fund (IMF), the World Development Indicators (WDI) from the World Bank (WB) and the Quarterly External Debt Statistics dataset (QEDS) from the Joint Effort of the WB and the IMF. The variables BudgetBal_NGDP, GGSB_NPGDP, GGXWDG_NGDP, GGXWDN_NGDP, NGDPDPC and ExtDebtPercGNI are expressed in terms of their year-to-year variation.

DOI: https://doi.org/10.14195/2183-203X_46_4

Investimento Empresarial em Portugal: Crise e Recuperação*

Fernando Alexandre
Pedro Bação
Carlos Carreira
João Cerejeira
Gilberto Loureiro
António Martins
Miguel Portela

Received for publication: April 12, 2018
Revision accepted for publication: May 12, 2018

RESUMO

Neste texto descrevemos e procuramos explicar a evolução do investimento empresarial em Portugal nos últimos anos. O facto mais saliente nessa evolução é naturalmente a queda abrupta do investimento após o início da crise financeira internacional. Em termos líquidos, o investimento privado tornou-se negativo, só tendo regressado a valores positivos em 2017. A nossa análise sugere que o elemento principal na explicação dessa evolução é o défice de procura face à capacidade instalada. Outro elemento importante tem sido o nível de confiança dos empresários. Pelo contrário, o custo do investimento parece ter um papel relativamente menor, embora o endividamento e o cash flow apareçam como fatores relevantes. As empresas mais dinâmicas são especialmente sensíveis ao cash flow. A nossa análise sugere igualmente que os últimos anos foram caracterizados por uma elevada incidência de empresas "zombie", embora tenha havido uma redução muito acentuada em 2016. Outra boa novidade recente é o aumento do peso dos setores transacionáveis no investimento. Estes elementos são indicadores de que estará em curso uma alteração estrutural na economia portuguesa.
Palavras-chave: Crise financeira, restrições financeiras, investimento, Portugal, empresas zombie.

* Este artigo baseia-se nos capítulos 2, 3 e 4 do livro 'Investimento Empresarial e o Crescimento da Economia Portuguesa", realizado com o patrocínio da Fundação Calouste Gulbenkian por iniciativa da Presidência da República Portuguesa, às quais agradecemos. Estamos igualmente gratos a Isabel Mota, Artur Santos Silva, Miguel Cadilhe, José Felix Ribeiro, Pedro Pita Barros e Pedro Ramos. Nos casos em que foram publicados dados novos desde a preparação daquele livro, os cálculos e os valores reportados foram atualizados. A afiliação completa dos autores é dada por: Fernando Alexandre, Escola de Economia e Gestão, Universidade do Minho (falex@eeg.uminho.pt); Pedro Bação, Faculdade de Economia, Universidade de Coimbra (pmab@fe.uc.pt); Carlos Carreira, Faculdade de Economia, Universidade de Coimbra (ccarreir@fe.uc.pt); João Cerejeira, Escola de Economia e Gestão, Universidade do Minho (jccsilva@eeg.uminho.pt); Gilberto Loureiro, Escola de Economia e Gestão, Universidade do Minho (gilberto@eeg.uminho.pt); António Martins, Faculdade de Economia, Universidade de Coimbra (amartins@fe.uc.pt); e Miguel Portela, Escola de Economia e Gestão, Universidade do Minho (mangelo@eeg.uminho.pt).

ABSTRACT
In this paper we provide a description and tentative explanations for the evolution of business investment in Portugal in recent years. Naturally, the feature that stands out is the abrupt fall in investment after the beginning of the international financial crisis. Net of capital depreciation, private investment actually became negative; only in 2017 did it became positive again. Our analysis suggests that the main driver of investment has been excess capacity relative to demand. Another important element has been the state of producers' confidence. On the other hand, the cost of capital seems to have had a relatively small role, although indebtedness and the cash flow appear to have some influence on investment decisions. The more dynamic firms are especially sensitive to cash flows. Our analysis also suggests that the last few years have been characterized by high levels of zombie firms, although in 2016 the percentage of zombie firms declined considerably. Another recent piece of good news is the increase of tradable sectors' share in corporate investment. These elements indicate that the Portuguese economy may be undergoing a process of structural change.
Keywords: Financial crisis, financial frictions, investment, Portugal, zombie firms.

JEL Classification: D22; D25; E22; E27

1. Introdução

A entrada no século xxi marcou o início de um período de baixo crescimento da economia portuguesa, pondo fim a quatro décadas de forte crescimento e de convergência para os níveis de rendimento dos países mais ricos da União Europeia. A crise financeira internacional de 2008 agravou drasticamente a situação e trouxe consigo a crise da dívida na Área do Euro. Os níveis elevados de endividamento do sector público, das famílias e das empresas (as mais endividadas do mundo, em termos do rácio da dívida em percentagem do PIB – ver, por exemplo, Alexandre et al., 2017a), tornaram Portugal especialmente vulnerável à turbulência que afetou os mercados financeiros.

O investimento das empresas em Portugal caiu de forma muito acentuada nos anos que se seguiram à crise, à semelhança do que aconteceu nas restantes economias da Área do Euro. Entre estas, as economias resgatadas pela troika foram as mais afetadas, tendo registado maiores quebras e por um período mais prolongado. A economia portuguesa foi uma das que registaram diminuições mais acentuadas, mantendo-se o investimento das empresas em valores ainda muito inferiores aos observados nos anos anteriores à crise. Com efeito, o investimento privado líquido tornou-se negativo e apenas em 2017 terá voltado a ser (apenas ligeiramente) positivo, situando-se muito abaixo dos valores observados antes do início da crise. Em consequência, apesar dos sinais de recuperação, o panorama continua a gerar preocupação, pois o investimento, em particular o investimento empresarial, é essencial para o aumento do produto potencial da economia. É o investimento que introduz nas organizações as instalações, os equipamentos e os processos que incorporam o avanço do conhecimento e que permitirão utilizar de forma mais produtiva os recursos disponíveis. Portanto, é importante refletir sobre a evolução recente do investimento das empresas em Portugal, para que eventuais intervenções possam ser devidamente fundamentadas. Foi isso que procurámos fazer em Alexandre et al. (2017b). Tendo em conta os dados estatísticos publicados entretanto, neste texto atualizamos parte da análise feita então, relativamente à caracterização da evolução do investimento privado e à explicação dessa evolução.

Assim, utilizando modelos macro- e micro-económicos, estimamos e analisamos um conjunto de relações que procuram representar o comportamento do investimento das empresas. A conjugação de modelos com características diferentes possibilitará uma discussão mais abrangente dos fatores determinantes do investimento das empresas. O longo regime de baixo crescimento em que se encontra a economia portuguesa coarta as possibilidades de crescimento das empresas cujas vendas se concentram no mercado interno. No entanto, também o retorno dos investimentos das empresas dos setores transacionáveis, em particular as exportadoras, é afetado por estarem localizadas numa economia com fracas perspetivas de crescimento. Por outro lado, o elevado endividamento das empresas e as dificuldades do setor bancário (que constitui a principal fonte de financiamento externo das empresas portuguesas) têm também sido identificados como restrições ao investimento das empresas. Com efeito, a análise aos dados do Inquérito ao Investimento do Instituto Nacional de Estatística (INE), apresentada no Boletim Económico do Banco de Portugal de outubro de 2016, identifica os seguintes fatores limitativos do investimento das empresas: a deterioração das perspetivas de vendas, a rentabilidade dos investimentos, a falta de capacidade de autofinanciamento, a dificuldade em obter crédito bancário e a taxa de juro. Este estudo

do Banco de Portugal refere que a evidência microeconómica atribui maior importância à posição financeira das empresas, em particular ao seu grau de endividamento, nas decisões de investimento das empresas.

Na Secção 2 começamos por utilizar dados agregados para descrever a evolução do investimento empresarial no contexto da Área do Euro. De seguida, recorremos à base de dados Sistema de Contas Integradas das Empresas (SCIE) do INE, para o período 2010-2016[1], para descrever a evolução do investimento das empresas, tendo em conta a região em que se localizam, a sua dimensão e o setor de atividade.

Na Secção 3 é estimado um modelo macroeconómico do investimento privado, para o período 1987-2017. O modelo permite discutir o papel do excesso de capacidade, do custo do investimento e das expectativas dos investidores. Os resultados obtidos atribuem ao excesso de capacidade uma importante contribuição para o mau desempenho do investimento privado nos últimos anos. Neste modelo macroeconómico, os custos do investimento não parecem constituir uma restrição ao investimento das empresas.

Da Secção 4 em diante, em vez de dados agregados, usamos dados desagregados por empresa. Na Secção 4, estimamos modelos microeconómicos tendo em vista a avaliação do impacto nas decisões de investimento das condições de financiamento, do excesso de capacidade e do potencial de crescimento das empresas. O modelo-base da Secção 4 é complementado na Secção 5 de forma a analisarmos diferenciadamente, de acordo com o dinamismo das empresas, o efeito de eventuais restrições financeiras sobre as decisões de investimento das empresas. Apesar da redução do endividamento nos últimos anos, as empresas portuguesas estão ainda entre as mais endividadas do mundo. O elevado endividamento é, em muitos casos, a origem de uma situação financeira e económica muito frágil, constituindo um entrave ao seu crescimento. De facto, as conclusões dos modelos microeconómicos estimados na Secção 4 sugerem que o endividamento tem um impacto negativo nas decisões de investimento das empresas.

Na Secção 6 caracterizamos o fenómeno das empresas "zombie" em Portugal. Veremos que representam uma parte muito significativa da nossa amostra, o que reforça a ideia de que o setor financeiro teve e tem um problema na afetação do crédito. Esse problema de afetação do crédito tem correspondência num problema macroeconómico de baixo crescimento, pois os recursos financeiros não foram afetados aos setores e às empresas mais dinâmicos, dificultando a transformação estrutural de que a economia portuguesa necessita. Ainda assim, os dados para 2016 mostram uma melhoria substancial da situação. A Secção 7 apresenta as conclusões deste estudo.

[1] Na construção da amostra utilizada na nossa análise, foram consideradas apenas as empresas que cumpriram cumulativamente os seguintes critérios: ter dois ou mais trabalhadores; volume de vendas igual ou superior a mil euros, capitais próprios positivos; não pertencer ao setor financeiro; total do ativo igual ou superior a 10 mil euros; não ter endividamento negativo.

2. A Evolução do Investimento Empresarial

2.1. A Evolução na Área do Euro

Na Tabela 1 apresentamos a evolução do investimento das empresas, em percentagem do PIB, na Área do Euro e para vários países dessa zona monetária, desde 1990. Os dados mostram que existem diferenças significativas entre os países da Área do Euro na taxa de investimento das empresas, o que faz crer que as características dos países são um fator importante na explicação do investimento das empresas – ver, a este propósito, por exemplo, Palenzuela e Dees (2016).

Tabela 1: Formação bruta de capital das empresas (% PIB, média no período indicado)

	2001-2007	2008-2017	2001-2017
Área do Euro (12 países)	12	12	12
Bélgica	15	15	15
Alemanha	12	11	12
Irlanda	12	15	14
Grécia	8	5	7
Espanha	16	14	15
França	12	13	13
Itália	11	10	10
Luxemburgo	11	10	11
Holanda	10	11	11
Áustria	16	15	15
Finlândia	13	11	12
Portugal	13	12	12

Nota: Cálculos dos autores com base em dados da AMECO.

Os dados da Tabela 1 sugerem que, no que diz respeito ao peso do investimento das empresas (em percentagem do PIB), os países da Área do Euro podem dividir-se em três grupos: Alemanha, França, Finlândia, Irlanda e Portugal estão próximos da média (12%-13%); Bélgica, Espanha e Áustria estão claramente acima da média (15%-16%); e a Grécia está significativamente abaixo da média, com uma taxa de investimento das empresas entre 5% e 8% do PIB.

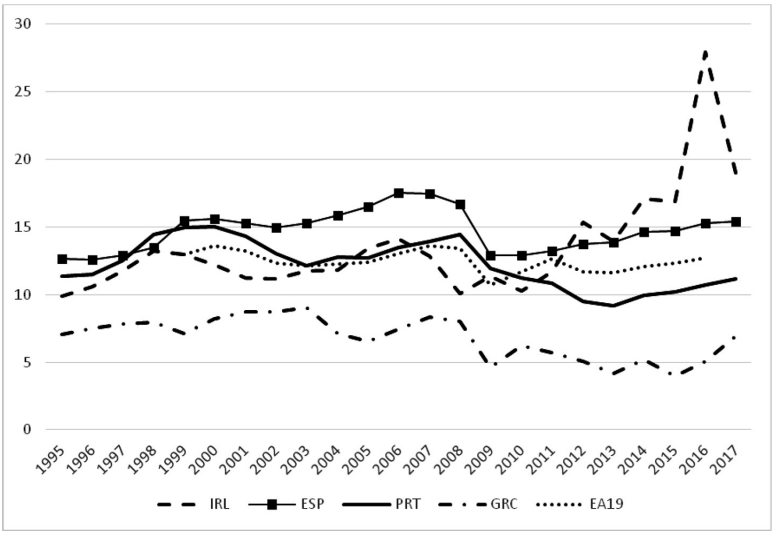
O investimento das empresas nos países intervencionados pela troika (Espanha, Grécia, Irlanda e Portugal) apresentou comportamentos muito diferenciados nas últimas décadas – ver Figura 1. Nos casos espanhol e irlandês, o investimento das empresas registou um aumento significativo nos anos anteriores à crise financeira internacional, atingindo valores máximos, em 2006, de 17,6% e 14,1% do PIB, respetivamente. Em ambos os casos, registou-se uma

queda abrupta do investimento das empresas nos anos seguintes. No caso da Irlanda, o valor mínimo de 10,1% do PIB foi atingido em 2008, tendo-se seguido uma recuperação que levou a que, em 2012, o investimento das empresas se fixasse num valor superior a 15% do PIB e, em 2016-2017, em valores superiores a 19% do PIB – isto é, valores muito superiores aos registados antes da crise financeira internacional. Em Espanha, o valor mínimo da taxa de investimento das empresas, de 12,9% do PIB, foi atingido em 2009 e 2010. Nesse ano, teve início uma recuperação lenta mas contínua do investimento das empresas, que ultrapassou os 15% do PIB em 2016. Deve ainda salientar-se o facto de a taxa de investimento das empresas em Espanha, apesar da gravidade da crise económica e financeira, se ter mantido sempre acima da taxa de investimento da Área do Euro.

A Grécia sobressai por apresentar uma taxa de investimento das empresas muito baixa relativamente aos outros países da Área do Euro – ver Tabela 1. A taxa de investimento das empresas gregas era cerca de 8% do PIB em 2007, tendo diminuído nos anos seguintes para valores em torno dos 5%. Os dados da Figura 1 sugerem que a tendência de queda terá sido invertida em 2016-2017.

Em Portugal, a taxa de investimento das empresas atingiu um valor máximo de 14,4% do PIB em 2008 (aproximando-se do valor máximo de 15% do PIB atingido em 2000) e iniciou uma trajetória descendente, atingindo um valor mínimo de 9,2% em 2013. Em 2014-2015, a taxa de investimento das empresas foi cerca de 10% do PIB, parecendo haver desde então uma tendência crescente – em 2017 foi já 11,1% do PIB. Até 2010, a taxa de investimento das empresas estava em linha com a taxa de investimento da Área do Euro, mas é, desde essa altura, significativamente inferior – ver Figura 1.

Figura 1: Formação bruta de capital (% PIB)

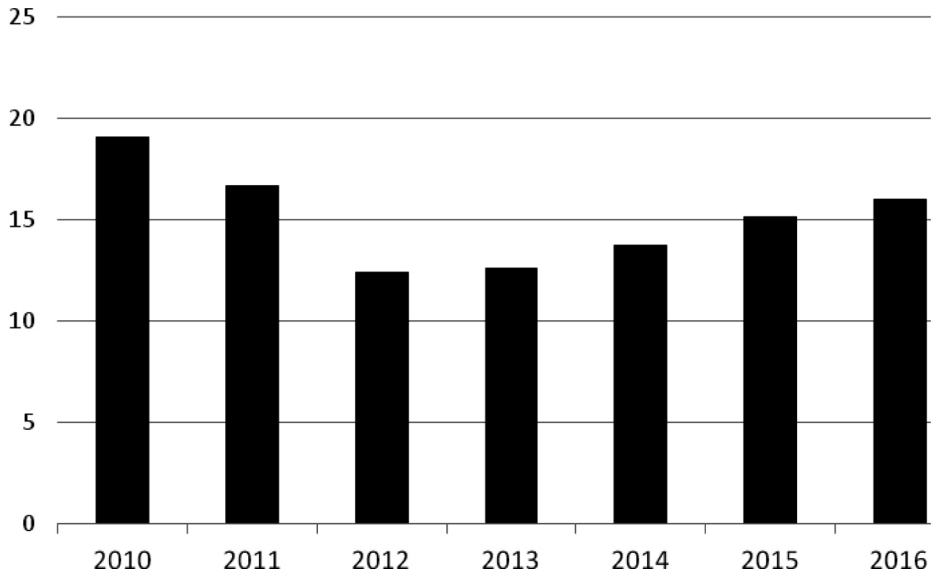


Notas: Cálculos dos autores com base em dados da AMECO. IRL: Irlanda; ESP: Espanha; PRT: Portugal; GRC: Grécia; EA19: Área do Euro (19 países).

2.2. A Evolução em Portugal

Na Figura 2 apresentamos a evolução do valor do investimento das empresas entre 2010 e 2016, em termos reais (preços de 2015), calculado a partir dos dados por empresa da base de dados SCIE do INE. No período em análise, o investimento das empresas atingiu um valor mínimo em 2012, de 12 mil milhões de euros, o que representa uma queda superior a 30% relativamente a 2010, ano em que o investimento chegou aos 19 mil milhões de euros. Em 2013, o investimento das empresas começou a recuperar, tendo ultrapassado os 16 mil milhões de euros em 2016 (ainda 16% abaixo do valor observado em 2010).

Figura 2: Evolução do investimento agregado das empresas (mil milhões de euros, a preços de 2015)



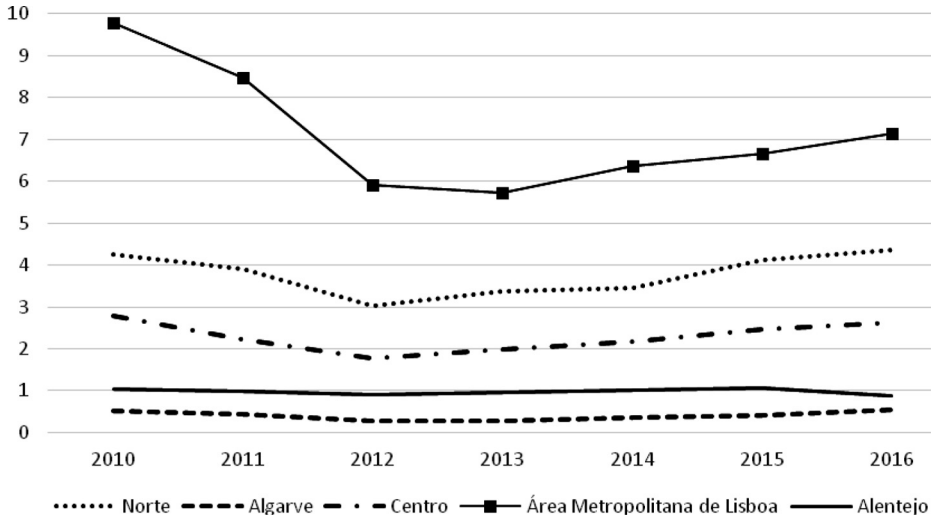
Nota: Cálculos dos autores com base em dados do SCIE, INE.

Investimento das empresas por região

A Figura 3 mostra a evolução do investimento das empresas por NUTS 2. Nessa evolução destaca-se a importância da quebra do investimento na Área Metropolitana de Lisboa (doravante, Região de Lisboa), que diminuiu cerca de 40% entre 2010 e 2012. Em 2016 o investimento das empresas na Região de Lisboa era ainda inferior em 27% ao registado em 2010. Nesta região, o valor mínimo do investimento foi registado em 2013. Ao nível dos setores de atividade, na quebra do volume de investimento na Região de Lisboa destacam-se os setores da "Indústria transformadora" (-27%, com 15% do investimento total em 2016), da "Construção" (-60%, com 2,3% do investimento total em 2016) e do "Comércio"

(-8%, com 15% do investimento total em 2016). Na Região de Lisboa, apenas os seguintes setores registavam, em 2016, um volume de investimento superior ao registado em 2010: "Agricultura" (+69%), "Atividades de consultoria" (+24%) e "Alojamento, restauração e similares" (+26%).

Figura 3: Evolução do investimento das empresas por região, NUTS 2 (mil milhões de euros, a preços de 2015).



Nota: Cálculos dos autores com base em dados do SCIE, INE.

Na Região Norte, a queda foi menor (29% de 2010 a 2012) e a recuperação ficou completa em 2016, ano em que o volume de investimento empresarial na Região Norte era já superior em cerca de 3,3% face ao registado em 2010. Os sectores de atividade que apresentaram uma recuperação mais forte do investimento foram a "Indústria transformadora" (+53%, e 43% do investimento total em 2016), as "Atividades imobiliárias" (+21%, 0,7% do investimento total em 2016), e "Atividades administrativas" (56%, e 2% do investimento total em 2016). Também na Região do Algarve o volume de investimento em 2016 ultrapassou ligeiramente o valor observado em 2010. Pelo contrário, no Centro e especialmente no Alentejo, o volume de investimento empresarial em 2016 ainda estava abaixo dos níveis de 2010, em 5% e 17%, respetivamente.

Em resultado das tendências acima descritas, entre 2010 e 2016, a Região de Lisboa perdeu importância no investimento total do País, passando o seu peso de 51,2% para 45%. Por outro lado, o peso do investimento das empresas da Região Norte e da Região Centro no investimento total aumentou, respetivamente, de 22,3% para 27,1% e de 14,6% para 16,5%. Concluindo, a não-recuperação do investimento das empresas para os valores de 2010 é explicada essencialmente pela não-recuperação do investimento das empresas localizadas na Região de Lisboa.

Investimento por dimensão das empresas

Na Tabela 2 apresenta-se a evolução da repartição do investimento das empresas por escalão de dimensão, definido pelo número de trabalhadores, e, na última linha, a variação do investimento para cada um dos grupos entre 2010 e 2016. Conclui-se que a maior quebra no investimento se registou nas empresas de maior dimensão (pelo menos 250 trabalhadores), com uma quebra de 22,9%. As empresas com um número de trabalhadores entre 10 e 49 foram as que registaram a menor quebra no investimento, com uma redução inferior a 1% entre 2010 e 2016.

A Tabela 2 permite também concluir que as grandes empresas, com pelo menos 250 trabalhadores, representam 30% a 40% do investimento empresarial total. O peso do investimento das microempresas no investimento total tem permanecido próximo dos 16%.

Tabela 2: Repartição do investimento total das empresas por dimensão da empresa

	Número de trabalhadores				
Anos	1-9	10-49	50-99	100-249	≥ 250
2010	16,1%	19,1%	10,8%	16,2%	37,8%
2013	16,3%	20,4%	10,1%	14,3%	38,9%
2016	16,0%	22,6%	10,8%	16,1%	34,6%
Δ 2010 – 2016	-15,8%	-0,7%	-16,1%	-16,5%	-22,9%

Notas: Cálculos dos autores com base em dados do SCIE, INE (investimento a preços de 2015). "Δ" indica a variação do investimento de cada categoria entre 2010 e 2016.

Investimento das empresas por setor de atividade

Na Tabela 3 podemos ver o peso de cada setor de atividade no investimento total das empresas em 2010 e em 2016. Em termos de importância no investimento total, destaca-se o investimento das empresas do setor "Indústrias transformadoras" (22,5% do investimento total das empresas em 2010 e 27,9% em 2016), do setor "Comércio" (14,5% do investimento total em 2010 e 15,7% em 2016), e do setor "Atividades de informação e de comunicação" (11,5% do investimento total das empresas em 2010 e 8,9% em 2016). Destacam-se ainda as fortes quebras do peso do investimento das empresas dos setores da "Água" (de 6% para 2%) e da "Construção". Em sentido inverso evoluiu o peso do sector "Alojamento, restauração e similares" (de 4,4% para 6,3%).

Tabela 3: Repartição do investimento total das empresas por ramo de atividade

	2010	2016
Agricultura, silvicultura e pesca	2,3%	3,1%
Indústrias extrativas	0,9%	0,7%
Indústrias transformadoras	22,5%	27,9%
Eletricidade, gás vapor, água quente e fria e ar frio	7,8%	6,4%
Captação, tratamento e distribuição de água; saneamento, gestão de resíduos e despoluição	6,0%	2,0%
Construção	5,4%	3,3%
Comércio por grosso e a retalho; reparação de veículos automóveis e motociclos	14,5%	15,7%
Transportes e armazenagem	7,7%	6,8%
Alojamento, restauração e similares	4,4%	6,3%
Atividades de informação e de comunicação	11,5%	8,9%
Atividades administrativas e dos serviços de apoio	7,2%	8,2%

Fonte: Cálculos dos autores com base em dados do SCIE, INE.

Na Figura 2 vimos que o investimento das empresas atingiu um mínimo em 2012 e que está em recuperação desde então. A Tabela 4 (taxas de variação do investimento por setor de atividade) permite concluir que apenas os setores da "Agricultura" (+18%), das "Indústrias transformadoras" (+4%) e do "Alojamento, restauração e similares" (+21%) tinham, em 2016, um volume de investimento superior ao registado em 2010. Setores como os da "Eletricidade" (-31%), "Água" (-72%), e "Construção" (-49%) tinham, em 2016, valores de investimento muito inferiores aos registados em 2010. No entanto, alguns setores registaram uma recuperação significativa entre 2013 e 2016, como o "Comércio" (+36%), os "Transportes e armazenagem" (+44%) e as "Atividades administrativas e dos serviços de apoio" (+66%).

Tabela 4: Taxas de variação do investimento das empresas

	2010-2013	2013-2016	2010-2016
Agricultura, silvicultura e pesca	3%	15%	18%
Indústrias extrativas	-11%	-25%	-33%
Indústrias transformadoras	-25%	39%	4%
Eletricidade, gás vapor, água quente e fria e ar frio	-7%	-26%	-31%
Captação, tratamento e distribuição de água; saneamento, gestão de resíduos e despoluição	-60%	-29%	-72%
Construção	-45%	-7%	-49%
Comércio por grosso e a retalho; reparação de veículos automóveis e motociclos	-35%	40%	-8%
Transportes e armazenagem	-48%	44%	-26%
Alojamento, restauração e similares	-34%	84%	21%
Atividades de informação e de comunicação	-41%	12%	-34%
Atividades administrativas e dos serviços de apoio	-42%	66%	-3%

Nota: Cálculo dos autores com base em dados do SCIE, INE (a preços de 2015).

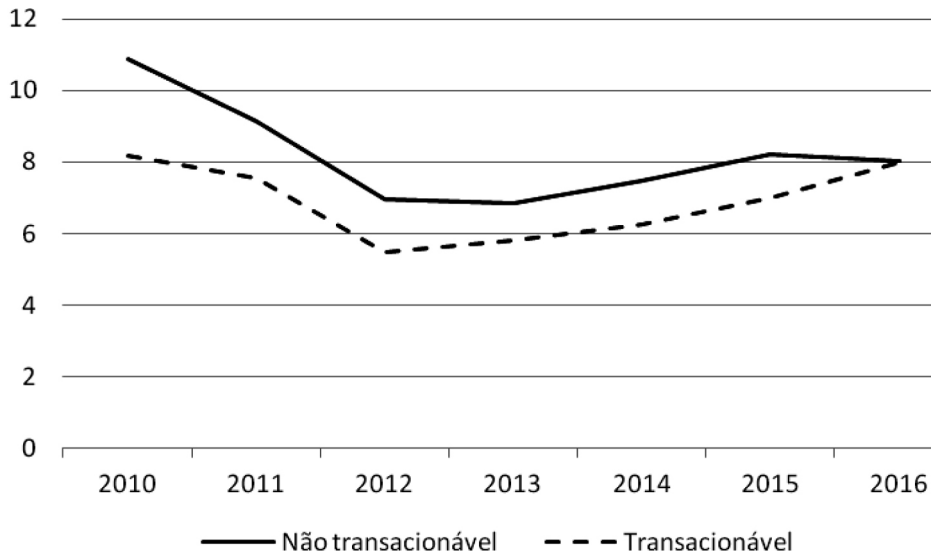
Transacionáveis e não transacionáveis

O peso dos setores de atividade tem sido alvo de atenção especial no contexto do debate sobre papel dos setores ditos "não-transacionáveis", isto é, que não estão sujeitos à concorrência internacional. Na prática, a distinção entre transacionáveis e não-transacionáveis não é evidente. Em Alexandre et al. (2017b) usámos como critério o peso das exportações e das importações no total dos recursos dos ramos de atividade, classificando como não-transacionáveis aqueles em que esse peso é inferior a 15%. Usamos a mesma classificação neste texto[2].

[2] Essa transposição resulta na seguinte classificação: Setores transacionáveis: Agricultura, silvicultura e pesca; Indústrias extrativas; Indústrias transformadoras; Transportes e armazenagem; Atividades de consultoria, científicas, técnicas e similares; Atividades administrativas e dos serviços de apoio. Setores não-transacionáveis: Eletricidade, gás vapor, água quente e fria e ar frio; Captação, tratamento e distribuição de água; Saneamento, gestão de resíduos e

De acordo com os dados apresentados na Figura 4, o investimento das empresas dos setores não-transacionáveis caiu 36% entre 2010 e 2012, e o das empresas dos setores transacionáveis caiu 33%. No período de recuperação do investimento das empresas, entre 2012 e 2015, o investimento nos setores não-transacionáveis aumentou 17,8%, tendo registado uma ligeira quebra de cerca de 2% em 2016. Por outro lado, o investimento das empresas dos sectores transacionáveis, entre 2012 e 2016, apresentou uma clara tendência de crescimento, tendo aumentado 32%. A conjugação das variações naqueles dois períodos refletiu-se num aumento do peso do investimento das empresas dos setores transacionáveis para 50% do investimento total em 2016. Este resultado é mais um indicador da alteração estrutural que está em curso na economia portuguesa.

Figura 4: Investimento das empresas, setores transacionáveis e não-transacionáveis (mil milhões de euros, preços de 2015)



Nota: Cálculos dos autores com base em dados do SCIE, INE.

3. Um Modelo para o Investimento Privado Agregado

Na secção anterior caracterizámos a evolução do investimento nos últimos anos. Nesta secção e na próxima procuramos explicar essa evolução; nesta adotando uma perspetiva

despoluição; Construção; Comércio por grosso e a retalho; Reparação de veículos automóveis e motociclos; Alojamento, restauração e similares; Atividades de informação e de comunicação; Atividades imobiliárias; Educação; Atividades de saúde humana e apoio social; Atividades artísticas, de espetáculos, desportivas e recreativas; Outras atividades de serviços.

agregada, na próxima uma perspetiva desagregada. Começamos então por estimar um modelo com vista a identificar os principais fatores que influenciam a evolução do investimento privado agregado. A variável dependente será o valor real (a preços de 2010) do investimento privado[3] líquido, isto é, deduzindo a depreciação do stock de capital existente, pois tal é o conceito relevante do ponto de vista económico.

Os modelos de inspiração neoclássica – ver, por exemplo, Hall e Jorgenson (1967) – sugerem que o investimento líquido será uma função da diferença entre o stock de capital necessário para satisfazer a procura de bens e serviços, medida pelo PIB, e o stock de capital existente. Por conseguinte, no modelo estimado incluímos o primeiro desfasamento do PIB real (recolhido da base de dados AMECO – base de dados macroeconómicos da Comissão Europeia), do stock de capital do setor privado (estimado a partir de dados da AMECO), duma medida do custo de oportunidade do investimento[4] (função da taxa de depreciação do stock de capital, da taxa de juro e da variação do preço dos bens de capital – todas estas variáveis recolhidas na AMECO) e dum indicador de confiança (Indicador de Confiança da Indústria Transformadora produzido pelo INE). Além destas variáveis, o modelo inclui também o primeiro desfasamento do investimento privado líquido. Desta forma, as variáveis referidas anteriormente definem o valor para o qual o investimento privado líquido tenderá a convergir. Os dados são anuais e a amostra utilizada abrange o período de 1987 a 2017. O investimento líquido (inv), o PIB real (PIB) e o stock de capital (K) estão medidos em milhares de milhões de euros e avaliados a preços de 2010. As outras duas variáveis são representadas abaixo como CCP, no caso do custo de oportunidade de capital, e IC, no caso do indicador de confiança. O modelo estimado tem a seguinte forma:

$$inv_t = \beta_1 + \beta_2 PIB_{t\text{-}1} + \beta_3 K_{t\text{-}1} + \beta_4 CCP_{t\text{-}1} + \beta_5 IC_{t\text{-}1} + \beta_6 inv_{t\text{-}1} + \varepsilon_t. \quad (1)$$

Este modelo é semelhante, embora bastante mais parcimonioso, aos modelos discutidos em trabalhos recentes do Banco Central Europeu (Palenzuela e Dees, 2016) e do FMI (IMF, 2015), sobre o investimento na Área do Euro e na economia mundial, respetivamente. O modelo do estudo do BCE, em que a variável dependente é a formação bruta de capital fixo das empresas (business investment), foi estimado para um painel de 16 países da Área do Euro, com dados anuais de 2003 a 2013. Na versão base do modelo, apenas a taxa de utilização da capacidade e a taxa de juro têm coeficientes significativamente diferentes de zero. As outras variáveis, como os custos unitários do trabalho, o grau de abertura, lucros retidos e um índice de perceção da corrupção, não são estatisticamente significativas.

Os modelos estimados no estudo do FMI também seguem a mesma linha, partindo da abordagem do modelo do acelerador – ver, por exemplo, Chenery (1952) – e experimentando acrescentar outras variáveis. De acordo com os resultados apresentados pelo FMI, no caso português, o modelo que inclui o índice de incerteza (para a Área do Euro) parece ser o mais adequado, pelo menos nos anos mais recentes. Julgamos que o indicador de confiança que incluímos no nosso modelo poderá ajudar a captar o mesmo tipo de efeito.

[3] Investimento total menos investimento público (calculados a partir de dados da AMECO).

[4] Aquilo que é habitualmente designado por "user cost of capital".

Os resultados da estimação estão na Tabela 5. Como esperado, o coeficiente do PIB desfasado tem sinal positivo e o coeficiente do stock de capital tem sinal negativo; portanto, o hiato entre o stock de capital necessário para satisfazer a procura e o stock de capital existente tem coeficiente positivo. Este resultado está de acordo com as previsões do modelo do acelerador, ou seja, o investimento das empresas aumenta quando o nível do produto excede a capacidade produtiva. Dito de outro modo, a existência de excesso de capacidade tem um efeito negativo no investimento das empresas.

Também como esperado, o coeficiente do custo de oportunidade tem sinal negativo, refletindo, nomeadamente, o efeito adverso do custo de financiamento sobre o investimento. Finalmente, o indicador de confiança, que capta as expectativas dos investidores em relação ao crescimento futuro, tem sinal positivo. Como veremos abaixo, o nosso modelo sugere que o indicador de confiança tem efetivamente um papel importante na determinação do investimento em Portugal. Desta forma, os resultados fazem lembrar uma das ideias centrais do modelo proposto por John Maynard Keynes, que atribuiu grande importância, na explicação das flutuações macroeconómicas, às variações do investimento originadas pelos "animal spirits" ("estados de alma") dos empresários.

Tabela 5: Resultados da estimação do modelo do investimento privado agregado

Variável	Coeficientes
PIB_{t-1}	0,208**
	(0,086)
K_{t-1}	-0,082**
	(0,031)
CCP_{t-1}	-0,015
	(0,094)
IC_{t-1}	0,117***
	(0,035)
inv_{t-1}	0,569***
	(0,119)

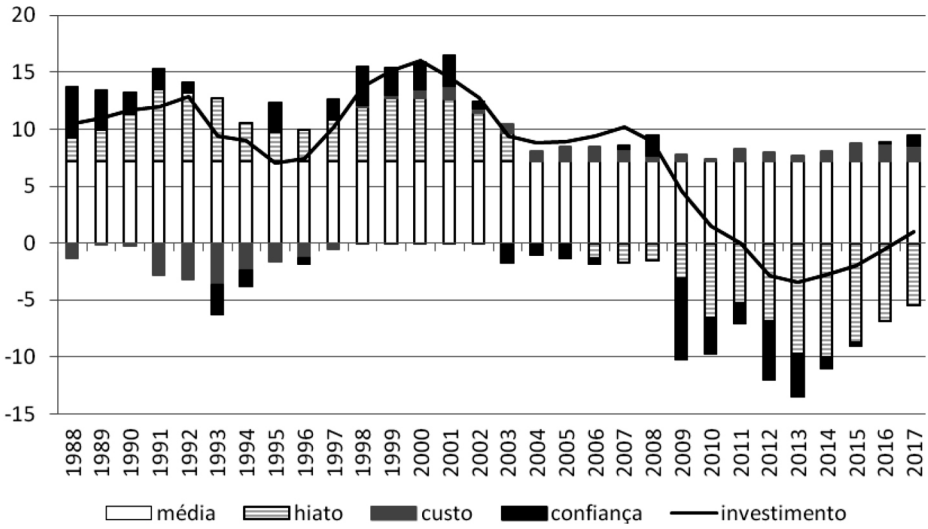
Notas: Cálculos dos autores com base em dados da AMECO e do INE. Níveis de significância: ***, 1%; **, 5%; *, 10%. Erros-padrão entre parêntesis. A média e o desvio-padrão da variável dependente são, respetivamente, 7,469 e 5,852. A variável dependente é a formação líquida de capital fixo das empresas a preços de 2010. O R^2 é 0,97 e a regressão é globalmente significativa: F(5,24) = 146,8. O número de observações é de 30. O modelo inclui uma constante não reportada na tabela.

A Figura 5 mostra a evolução do investimento privado líquido, juntamente com a contribuição de cada elemento do modelo para o valor para o qual o investimento irá tender (que, portanto, corresponderá à sua "tendência"). A contribuição é medida como a diferença em relação à média dessa contribuição ao longo do período. O gráfico inclui um elemento constante que representa exatamente a média do investimento ao longo do período.

De acordo com os dados apresentados na Figura 5, o excesso de capacidade foi o fator que mais contribuiu para a quebra do investimento privado após o início da crise financeira internacional, bem como para a sua recuperação recente. Note-se que o investimento líquido voltou a ser positivo em 2017, ou seja, a capacidade produtiva voltou a aumentar, apesar de as nossas estimativas indicarem que o excesso de capacidade continua a ser significativo (ainda que esteja a diminuir).

Nos anos da crise financeira, entre 2009 e 2013, o investimento privado líquido foi também muito afetado negativamente pelo estado da confiança dos investidores, mas este efeito dissipou-se nos últimos anos. Em termos agregados, o custo do investimento, nos últimos anos, não parece constituir a principal restrição à realização de investimento pelos privados. Os modelos estimados na próxima secção, utilizando dados microeconómicos, analisarão a questão das condições de financiamento do investimento com maior detalhe.

Figura 5: Investimento privado líquido e contributos para a evolução da sua tendência (desvios em relação à média)



Nota: Cálculos dos autores.

4. Determinantes do Investimento das Empresas: Uma Análise Microeconómica

Nesta secção são estimados três modelos, usando dados do SCIE ao nível da empresa, com o objetivo de identificar os fatores que influenciam as decisões de investimento das empresas. O primeiro modelo é estimado pelo método dos mínimos quadrados simples, isto é, o modelo supõe que, pelo menos aproximadamente, nada distingue as decisões de investimento das empresas a não ser o facto de as variáveis que, de acordo com o modelo,

determinam o investimento (mais propriamente, a taxa de investimento da empresa, definida como o quociente entre o investimento e o ativo) tomarem valores diferentes para cada empresa. O segundo modelo admite que haja heterogeneidade entre as empresas, ou seja, que, mesmo que as variáveis que, de acordo com o modelo, determinam o investimento tomassem valores iguais para todas as empresas, haveria fatores específicos a cada empresa que diferenciariam sistematicamente os níveis de investimento. Este modelo será estimado através do estimador de "efeitos fixos", aproveitando a dimensão longitudinal dos dados. O terceiro modelo tem um carácter diferente do dos dois primeiros. Enquanto os dois primeiros modelos procuram explicar o montante investido (em proporção do stock de capital da empresa), o terceiro modelo procura explicar por que é que certas empresas investem (querendo com isto significar que aumentam o stock de capital) e outras não. Portanto, a variável dependente neste modelo será binária (dummy) – tomará o valor um quando a empresa investir e o valor zero quando não investir – e o modelo será um Probit. Este modelo permite-nos avaliar o efeito que as variáveis explicativas têm sobre a probabilidade de uma empresa investir e não sobre o montante investido, como era o caso nos dois primeiros modelos.

A escolha das variáveis utilizadas nos modelos baseou-se na extensa literatura existente sobre este tema[5]. Com as variáveis por nós selecionadas, pretendemos avaliar a importância das condições de financiamento do investimento – em particular, do nível de endividamento –, do excesso/défice de capacidade produtiva, do potencial de crescimento das empresas e da sua dimensão.

Como indicadores das condições de financiamento do investimento, considerámos o nível de endividamento das empresas, a capacidade de gerar recursos internamente (medida pelo cash flow) e a tangibilidade dos ativos (que pode ser vista como uma medida da facilidade de acesso ao crédito, pois os bancos preferem conceder crédito a empresas com bens que possam ser apresentados como garantia). O interesse no estudo do impacto das condições financeiras sobre as decisões de investimento, na linha dos trabalhos realizados por Fazzari et al. (1988) ou Bernanke et al. (1999), tem sido grande nos últimos tempos, pois a atual crise económica teve na sua origem uma crise financeira que afetou de sobremaneira o setor financeiro e a concessão de crédito. No modelo estimado na secção anterior, utilizando dados macroeconómicos, o custo do financiamento não parece constituir uma restrição ao investimento. Porém, tal resultado não significa que não haja empresas para as quais o custo do financiamento seja um obstáculo à realização de projetos de investimento. Num contexto de elevado endividamento das empresas e de dificuldades no setor bancário, é natural que muitas empresas se sintam limitadas no acesso ao crédito, possibilidade que pretendemos ter em atenção nos modelos microeconómicos estimados nesta secção.

As questões do excesso ou défice de capacidade e potencial de crescimento das empresas são reconhecidamente importantes para as decisões de investimento, como vimos a respeito do modelo macroeconómico da secção anterior. No entanto, uma situação de excesso/défice de capacidade em termos agregados pode corresponder a situações muito diferentes consoante os ramos de atividade, ou mesmo de empresa para empresa dentro

[5] Sobre o tema da identificação dos determinantes de investimento, salientamos os estudos de Fazzari, Hubbard e Petersen (1988), Chirinko (1993), Lang, Ofek e Stulz (1996), Aivazian, Ge e Qiu (2005) e Bond, Rodano e Serrano-Velarde (2015).

do mesmo ramo. Tal possibilidade ganha especial acuidade numa economia com baixo crescimento e em processo de alteração estrutural como a portuguesa – ver Alexandre et al. (2017a). Para ter esta situação em linha de conta, incluímos como variáveis explicativas no modelo a taxa de crescimento das vendas e a dimensão exportadora das empresas. A taxa de crescimento das vendas poderá refletir a eventual existência de excesso (se for baixa) ou de défice (se for elevada) de capacidade, ou o potencial de crescimento da empresa (maior se a taxa de crescimento for elevada). Por seu turno, as empresas exportadoras poderão ter um potencial de crescimento mais elevado, pois o seu mercado não estará limitado à economia nacional.

Por fim, teremos em atenção a possibilidade de o comportamento das empresas ser diferente em função da sua dimensão, possivelmente por as necessidades de investimento serem diferentes, ou por as condições de acesso ao financiamento terem outras características. Com este fito, incluiremos no modelo variáveis relacionadas com o nível de vendas e de emprego da empresa. Em complemento, são feitas estimações utilizando diferentes subamostras, definidas com base na dimensão das empresas. Ao procedermos desta forma, estaremos a colocar a possibilidade de os parâmetros que descrevem o comportamento das empresas serem diferentes conforme a dimensão da empresa.

A amostra da base de dados do SCIE do INE foi utilizada para estimar pelo método dos mínimos quadrados simples (OLS) o seguinte modelo:

$$inv_{it} = \beta_0 + \beta_1 endiv_{i,\,t-1} + \beta_2 export_{i,\,t-1} + \beta_3 txvendas_{i,\,t-1} + \beta_4 tang\_ativos_{i,\,t-1} + \beta_5 cflow_{it} + \beta_6 lnvendas_{i,\,t-1} + \beta_7 lnpessoal_{i,\,t-1} + \beta_8 nova_i + \varepsilon_{it}. \quad (2)$$

Na equação (2), o primeiro índice (i) identifica a empresa, e o segundo (t) identifica o período temporal. O modelo estimado pelo estimador de efeitos fixos apenas difere do modelo acima pelo facto de o termo constante (β_0) variar entre empresas. Em ambos os modelos, a variável dependente, *inv*, é a taxa de investimento da empresa, medida pelo quociente entre o investimento líquido total no ano *t* (em ativos fixos tangíveis e intangíveis) e o total do ativo do ano anterior.

As variáveis explicativas são as seguintes: *endiv* é uma medida do nível de endividamento da empresa, dada pelo rácio do passivo não-corrente sobre o capital próprio; *export* mede as exportações da empresa como percentagem do volume de negócios, podendo ser entendida como uma proxy para oportunidades de crescimento das empresas, pois as empresas exportadoras têm um mercado potencial maior; *txvendas* representa a taxa de crescimento das vendas, podendo também servir como uma medida das oportunidades de crescimento das empresas, em particular de empresas não-cotadas; $tang_{ativos}$ é uma medida da tangibilidade dos ativos, dada pelo peso dos ativos fixos tangíveis nos ativos totais, procurando medir o nível de garantias que a empresa pode oferecer aos seus credores; *cflow* é o rácio dos resultados líquidos mais depreciações sobre o total do ativo no período anterior, medindo os fundos gerados pela atividade da empresa; *lnvendas* é o logaritmo das vendas, sendo uma proxy para a dimensão da empresa; *lnpessoal* é o logaritmo do número de trabalhadores, outra proxy para a dimensão da empresa; *nova* é uma variável que identifica as empresas mais jovens, definidas como tendo sido criadas depois de 2009.

No modelo Probit, a variável dependente é, como escrevemos atrás, uma variável binária que, em cada ano, identifica as empresas com investimento líquido positivo. As variáveis explicativas são as mesmas que descrevemos acima.

Tabela 6: Determinantes do Investimento

		OLS	FE
endiv	Dívida LP/ Capital Próprio $_{t-1}$	-0,0002***	-0,0005***
		(0,0000)	(0,0001)
export	Rácio de Exportações $_{t-1}$	0,0036***	0,0116***
		(0,0012)	(0,0035)
txvendas	Taxa Cresc. Vendas $_{t-1}$	0,0130***	0,0050***
		(0,0003)	(0,0004)
$tang_{ativos}$	Tangibilidade dos Activos $_{t-1}$	0,0074***	0,0747***
		(0,0011)	(0,0031)
cflow	Cash Flow $_{t}$	0,0843***	0,0746***
		(0,0007)	(0,0012)

Notas: Cálculos dos autores com base em dados da SCIE. Níveis de significância: ***, 1%; **, 5%; *, 10%. Desvio padrão entre parêntesis. A variável dependente de ambos os modelos é a taxa de investimento da empresa. O número de observações é de 575621. As regressões incluem, adicionalmente, controlos para a dimensão, vendas e número de trabalhadores, bem como a indicação se se trata de uma empresa nova. A regressão OLS inclui ainda *dummies* para o ano e para o setor, enquanto a regressão FE (modelo de efeitos fixos) inclui apenas as *dummies* para o ano.

Tanto no caso do modelo estimado por OLS como no caso do modelo de efeitos fixos (FE) por empresa, os resultados das estimações – ver Tabela 6 – corroboram as previsões da teoria relativamente aos fatores determinantes do investimento, sendo todos os coeficientes estatisticamente significativos. Em particular, observa-se que as empresas mais endividadas – isto é, com maior peso da dívida de longo prazo no total dos capitais próprios – têm taxas de investimento mais baixas. Ainda relativamente ao efeito da capacidade de financiamento das empresas nas decisões de investimento, os resultados da estimação sugerem que as empresas com maior nível de colateral têm mais facilidade em aceder a financiamento para a realização de investimento.

Por outro lado, conclui-se que as empresas com mais recursos financeiros gerados internamente, medidos pelo cash flow, investem mais. Numa situação de mercados financeiros eficientes, a relação desta variável com o investimento tende a ser nula. A existência de uma correlação positiva entre o cash flow e o investimento das empresas sugere que a política de investimento está, pelo menos parcialmente, dependente da geração de fundos internos. Uma situação deste tipo é tipicamente interpretada como evidência de restrições ao financiamento. Esta questão será analisada mais em detalhe nos modelos apresentados na secção seguinte.

Os resultados corroboram igualmente a hipótese de que as empresas com maior potencial de crescimento (ou com necessidade de suprir um eventual défice de capacidade produtiva instalada), quer este seja medido pelo crescimento das vendas quer pela sua capacidade exportadora, investem mais.

Os resultados da estimação do modelo Probit vão ao encontro dos resultados dos dois modelos anteriores. Resumidamente, a probabilidade de investir é maior em empresas com uma estrutura de capitais mais favorável ao uso de capitais próprios, empresas mais exportadoras, com taxas de crescimento das vendas mais elevadas, maiores rácios de ativos tangíveis e maiores cash flows. Dada a estabilidade das estimativas, optou-se por não listar estes resultados.

Tendo em conta que a natureza das decisões de investimento das empresas mais pequenas pode ser diversa da das empresas com maior dimensão, o modelo descrito acima foi estimado eliminando da amostra, alternadamente, as empresas com menos de 10 trabalhadores, as empresas com menos de 10 trabalhadores e 250 ou mais trabalhadores, e as empresas com 250 ou mais trabalhadores. No entanto, o resultado das estimações com qualquer uma daquelas subamostras são muito semelhantes aos obtidos anteriormente, quer no modelo OLS, quer no modelo com efeitos fixos, pelo que não são aqui reportados.

5. Investimento, Dinamismo e Restrições Financeiras

Nesta secção analisamos com maior detalhe a relevância das restrições financeiras nas decisões de investimento das empresas, distinguindo as empresas com base no seu dinamismo económico. A estimação levada a cabo nesta secção permitirá avaliar se as restrições no financiamento do investimento das empresas constituem um obstáculo ao crescimento da economia, na medida em que limitem o investimento e o crescimento das empresas mais dinâmicas.

As empresas mais dinâmicas são definidas como aquelas que apresentam taxas de crescimento das vendas superiores à mediana do ano ou do setor no ano. Dada a importância do setor exportador para o crescimento sustentável da economia portuguesa, identificam-se também as empresas com maior peso das exportações nas vendas totais e as restrições financeiras por elas sentidas nas decisões de investimento.

Nesta análise, as restrições financeiras são avaliadas com base num modelo que testa a sensibilidade do investimento ao cash flow. Este modelo, que tem sido amplamente usado na literatura, foi inicialmente proposto por Fazzari et. al. (1988). O modelo assenta na hipótese de que as empresas com mais dificuldades em obter financiamento externo estão mais dependentes dos fundos que geram internamente. Por essa razão, os níveis de investimento realizados estarão positivamente correlacionados com os cash flows gerados.

O modelo é o seguinte:

$$inv_{it} = \beta_0 + \beta_1 cflow_{it} + \beta_2 lnactivo_{i,\,t-1} + \beta_3 txvendas_{i,\,t-1} + \beta X_{it} + \varepsilon_{it}. \quad (3)$$

As variáveis *inv*, *cflow* e *txvendas* têm a mesma definição que demos na secção anterior. Assim, *inv* é a taxa de investimento da empresa (relativamente ao total do ativo do ano

anterior); *cflow* é o rácio dos resultados líquidos mais depreciações sobre o total do ativo no período anterior; e *txvendas* representa a taxa de crescimento das vendas. A variável *lnactivo* é o logaritmo do total do activo, usado como controlo para a dimensão da empresa. Para além das variáveis indicadas, neste modelo são também incluídas dummies por setor e ano e um conjunto de outras variáveis, representadas por *X* na equação (3). Estimámos várias versões da equação (3), correspondendo a diferentes escolhas para as variáveis representadas por *X*. Procurámos com as nossas escolhas para essas variáveis destacar grupos de empresas com características que possam ser vistas como indiciadoras de "maior dinamismo", ou seja, de maior crescimento atual ou potencial. A introdução dessas variáveis no modelo permitirá verificar se o comportamento desses grupos de empresas é diferente do comportamento das empresas "menos dinâmicas".

A identificação dos grupos de empresas mais dinâmicas foi feita utilizando as seguintes variáveis binárias, incluídas à vez no vetor *X*: "crescimento 1", que identifica as empresas com uma taxa de crescimento das vendas superior à mediana do ano; "crescimento 2", que identifica as empresas com taxa de crescimento das vendas superior à mediana do setor/ano; "exporta", que identifica as empresas exportadoras; "setor exportador", que identifica as empresas pertencentes aos setores com maior peso das exportações nas vendas totais, sendo essa classificação baseada na mediana daquele rácio em cada ano.

Para testar se as empresas com mais potencial de crescimento apresentam ou não uma maior sensibilidade do investimento ao cash flow, incluímos igualmente no vetor *X* a interação das variáveis binárias acima descritas com o cash flow.

A coluna (1) da Tabela 7 mostra as estimativas dos coeficientes da equação (3) no caso em que o vetor *X* é omitido do modelo. Esta coluna corresponde no essencial aos modelos estimados na secção anterior e serve como referência. Note-se que os resultados apresentados na coluna (1) mostram que os níveis de investimento das empresas estão positivamente correlacionados com os cash flows, o que sugere que as empresas podem estar a enfrentar restrições no acesso ao financiamento externo, situação que pode condicionar os planos de investimento das empresas.

Tabela 7: Sensibilidade do investimento ao *cash flow*

	(1)	(2)	(3)	(4)	(5)
Cash Flow $_{t}$	0,0867***	0,0451***	0,0468***	0,0808***	0,0683***
	(0,0007)	(0,0010)	(0,0010)	(0,0007)	(0,0010)
Log Ativo Total $_{t-1}$	-0,0064***	-0,0057***	-0,0055***	-0,0069***	-0,0063***
	(0,0001)	(0,0001)	(0,0001)	(0,0001)	(0,0001)
Taxa Cresc. Vendas $_{t-1}$	0,0125***	0,0124***	0,0124***	0,0122***	0,0124***
	(0,0003)	(0,0003)	(0,0003)	(0,0003)	(0,0003)
Crescimento $_{1}$		0,0211***			
		(0,0003)			
Crescimento 1 x Cash Flow		0,0668***			
		(0,0014)			
Crescimento 2			0,0215***		
			(0,0003)		
Crescimento 2 x Cash Flow			0,0614***		
			(0,0014)		
Exporta				0,0039***	
				(0,0005)	
Exporta x Cash Flow				0,1019***	
				(0,0030)	
Setor Exportador					-0,0063***
					(0,0011)
Setor Exportador x Cash Flow					0,0396***
					(0,0014)

Notas: Cálculos dos autores com base em dados da SCIE. Níveis de significância: ***, 1%; **, 5%; *, 10%. Desvio padrão entre parêntesis. A variável dependente dos vários modelos é a taxa de investimento da empresa. O número de observações é de 575621; quer o R^2 quer o RMSE são idênticos nos 5 modelos: 0,06 e 0,12, respetivamente.

As restantes colunas da Tabela 7 correspondem à introdução no modelo das variáveis binárias (e correspondente interação com a variável cash flow) referidas acima: "Crescimento 1", "Crescimento 2", "Exporta" e "Setor Exportador". Independentemente da medida de dinamismo utilizada, os coeficientes das interações com o cash flow são positivos

e estatisticamente significativos. De acordo com estes resultados, as empresas com maior potencial de crescimento são aquelas em que o investimento depende mais da capacidade de gerar recursos próprios. Ou seja, as empresas mais dinâmicas parecem ser aquelas que têm maiores restrições financeiras no financiamento do investimento. Os resultados desta análise estão sintetizados na Tabela 8, que mostra que o efeito do cash flow no investimento é positivo e muito significativo. Com a exceção da medida de dinamismo baseada nas exportações por setor, a sensibilidade do investimento relativamente ao cash flow (indicador do efeito das restrições financeiras) é mais do dobro nas empresas dinâmicas.

Tabela 8: Investimento das empresas mais dinâmicas e *cash flow*

Definição de dinamismo	(-) Dinâmicas	(+) Dinâmicas
Crescimento 1	0,0451	0,0451 + 0,0668 = 0,1119
Crescimento 2	0,0468	0,0468 + 0,0614 = 0,1082
Exporta	0,0808	0,0808 + 0,1019 = 0,1827
Setor Exportador	0,0683	0,0683 + 0,0396 = 0,1079

Notas: ver estimativas apresentadas na Tabela 7. Definição de dinamismo: "Crescimento 1" corresponde a "empresas com uma taxa de crescimento das vendas superior à mediana do ano"; "Crescimento 2" corresponde a "empresas com taxa de crescimento das vendas superior à mediana do setor/ano"; "Exporta" corresponde a "empresas exportadoras"; "Setor Exportador" corresponde a "empresas pertencentes aos setores com maior peso das exportações nas vendas totais, sendo essa classificação baseada na mediana daquele rácio no ano".

Os modelos apresentados na Tabela 7 foram também estimados numa subamostra que exclui as empresas com menos de 10 trabalhadores, sendo os resultados muito semelhantes aos da Tabela 7.

6. Empresas "Zombie"

Passamos agora à análise da incidência de empresas "zombie" na nossa amostra e, em particular, nos grupos de empresas mais e menos dinâmicas. Genericamente, a designação de empresas "zombie" aplica-se a empresas com situação financeira débil, com forte dependência em relação aos bancos e incapazes de cumprirem as suas obrigações financeiras para com estes.

Ambos, bancos e devedores em dificuldade de solver os seus compromissos financeiros, têm pouco incentivo para revelarem que os empréstimos são incobráveis. Com efeito, é do interesse dos bancos assumirem uma política de tolerância para com os seus mutuários em dificuldade financeira, a fim de não terem de registar imparidades com empréstimos incobráveis. Um banco pode poupar-se a um aumento das provisões para cobrir empréstimos em incumprimento se disponibilizar crédito suficiente às empresas insolventes para que estas

paguem os juros e amortizem os empréstimos pendentes, evitando que as empresas declarem imediatamente falência. Naturalmente, esta "luz verde" de empréstimos também beneficia as empresas insolventes, pois pelo menos adia a sua falência. No entanto, como mostram Caballero et al. (2008), o prolongamento da vida das empresas "zombie" impede a realocação de recursos humanos e do investimento das empresas inviáveis para as empresas mais lucrativas. No contexto da economia portuguesa, cujo processo de alteração estrutural tem enfrentado desde o início do século XXI bloqueios difíceis de ultrapassar, o prolongamento artificial da vida de empresas pelos bancos pode ser um obstáculo adicional à alteração estrutural da economia e ao aumento do seu potencial de crescimento.

Classificamos uma empresa como "zombie" com base em dois critérios: (1) se a empresa recebe "crédito subsidiado" de acordo com a metodologia definida por Caballero et al. (2008); e (2) se o lucro da empresa é menor que a bonificação de juros, conforme sugerido por Fukuda e Nakamura (2011). Assim, o procedimento para identificar as empresas "zombie" envolve três etapas:

Etapa 1: Cálculo do valor mínimo do juro que se esperaria que fosse exigido à empresa em cada ano. Este juro mínimo é o juro que a empresa pagaria caso as taxas de juro (de curto e de longo prazo) das suas dívidas fossem aquelas que são aplicadas a devedores de risco baixo[6].

Etapa 2: Cálculo do hiato das taxas de juros, ou seja, a diferença entre a taxa de juro média efetivamente paga pela empresa em cada ano e a taxa de juro média que corresponderia ao juro mínimo calculado na Etapa 1. De acordo com a interpretação dada, por exemplo, por Caballero et al. (2008), se o hiato for negativo, a empresa recebeu um "subsídio" (pagou menos juros do que pagariam devedores de risco baixo com as mesmas dívidas) e o seu índice *"zombie"* é 1; caso contrário, é 0.

Note-se a filosofia conservadora assumida na Etapa 1 ao selecionarmos taxas de juro extremamente vantajosas para o mutuário: o juro mínimo é de facto menor do que a maioria das empresas pagaria na ausência de juros subsidiados. Deve-se também notar que, dado o procedimento seguido na determinação do hiato, não seremos capazes de detetar todos os tipos de empréstimos subsidiados. Em particular, se um banco faz novos empréstimos a uma empresa a taxas de juros normais para pagar os empréstimos passados, a nossa variável hiato não revela a existência de subsídio. Do mesmo modo, se um banco compra ativos a um cliente a preços excessivamente generosos, a nossa medida não deteta a assistência.

Etapa 3: Posto que na Etapa 2 existe a possibilidade de classificarmos erroneamente as não- "zombie" como "zombie" apenas porque funcionam bem e têm um custo com juros relativamente baixo, se o lucro da empresa for suficiente para pagar o valor "subsidiado" (diferença entre o juro mínimo e o pagamento efetivo de juros), então a empresa é reclassificada como uma não- "zombie".

[6] Como taxa de juro de referência de curto prazo, usamos a taxa de juro do mercado interbancário a 3 meses e, como taxa de longo prazo, a taxa de juro das obrigações do Estado a 10 anos. Fonte: OCDE.

Incidência de empresas "zombie" na amostra

Nas Tabelas 9 e 10 reportamos a incidência de empresas *"zombie"* por região e por dimensão para todos os anos da amostra[7]. Em termos agregados, é no ano de 2012 (35%) que se verifica uma maior incidência de empresas *"zombie"*, tendo essa percentagem vindo a decrescer. Em 2015 (26,1%) ainda estava acima dos valores de 2011 (25,5%), mas em 2016 já se situava claramente abaixo (16,4%), possivelmente refletindo uma mudança de comportamento da banca portuguesa, a limpeza dos balanços dos bancos e a própria melhoria da situação económica.

É nos segmentos onde há empresas mais endividadas que se encontra a maior percentagem de empresas *"zombie"* (ver Alexandre et al., 2017b). Assim, em termos de incidência por região, no total dos cinco anos, é nas regiões da Madeira, Algarve e Área Metropolitana de Lisboa que se encontra a maior concentração de empresas *"zombie"*. É também nas empresas de menor dimensão que se encontra a maior incidência destes casos, em qualquer dos anos.

Tabela 9: Incidência de empresas *"zombie"* região/ano

Região	2011	2012	2013	2014	2015	2016
Norte	25,9%	34,2%	30,5%	27,0%	24,6%	16,5%
Algarve	29,8%	38,4%	34,3%	29,4%	24,7%	15,1%
Centro	23,7%	33,9%	30,3%	27,0%	23,9%	17,3%
Lisboa	27,6%	37,0%	32,4%	28,4%	25,0%	14,1%
Alentejo	24,0%	32,7%	30,5%	27,3%	24,8%	16,9%
Açores	19,0%	28,1%	30,8%	30,2%	25,4%	16,4%
Madeira	28,8%	44,1%	39,3%	36,6%	34,1%	18,7%
Total	25,5%	35,5%	32,6%	29,4%	26,1%	16,4%

Nota: Cálculos dos autores com base em dados do SCIE.

Apesar da redução da incidência das empresas 'zombie' desde 2012, em 2016, na nossa amostra, aquelas empresas tinham ainda um peso significativo na economia, representando 13,9% dos ativos tangíveis (30,6% em 2012), 9,9% do emprego (25,5% em 2012) e 8,3% do endividamento (35,7% em 2012).

[7] O ano de 2010 não está incluído nas tabelas porque a metodologia usada para identificar as empresas *"zombie"* obriga a recorrer a variáveis com desfasamento de um ano.

Tabela 10: Incidência por dimensão da empresa/ano

Dimensão	2011	2012	2013	2014	2015	2016
Trabalhadores 1-9	28,2%	37,3%	33,6%	30,1%	26,7%	17,0%
Trabalhadores 10-49	19,1%	28,2%	23,9%	20,5%	19,1%	13,7%
Trabalhadores 50-99	13,4%	21,3%	17,1%	13,0%	13,2%	9,8%
Trabalhadores 100-249	9,8%	20,4%	16,1%	11,2%	9,3%	7,3%
Trabalhadores ≥250	6,4%	22,0%	16, 5%	5,7%	4,5%	2.9%
Total	25,5%	35,5%	32,6%	29,4%	26,1%	16,4%

Nota: Cálculos dos autores com base em dados do SCIE.

Empresas "zombie" e empresas dinâmicas

Na Tabela 11 apresentamos a percentagem de empresas "zombie" por grupo (empresas mais dinâmicas versus empresas menos dinâmicas, usando os mesmos indicadores do dinamismo das empresas que usámos na Secção 5: crescimento e exportações), a diferença de incidência entre os grupos e um teste para averiguar a significância estatística dessa diferença.

Os resultados sugerem que a percentagem de empresas "zombie" é significativamente menor entre as empresas mais dinâmicas. Em média, o grupo de empresas mais dinâmicas (usando as várias classificações identificadas na tabela) apresenta uma incidência de empresas "zombie" menor em 11 a 13 pontos percentuais – uma diferença que é estatisticamente significativa e de magnitude relevante. Este resultado evidencia, mais uma vez, que as empresas menos dinâmicas são, em média, as que apresentam maiores problemas de solvabilidade, sendo por isso mais "subsidiadas" pelos bancos, os quais, dessa forma, adiam um eventual registo de imparidades. Empréstimos bancários adicionais concedidos a estas empresas sob a forma de um "subsídio" (i.e., com taxas de juros baixas que não refletem o seu nível de risco) representam crédito que não está ao serviço das melhores oportunidades de crescimento das empresas nacionais. Em vez disso, servem um propósito de camuflagem de eventuais perdas para os bancos.

Tabela 11: Empresas *"zombie"* por grupos de empresas mais e menos dinâmicas

Crescimento (1)	24,69%	
Restantes (0)	38,05%	
Dif. (1) - (0)	-13,36%	***
Exportadoras (1)	21,79%	
Restantes (0)	33,16%	
Dif. (1) - (0)	-11,37%	***

Notas: Cálculos dos autores com base em dados do SCIE. O teste usa a estatística *t*. Níveis de significância: ***, 1%; **, 5%; *, 10%. Em caso de a estatística ter significância, a hipótese rejeitada é a da igualdade da média nos dois grupos ("1" e "0").

7. **Conclusão**

A descrição da evolução do investimento das empresas que apresentámos mostra que o investimento caiu a pique a seguir ao início da crise financeira internacional, que há sinais de retoma desde 2013, mas que essa retoma é ainda muito ténue face à dimensão da queda do investimento. De facto, os dados da base SCIE mostram uma diminuição do investimento empresarial superior a 30% entre 2010 e 2012, e de 16% entre 2010 e 2016. O investimento privado (investimento total menos investimento público) líquido foi negativo entre 2012 e 2016, ou seja, o stock de capital privado esteve em queda durante cinco anos.

A redução do investimento das empresas foi especialmente forte na Região de Lisboa (40% entre 2010 e 2012), nos sectores da "Água" (72% entre 2010 e 2016) e da "Construção" (49% entre 2010 e 2016) e nas maiores empresas (23% entre 2010 e 2016). O investimento diminuiu tanto nos setores transacionáveis como nos não-transacionáveis. No entanto, em 2016 o investimento nos setores transacionáveis já estava ao nível de 2010, algo que não aconteceu nos setores não-transacionáveis. Como consequência, o investimento das empresas em 2016 encontrava-se repartido igualmente entre os setores transacionáveis e os não-transacionáveis, contrariando o enviesamento a favor dos não-transacionáveis que tinha sido característico. Este facto sugere que está em curso uma alteração estrutural da economia portuguesa.

Numa primeira abordagem à questão das causas da evolução do investimento nas últimas décadas estimámos um modelo com dados agregados para o investimento privado líquido. As estimativas obtidas indicam que, nos últimos anos, o investimento em Portugal foi prejudicado essencialmente por um excesso de capacidade produtiva instalada face à procura. Igualmente importante desde o início da crise financeira internacional, mas a mostrar mais recentemente sinais claros de recuperação, foi a falta de confiança dos empresários. No modelo agregado, o custo do investimento, nos últimos anos, não parece constituir a principal restrição ao investimento.

Numa segunda abordagem, estimámos modelos com dados desagregados ao nível da empresa. De acordo com estas estimações, os seguintes fatores estão associados a níveis de

investimento mais elevados: o "dinamismo" da empresa (exportar ou ter taxas de crescimento das vendas elevadas); ter níveis elevados de ativos tangíveis (eventualmente úteis como garantia para empréstimos) e de cash flow (financiamento interno do investimento). Pelo contrário, níveis elevados de endividamento estão associados a taxas de investimento mais baixas. Portanto, do ponto de vista microeconómico, a situação financeira da empresa é um fator relevante na decisão de investimento.

Os resultados indicam que as empresas "mais dinâmicas" são mais sensíveis ao cash flow. Isto é, as suas decisões de investimento dependem em maior grau dos fundos que conseguirem gerar a partir da sua atividade. Este resultado tem sido interpretado na literatura financeira como uma indicação de dificuldades no acesso a financiamento com fundos externos. No entanto, estes resultados podem refletir uma opção mais conservadora de gestão financeira das empresas, sobretudo quando o período em análise foi antecedido de uma grave crise financeira. De qualquer forma, estes resultados sugerem a existência de empresas com potencial para apresentarem um crescimento mais rápido assente numa maior diversidade de fontes de financiamento, nomeadamente com o reforço da componente de financiamento externo.

Por outro lado, estes resultados mostram também a existência de oportunidades para as instituições financeiras, em particular para o setor bancário. Nos casos em que a opção pelo financiamento assente fundamentalmente em recursos próprios resultar de dificuldades na gestão, as instituições financeiras poderão contribuir para ultrapassar essas dificuldades, fornecendo competências técnicas e os meios financeiros que permitam um crescimento mais rápido e sustentado das empresas.

No entanto, em termos gerais, o setor financeiro tem tido problemas com a atribuição de crédito às empresas. A nossa análise sugere que o setor financeiro tem estado a suportar empresas "zombie", procurando evitar o reconhecimento da perda dos créditos concedidos a essas empresas. Os resultados apresentados indicam que, nos últimos anos, mais de um quarto das empresas poderá ter estado nessa situação, com destaque para o ano de 2012, em que poderão ter sido mais de um terço. No entanto, deve destacar-se a forte diminuição registada nos últimos anos no número de empresas 'zombie'. As decisões do setor bancário sobre as empresas "zombie" que continuará a apoiar e as que deixará cair terão um importante impacto na velocidade e na qualidade da alteração da estrutura setorial da economia. Há sinais positivos neste domínio, tal como na evolução do volume de investimento e no peso dos setores transacionáveis. A confirmação desses sinais será muito importante para o crescimento da economia portuguesa e para o bem-estar dos portugueses.

Referências Bibliográficas

Aivazian, V.; Ge, Y.; Qiu, J. (2005) The impact of leverage on firm investment: Canadian evidence, *Journal of Corporate Finance*, 11(1-2), 277-291.

Alexandre, F.; Aguiar-Conraria, L.; Bação, P.; Portela, M. (2017a) *Poupança e Financiamento da Economia Portuguesa*, Lisboa: Imprensa Nacional-Casa da Moeda e Associação Portuguesa de Seguradores.

Alexandre, F. (Coordenação); Bação, P.; Carreira, C.; Cerejeira, J.; Loureiro, G.; Martins, A.; Portela, M. (2017b) *Investimento Empresarial e o Crescimento da Economia Portuguesa*, Lisboa: Fundação Calouste Gulbenkian.

Bernanke, B.S.; Gertler, M.; Gilchrist, S. (1999) The financial accelerator in a quantitative business cycle framework, in J. B. Taylor; M. Woodford (eds.), *Handbook of Macroeconomics,* Vol. 1, Amsterdam: Elsevier, 1341-1393.

Bond, S.; Rodano, G.; Serrano-Velarde, N. (2015) Investment dynamics in Italy: Financing constraints, demand and uncertainty, *Working paper*, Banca d'Italia.

Caballero, R.J.; Hoshi, T.; Kashyap, A.K. (2008) Zombie Lending and Depressed Restructuring in Japan, *American Economic Review*, 98(5), 1943-1977.

Chenery, H.B. (1952) Overcapacity and the Acceleration Principle, *Econometrica*, 20(1), 1-28.

Chirinko, S. (1993) Business fixed investment spending: Modelling strategies, empirical results and policy implications, *Journal of Economic Literature*, 31(4), 1875-1911.

Fazzari, S.; Hubbard, R.; Petersen, B. (1988) Financing constraints and corporate investment, *Brookings Papers on Economic Activity*, 1, 141-206.

Fukuda, S.; Nakamura, J. (2011). Why Did "Zombie" Firms Recover in Japan? *The World Economy*, 34(7), 1124-1137.

Hall, R.E.; Jorgenson, D. (1967). Tax Policy and Investment Behavior, *American Economic Review*, 57(3), 391-414.

IMF (2015) World Economic Outlook: Uneven Growth -Short- and Long-Term Factors, Washington DC: International Monetary Fund.

Lang, L.; Ofek, E.; Stulz, R. (1996) Leverage, investment, and firm growth, *Journal of Financial Economics,* 40(1), 3-29.

Palenzuela, D.R.; Dees, S. (eds.) (2016) Savings and investment behavior in the euro area, ECB Occasional Paper Series no. 167.

Sara Cerdeira
sara.jm.cerdeira@gmail.com

DOI: https://doi.org/10.14195/2183-203X_46_5

The Business Plan in the Context of Corporate Entrepreneurship: A Literature Review

Sara Cerdeira

Received for publication: January 8, 2018
Revision accepted for publication: April 9, 2018

ABSTRACT
Business planning is seen by many authors as an important step in the entrepreneurial process, supporting companies in business development. However, there is no unanimity about the importance of the business plan, since both benefits and costs arise from business planning. This paper intends to contribute to this debate by analyzing the business plan in the context of corporate entrepreneurship through a literature review. Since business planning requires the spending of company's resources but helps the company obtain financing, our analysis leads to the conclusion that business planning may be like an investment, that is, it is necessary to spend resources to gain more resources and, hence, the choice of the company to plan will depend on the evaluation of whether the funding obtained will surpass the resources spent.
Keywords: Entrepreneurship; corporate entrepreneurship; entrepreneurial process; business plan.

JEL Classification: D01; C70; B41.

Acknowledgments: The author would like to thank Miguel Torres Preto for advising and guiding in the writing of this paper. I would also like to thank the Faculty of Economics of the University of Coimbra for recognizing my effort during the Master's Degree in Economics with the Award FEUC Exemplar 2.º Ciclo.

RESUMO
Planear um negócio é visto por muitos autores como um passo importante no processo empreendedor, apoiando as empresas no seu desenvolvimento. No entanto, não há unanimidade sobre a importância do plano de negócios, uma vez que do planeamento surgem tanto benefícios como custos. Este artigo pretende contribuir para este debate, analisando o plano de negócios no contexto do empreendedorismo corporativo através de uma revisão da literatura. Dado que planear o negócio requer o uso de recursos, mas ajuda a empresa a obter financiamento, a nossa análise leva à conclusão de que o plano de negócios assemelha-se a um investimento: é necessário despender recursos para obter mais recursos e, portanto, a escolha da empresa em planear dependerá da sua avaliação de que o financiamento obtido ultrapassará os recursos gastos.

1. Introduction

Entrepreneurship has been linked with the concepts of new entry and innovation and described as the exploitation of new opportunities by entrepreneurs (Ardichvili et al., 2003; Choi and Shepherd, 2004; Lumpkin and Dess, 1996; Shane and Venkataraman, 2000; Stevenson and Gumpert, 1985). It refers to companies' entry into new markets (when they are able to identify and exploit a new business opportunity), or to innovations or renewals within the companies (Lumpkin and Dess, 1996; Shane and Venkataraman, 2000; Sharma and Chrisman, 1999). While entrepreneurship can take place in the creation of a company (start-up entrepreneurship), the focus of this paper will be the entrepreneurship that takes place within established companies – corporate entrepreneurship (Thornberry, 2001), which results from their need to continue innovating and acting entrepreneurially so that they can maintain or gain a competitive advantage and continue to grow (Covin and Miles, 1999; Thornberry, 2001).

Entrepreneurship is also an important contribution to economic growth. Several studies have analyzed the relationship between entrepreneurship and economic growth. For instance, Castaño et al. (2016), that used different datasets to determine the effect of several indicators related to entrepreneurship (e.g. characteristics of the entrepreneur, Research & Development (R&D) policy) on economic growth (Gross Domestic Product (GDP) per capita) and found a positive effect. Galindo and Méndez, 2014, used panel data of 13 countries to assess the effect of entrepreneurship and innovation on economic growth (GDP) and found a positive effect. In turn, Audretsch and Keilbach, 2004, found that entrepreneurship capital (defined by the authors as the endowment in factors that contribute to the development of new businesses) leads to economic growth (GDP).

However, it is not a one way relationship. Entrepreneurship can have a positive contribution on economic growth but economic development will also lead to new opportunities for entrepreneurs, acting as an incentive to act entrepreneurially (Galindo and Méndez, 2014). The positive effect of entrepreneurship on economic growth can occur in several ways. For instance, the knowledge created by entrepreneurial activities leads to the development of new things (e.g. new products, capacities), through innovation and investment on R&D (Audretsch et al., 2008; Cumming et al., 2014). Even when entrepreneurs do not develop something new they can contribute to economic growth just by creating new firms, increasing total production in the economy and, thus, increasing employment, wages and competition (Audretsch and Thurik, 2004; Minniti and Lévesque, 2010). In short, Given the relevance of entrepreneurship to economic growth it seems vital that governments create incentives and policies to stimulate entrepreneurial activities, thus generating more positive prospects to potential entrepreneurs that the benefits will surpass the costs (Audretsch and Keilbach, 2004; Cumming et al., 2014; Freytag and Thurik, 2007).

To act entrepreneurially companies follow a process (that may be different among them) – the entrepreneurial process, which is essentially the pursuit of a possible opportunity (Ardichvili et al., 2003; Choi and Shepherd, 2004; Stevenson and Gumpert, 1985). Authors have different perspectives of the stages that integrate this process. One stage that is mentioned by some is the conception of a business plan (Ardichvili et al., 2003; Haber and Reichel, 2007), which acts as a support in the exploitation of the opportunity, allowing

entrepreneurs to gather and analyze crucial information and to make forecasts about what will be the value created to the company (Chwolka and Raith, 2012; Honig, 2004). Although the business plan is thought by some as a fundamental support in business development, others disagree given that it also generates costs (Karlsson and Honig, 2009). Thus, the business plan will be further analyzed in this paper to help understand the debate between planning and not planning.

The paper will begin with an analysis of corporate entrepreneurship to determine its benefits and the different types of corporate entrepreneurship that can take place in a firm. Then it will be presented the entrepreneurial process and its various stages. The paper will then focus on the topic business plan, namely its goals, the planning process, and the benefits and costs of the business plan formulation.

2. Corporate Entrepreneurshirp

Corporate entrepreneurship is according to Sharma and Chrisman (1999: 18): "the process whereby an individual or a group of individuals, in association with an existing organization, create a new organization or instigate renewal or innovation within that organization." That is, it occurs within a company or organization already established in the market and can consist on the formation of a new business, an innovation or renewal in the company or the creation of a new organization, using the resources of the existing company (Sharma and Chrisman, 1999; Thornberry, 2001; Wolcott and Lippitz, 2007).

The main purpose of corporate entrepreneurship is to create economic value for the company, contributing to the company's performance and competitive advantage. The development of something new in the company, which occurs in corporate entrepreneurship, will also result in an increase in the entrepreneur's knowledge and skills (Covin and Miles, 1999; Thornberry, 2001), and may later reflect in more benefits for the company. However, corporate entrepreneurship does not only result in benefits for companies. When a company chooses to be entrepreneurial, it has to take into account that it is also engaging in a risky activity since a significant investment of the company's resources is required and there is no certainty of success (Thornberry, 2001; Wolcott and Lippitz, 2007). Thus, risk-taking behavior becomes an essential characteristic of entrepreneurs. They must be willing to face risky situations so that they can have truly entrepreneurial initiatives (Miller, 1983).

Corporate entrepreneurship can then occur in four ways according to Thornberry (2001). First, by encouraging the company's employees to have entrepreneurial initiatives – Intrapreneuring; second, by developing a new business within the company – Corporate Venturing; third, through the Organizational Transformation of the company; and finally, by the company's change of the competition rules to which it is subject – Industry Rule-Bending. These four types of corporate entrepreneurship are presented in Table 1.

Table 1: Types of corporate entrepreneurship described by Thornberry (2001)

Types of Corporate Entrepreneurship	
Intrapreneuring	Companies aim to turn their workers into entrepreneurs, allowing them to create innovations within the company.
Corporate Venturing	Creation of a new business within a company that distinguishes itself from the current business.
Organizational Transformation	An innovation or a reorganization of resources leads to the development of new business opportunities.
Industry Rule-Bending	The company tries to change the competition rules in the market in which it operates.

Intrapreneuring, also defined as intrapreneurship, is a concept that is also used in the literature to refer to corporate entrepreneurship in the general sense (Russell, 1999) or to identify the situation in which ideas for new products emerge from one or more individuals within a company (Covin and Miles, 1999). However, this paper focuses on the definition given by Thornberry (2001). Intrapreneuring is therefore the type of corporate entrepreneurship that occurs when companies aim to turn their workers into entrepreneurs so that the company's employees create innovations in the company's business (Thornberry, 2001). The increased involvement of the company's workers in the business can lead to more opinions, which may increase innovation activities in the company. However, there are also consequences, such as higher employee management costs and the risk of confidential information escaping, as there may be more people accessing potentially confidential information about the company (Barringer and Bluedorn, 1999).

Corporate Venturing is the concept used to describe the creation of a new business within a company that distinguishes itself from the current business, resulting this new business from the creation of a new product, an innovation or a new market opportunity. This can lead to the formation of new divisions within the company and a change in its strategy (Narayanan et al., 2009; Sharma and Chrisman, 1999; Thornberry, 2001; Vanhaverbeke and Peeters, 2005). Companies may incur in internal or external Corporate Venturing. In the first case, the new business is formed within the domain of the company, while in the second case there are investments by the company in businesses external to its domain, either to develop or to create them (Covin and Miles, 2007; Sharma and Chrisman, 1999). This type of corporate entrepreneurship is positive for companies since it contributes to their heterogeneity (by increasing their business portfolio), to their competitive advantage and to the development of skills (Narayanan et al., 2009; Vanhaverbeke and Peeters, 2005). However, it can be difficult for companies to change their organizational structure and processes, which is essential when they incur in Corporate Venturing (Narayanan et al., 2009).

Corporate entrepreneurship through the Organizational Transformation of a company results from an innovation (e.g. product innovation) or a reorganization of resources, which leads to the development of new business opportunities (Covin and Miles, 1999; Dougherty, 1992; Thornberry, 2001). It is possible to find this type of transformation associated with other concepts in the literature such as the terms strategic renewal (Guth and Ginsberg, 1990;

Sharma and Chrisman, 1999), organizational renewal (Dougherty, 1992) or organizational rejuvenation (Covin and Miles, 1999), since it consists in the renewal of an existing company through a significant change in its strategy (e.g. marketing strategy), structure, processes and resources combination. What happens is therefore a transformation of the company and not the formation of a new business (Covin and Miles, 1999; Guth and Ginsberg, 1990; Sharma and Chrisman, 1999). This will lead to the creation of wealth and economic value for the company, to the increase of its competitiveness, and possibly to the creation of value for its customers (Covin and Miles, 1999; Guth and Ginsberg, 1990; Thornberry, 2001).

In the studies of Guth and Ginsberg (1990) and Sharma and Chrisman (1999), the concept of strategic renewal is associated with a transformation of the company and its businesses, but the concept defined by Covin and Miles (1999) bares more similarities to what Thornberry (2001) calls Industry Rule-Bending – the company's change of the competition rules in the market in which it operates. Covin and Miles (1999) describe the strategic renewal as the situation in which the company considerably changes its form of competition and its business strategy to influence its position in the market and its relation with competing companies, and to better exploit market opportunities.

Thornberry (2001) presents these types of corporate entrepreneurship, but it is possible to find others in the literature. For instance, Covin and Miles (1999) describe the following types: sustained regeneration, organizational rejuvenation, strategic renewal and domain redefinition. Organizational rejuvenation has already been mentioned in the topic of Organizational Transformation and strategic renewal in the topic of Industry Rule-Bending, so only sustained regeneration and domain redefinition will be further briefly introduced.

Sustained regeneration refers to the companies "that regularly and continuously introduce new products and services or enter new markets" (Covin and Miles, 1999: 51), actively innovating to exploit market opportunities and, therefore, engaging in strong entrepreneurial activity. The firm must resort to its technical knowledge to be able to introduce new products/services or its current products/services in new markets, which might lead to a new business and to an increase of the firm's competitive advantage, allowing the company to react to products short life cycles or to rapid technological changes (Dess et al., 2003; Kantur, 2016; Kuratko and Audretsch, 2009). Domain redefinition, on its turn, results from an innovation not only at the firm level, but at the market level as well. In this case, the company "creates a new product-market arena that others have not recognized or actively sought to exploit" (Covin and Miles, 1999: 54). This way, the company creates a new business in a market space that has not been exploited, achieving a first mover advantage and, thus, gaining a competitive advantage against the later entrants (Covin and Miles, 1999; Kuratko and Audretsch, 2009).

3. Entrepreneurial Process

The entrepreneurial process consists of "the methods, practices, and decision-making styles managers use to act entrepreneurially" (Lumpkin and Dess, 1996: 136). Several authors present the process as the identification and exploitation of an opportunity (Ardichvili et al., 2003; Choi and Shepherd, 2004; Stevenson and Gumpert, 1985). Thus, the entrepreneurial

process is a dynamic process of recognition and development of an opportunity, in which there must be continuous evaluation and permanent search for new opportunities (Ardichvili et al., 2003). The stages that constitute the entrepreneurial process differ among different authors, and it is not possible to identify a single process. Cardon et al. (2005), for example, differ from other authors in describing the entrepreneurial process through a metaphor with the paternity process, defining the stages of conception, gestation, infancy and toddlerhood, childhood and adolescence, and finally maturity, as the company is formed and develops. However, it is possible to find some similarities in the stages of these processes, which will be presented in the perspective of a new business development. In Table 2 we present the stages of the entrepreneurial process as described by several authors.

The initial stage is closely linked to the concept of market. Stevenson and Gumpert (1985) describe the first stage of the process as the identification of the business opportunity, in which the entrepreneur must have a market orientation to be able to identify the opportunity. It is at this stage that the idea of the business to be developed arises, whether it is a new idea or a new application of old ideas (Haber and Reichel, 2007; Stevenson and Gumpert, 1985). Also, for Ardichvili et al. (2003), the first stage of the process is related to the identification of the opportunity, in order to identify a market need that can be suppressed with a new combination of resources. The authors consider this initial stage the definition of the business concept through the identification of the market need, the definition of the desired benefits and the establishment of how the resources will be used, that is, the concept must include how the entrepreneur intends to supply the market need and how the resources will be applied for this purpose.

Brockner et al. (2004) consider that the entrepreneurial process begins with the identification and screening of an idea. In order to evaluate the idea, the company must carry out an analysis, questioning several factors, such as whether there is a market for the product/service it intends to offer, whether it has the capacity to supply it to the market and if it has any competitive advantage over companies already established in the market. In addition to these issues, the company should also regard other factors such as the risks to which will be subjected to and how to manage them. Finally, the company must consider whether all the investment needed for the project will be offset by the returns. These are questions that must be considered at the beginning of the process to help the company understand if it is possible to put the idea into practice but also throughout the process to reflect whether to give up at some point in the process if the company comes to the conclusion that the returns of the business will not be enough to offset the investment.

Table 2: Stages of the entrepreneurial process described by different authors

Entrepreneurial Process					
Author	Stage				
Stevenson and Gumpert (1985)	1. Identify opportunity	2. Identify how to capitalize on opportunity	3. Identify required resources	4. Determine how to control resources	5. Determine organizational structure
Ardichvili et al. (2003)	1. Define business concept	2. Define business model	3. Define business plan		
Brockner et al. (2004)	1. Identifying and screening the idea	2. Procuring resources	3. Proving the business model	4. Rollout phase	5. Product life cycle
Cardon et al. (2005)	1. Conception (commitment to the venture)	2. Gestation (resource acquisition)	3. Infancy and toddlerhood (high dependence of the venture from the entrepreneur)	4. Childhood and adolescence (increasing independence of the venture from the entrepreneur)	5. Maturity (separation of firm and entrepreneur)
Haber and Reichel (2007)	1. Idea formulation	2. Feasibility study of the business concept	3. Establishment phase (writing business plan)	4. Operation stage (initiation of the life cycle of the venture)	

The remaining stages of the process are more divergent across authors, but some point out that the second stage is the moment to identify and/or obtain the necessary resources (Ardichvili et al., 2003; Brockner et al., 2004; Cardon et al., 2005), in particular financial, technological and human resources. Investors may be needed to obtain the financial resources and, as such, the company must be able to prove the idea will be beneficial to them (Brockner et al., 2004). Ardichvili et al. (2003) consider that in the second stage of the process the company must determine what resources will use and how, through the definition of the business model. This model should include a financial model, which explains the value that the development of the opportunity will create and how it will be distributed to the stakeholders. This should include the more detailed business concept as well (Ardichvili et al., 2003), which feasibility should be analyzed at this second stage of the process, according to Haber and Reichel (2007).

For Stevenson and Gumpert (1985), resources are only related to the stages 3 and 4 of the process. Thus, before evaluating the resources, the company must identify how to capitalize on the entrepreneurial idea (stage 2), identifying the circumstances that can make the idea profitable. It is only then that the necessary resources are identified (stage 3) and how to control them is determined (stage 4), the fundamental being not the amount of resources that are applied in the project, but the innovation of the company in the use of these resources. The entrepreneurial process ends with the determination of the organizational structure (stage 5).

In the third stage of the process, according to Brockner et al. (2004), the company has to be able to demonstrate the feasibility of the business model, developing a prototype of the product/service and testing it with potential clients. For Ardichvili et al. (2003), the third (and last) stage is the time for the company to define a business plan, which should include the estimation of the expected cash flows, the description of the opportunity development activities, and the resources needed to develop it. Also for Haber and Reichel (2007) it is in the third stage (which they refer to as establishment stage) that the business plan is developed with the purposes of analyzing the financial viability of the business and obtaining external financing. The fourth stage of the process, according to Brockner et al. (2004), is the rollout phase, which entails the business/product launch. If the launch is successful, the final stage begins, the business/product life cycle – maturity and renewal/ growth or maturity and decline (Brockner et al., 2004), which Haber and Reichel (2007) call the operation stage.

4. The Business Plan and its Goals

Since the business plan is an important component and support of the entrepreneurial process (Ardichvili et al., 2003; Haber and Reichel, 2007; Honig, 2004) it should be better analyzed. The business plan is a formal document, which describes and develops the opportunity of a business identified by the entrepreneur and the strategy defined to explore it, and is designed to improve the company's performance in the market (Chwolka and Raith, 2012; Gruber, 2007; Honig, 2004). In addition to comprise a document that allows analyzing the viability of the business, the business plan also consists of a project constituted by the strategy that the company must follow to develop the new business (Fernández-Guerrero et al., 2012). Essentially, the business plan evaluates the current situation of the company and presents the company's vision for the future, through the prediction of the expected situation for the company in the future after the development of the business (Delmar and Shane, 2004; Honig, 2004).

The business plan is designed with the purposes of defining the business concept and developing the ideas about the new business (Gruber, 2007). Planning is a support in the entrepreneurial process by enabling companies to make decisions regarding the various steps to be taken in this process, including the fundamental decision on whether they should actually enter the market, thus contributing to their survival (Chwolka and Raith, 2012). The plan also has the fundamental goal of gaining financing for the development of the business which is sometimes the only reason why entrepreneurs decide to make a plan (Bewayo, 2010; Bianchi et al., 1998; Fernández-Guerrero et al., 2012), since banks and investors typically require a business plan before investing in businesses (Honig, 2004) as it gives them a perspective of the entrepreneur's ideas allowing them to assess its potential and to assess if the expected revenues are consistent with the actions planned for the business development. This financing will be an important contribution from the business plan to the business survival but, if entrepreneurs have many resources they may choose not to write a business plan since external financing will not be so important to the business startup (Burke et al., 2010; Castrogiovanni, 1996; Delmar and Shane, 2004).

Another goal of some companies when designing a business plan, which departs from the main purposes of the plan but may be equally important, is to gain legitimacy and credibility for their business. Sometimes, in this case, entrepreneurs do not really mean to use the business plan as a support for the new business (Delmar and Shane, 2004; Honig, 2004; Ivanisevic et al., 2016; Karlsson and Honig, 2009), but because they think it is something supposed to be done when starting a business and that makes their business show potential for success (Castrogiovanni, 1996). However, while legitimizing the business might seem not that fundamental to its start, it can be important to gather the support of stakeholders and to obtain resources, which in turn are fundamental to the business success (Delmar and Shane, 2004).

In terms of the contents and structure of the business plan, these diverge from plan to plan, companies do not follow a single model (Ivanisevic et al., 2016). However, there are certain topics accepted as key elements of a plan. The plan should include among its contents the description of the product or service, the definition of the business goals, the identification of the steps necessary to achieve these goals, and a financial projection of the business. It should also have a delineation of the company's strategies (organizational and financial), the expected results of these strategies and possible corrective measures in case the expected results are not met. Thus, the plan consists of a set of commercial, financial, statistical and economic information that allows the entrepreneur to understand the system where the company will act, the restrictions that will be subject to and the resources available, and thus determine how these factors will affect it (Bianchi et al., 1998; Bracker and Pearson, 1986; Brinckmann et al., 2010; Honig, 2004).

With the gathered information, it becomes possible for the entrepreneur to plan the fundamental factors of a business, namely to predict production and to establish marketing and management methods (Honig, 2004). Since the business plan is a support in the preparation of a new business, it should also provide information about potential customers, the market in which the product is to be offered and the company's competition (Brinckmann et al., 2010; Honig, 2004). Finally, another common element in business plans is the SWOT (Strengths, Weaknesses, Opportunities, Threats) analysis (Bracker and Pearson, 1986), which allows to analyze the new business both at the company level (strengths and weaknesses) and the external level (opportunities and threats).

Bracker and Pearson (1986) defined some types of business plans according to their structures and content, distinguishing between structured plans (written plans), intuitive plans (plans that are only in the mind of the entrepreneur) and unstructured plans (when there is no structured planning in the firm). They also divided the structured plans into strategic plans (long-range plans) and operational plans (short-range plans). However, the authors concluded that the planning process has more influence on the company's performance than the business plan itself.

5. The Planning Process

It is not only the business plan that affects business performance but also the planning process itself. That is: the positive contribution of the business plan to companies stems

not only from the business plan itself, but also from the whole process of defining the plan (Brinckmann et al., 2010). The planning process is the process of researching and collecting information that is fundamental to the business and its consequent analysis (Honig, 2004). This process will start when the entrepreneur identifies an opportunity, making it necessary to verify if this is in fact an idea that could benefit the company and what is the best strategy for the idea to be of value to the company (McGrath, 2010). For this process it will be fundamental that the entrepreneur has a developed planning capacity, since the more time is spent in the formulation of the business plan, the less likely it is to achieve the goals of the plan and, thus, the chances of business success will decrease (Gruber, 2007; Van de Ven et al., 1984). The benefits of planning depend from the activities developed in the planning and the time invested in the planning process as well, therefore, the entrepreneur must be able to choose what is worth of time investment and focus only on the fundamental activities (Gruber, 2007).

The process of developing the business plan should be gradual, beginning with simple business planning activities (e.g. meetings and market analysis) that enable the entrepreneur to acquire some knowledge. As the plan is developed, the entrepreneur and the company gain experience and acquire more knowledge and as such can increase the investment in planning activities, applying more and more resources to planning as the process unfolds. This type of activity must occur simultaneously with other activities associated with the development of the business (Brinckmann et al., 2010). However, it is not always beneficial that business development activities occur simultaneously with the preparation of the business plan. In the case of marketing activities, it may be more beneficial for companies to develop them only after the business plan is complete. For example, the business plan can help define the target customers of the business, so it might be more beneficial for the entrepreneur to talk to potential clients only after the business plan is completed (Shane and Delmar, 2004). Finally, when the planning is completed and the business plan written, entrepreneurs should send the plan to the maximum number of people connected to the business to be analyzed, increasing the chances of business success if more people analyze it (Van de Ven et al., 1984).

Entrepreneurs are not necessarily the ones who formulate the business plan, sometimes they approach consultants to prepare them, especially in the early stages of business development (Bianchi et al., 1998). This aid can be beneficial for companies, like Van de Ven et al. (1984) concluded in their study on new firms – success was superior in companies where there was greater support from consultants. Chrisman et al. (2005) argue as well that consultant support in the planning process contributes to the performance of the business, even though they recognize that the planning process leads to the learning of entrepreneurs when they develop the plans without external help. This learning factor should not be overlooked. Learning is one of the fundamental outcomes of a business plan, since formulating the business plan will help the entrepreneur to acquire knowledge about the new business (e.g. about competition or the market) which in turn will help the entrepreneur to determine the best actions to develop the business, contributing to the business survival. Learning will also help the entrepreneur to select the correct information about the business to present to potential financers, increasing the chances of gaining the financing needed (Castrogiovanni, 1996).

Business planning is also affected by the type of companies in which it occurs, with a difference in the planning between emerging companies and companies already established in the market. In emerging companies there is a great deal of uncertainty that will affect planning, given that as companies are starting their activity, they have little knowledge and experience (Brinckmann et al., 2010; Gruber, 2007). This uncertainty can lead companies to invest in a business plan as a mean to decrease it. Uncertainty can hinder the learning gained by business planning and diminish the chances of business success (Castrogiovanni, 1996). Still, Delmar and Shane (2003) found in their study that planning is beneficial even with the uncertainty present in new firms.

On the other hand, established companies have greater knowledge and information resulting from their experience, which is reflected in a lower degree of uncertainty in the business plans. The lower degree of uncertainty in established firms leads to a more positive influence of business planning in these companies than in emerging ones. Also, the fact that there is greater concern in emerging companies to reduce losses and that there is a great shortage of information (thus planning costs can greatly outweigh the benefits) results in a limitation of the costs incurred in market analysis and, therefore, these companies may choose not to invest in the acquisition of fundamental information (Brinckmann et al., 2010; Gruber, 2007). Even when emerging companies opt to invest in a business plan, they often do not use them, as Karlsson and Honig (2009) found in a study of a sample of new firms, where they discovered a progressive departure of the business from what was originally defined in the business plan, since the main concern of concern of companies to write the business plan was to gain legitimacy.

6. Benefits and Costs of the Business Plan

The business plan may have a positive effect on the development of companies, but some costs arise from its formulation. After many studies there is still no consensus about whether the business plan is important to firm survival. While some studies have shown a positive impact of the business plan in firm performance (e.g. Brinckmann et al., 2010; Burke et al., 2010; Delmar and Shane, 2004), others have shown no impact (e.g. Honig and Karlsson, 2004; Lange et al., 2007; Tornikoski and Newbert, 2007). Some authors oppose to business planning because it is time consuming and deviates the focus from activities that have a more direct contribution to the beginning of the new business (Karlsson and Honig, 2009; Shane and Delmar, 2004). According to Karlsson and Honig (2009), the fact that there are successful companies that did not develop business plans at the beginning of their activity can be given as evidence against business planning, yet, the business plan is still seen as an important support for business development.

While the business plan is considered positive to firm performance because it is a support for the new business, Chwolka and Raith (2012) argue that in fact business planning is beneficial to companies because it will contribute to a better analysis of business ideas, so the chances that bad business ideas stay out of the market will increase, and this will lead to only good ideas ever reaching the market. This way, the chances of business survival will be greater.

The benefits of the business plan will result from the predictions that can be made through the formulation of the plan and from the business planning itself that will help the company to choose the most beneficial approach for its business (Castrogiovanni, 1996; Chwolka and Raith, 2012). An example of a fundamental forecast is the expected cash flows. Their prediction and analysis allow the entrepreneur to understand whether it will in fact be beneficial to enter the market (Chwolka and Raith, 2012). The business plan also contributes to the acquisition of resources through financing (Burke et al., 2010) and to the economic use of the company's resources (Brinckmann et al., 2010). It may also improve the company's financial performance; however, this obviously depends on the fulfillment of the plan. Sometimes, as mentioned earlier, entrepreneurs draw up plans without having the intention of consulting and following them, so it is necessary to control the compliance of the business with what is established in the plan, for the plan to have a positive effect in the financial performance (Karlsson and Honig, 2009).

Essentially, the benefits of the business plan derive from the information it provides, since its elaboration eases the collection and management of key business startup information (Gruber, 2007; Shane and Delmar, 2004). This collection of information allows to define the business concept and to better understand the market in which the company will operate, and helps in the development of marketing activities, which results in the definition of the necessary steps for the commercialization of the new product or service, and, therefore, allows the business launch (Brinckmann et al., 2010; Gruber, 2007; Shane and Delmar, 2004). This information will be important not only for the entrepreneur but to everyone involved in the business development as well. Through the business plan the entrepreneur will convey the information to the company and to stakeholders, allowing them to understand the entrepreneur's vision for the business (Delmar and Shane, 2003).

In terms of costs, a relevant cost arising from the business plan is the opportunity cost that results from the time spent in its preparation (Chwolka and Raith, 2012; Shane and Delmar, 2004). The time devoted to the business plan could be used in other activities with a more direct effect on the origin of the business, such as marketing activities, rather than being spent on an activity that does not directly result in the formation of the business (Shane and Delmar, 2004). This time spent will be reflected in a delay in the development of the business, which according to Chwolka and Raith (2012) has two consequences – a possible loss of the present value of future revenues (related to interest expenses) and a possible reduction of the revenues, since market conditions might change and new competition may arise. However, while it is true that the company may be investing time that could be useful in other activities, business planning entails a planning of activities that will result in the saving of time when the business starts its development, since the entrepreneur will only focus on the necessary activities to achieve the business goals. Also, the information collected to plan the business will allow a faster decision making in the business development (Delmar and Shane, 2003). This shows indeed that benefits and costs of the business plan are opposite to each other – as summarized in Table 3 – further dividing the debate between supporters and opponents of the business plan.

Table 3: Opposed benefits and costs of the business plan

Business Plan	
Benefits	Costs
- Planning of activities will result in the saving of time when the business starts its development. - Planning allows faster decision making when the business takes off (Delmar and Shane, 2003).	- Time devoted to the business plan could be used in other activities with a more direct effect on the origin of the business (leads to opportunity cost) (Shane and Delmar, 2004).
- Business planning can improve the adaptability of the business (Castrogiovanni, 1996). - Business plan improves decision making when it is necessary to improvise (Burke et al., 2010).	- The business plan may make it difficult for companies to adapt to new conditions (Honig, 2004).
- Business planning contributes to the economic use of the company's resources (Brinckmann et al., 2010). - The business plan contributes to the acquisition of resources through financing (Burke et al., 2010).	- The writing of the business plan requires the spending of company's resources (Karlsson and Honig, 2009).

Other costs include expenses with consultants who support the preparation of the plan, the effort required by the planning activities, and the spending of company's resources (such as financial resources) that could be applied to other activities that would actually start the business, such as searching for customers and suppliers (Chwolka and Raith, 2012; Karlsson and Honig, 2009). The plan may also make it difficult for companies to adapt to new conditions if they are in dynamic markets where product changes are frequent (Honig, 2004). However, according to Castrogiovanni (1996), business planning can actually improve the adaptability of the business, since the learning gained from it can help the entrepreneur to understand how to adapt to certain situations before they occur.

Given that the business plan has costs and benefits, the entrepreneur must decide if the business plan will be useful, according to the information available. If the entrepreneur does not have the necessary knowledge and experience the business plan will be an important support to the new business; on the other hand, if the entrepreneur and the company have extensive experience in business development and extensive knowledge about the new business and market, the business plan might not be so relevant (Burke et al., 2010). However, as discussed before, even in established firms the business plan can be an important support given that these companies have less uncertainty and can benefit more from the information it provides (Brinckmann et al., 2010). As such the business plan can be beneficial both when the entrepreneur has great or little knowledge.

As discussed above, studies usually find a positive correlation between the business plan and the firm's performance or no correlation. This means that at most the business plan has no effect on firm's performance, which shows that usually the costs of the business plan will not exceed the benefits.

7. Conclusion

The purpose of this paper was to explore entrepreneurship at the firm level – corporate entrepreneurship, and to determine how an entrepreneur within a firm should proceed in the exploitation of an opportunity. The stages of the entrepreneurial process might be different among authors, but it was possible to find some similarities, essentially the need to first analyze the business and to eventually assemble the necessary resources. Some authors mention the stage of writing the business plan as well, but there is no unanimity about whether this is really a fundamental tool in the business development. Some studies have demonstrated that the business plan has a positive effect on companies' performance, helping them to thrive in the market and succeed (Brinckmann et al., 2010), while others oppose by claiming that the time spent in its formulation should be spent in activities that would result in a direct influence in the start of the business (Karlsson and Honig, 2009).

The fact that some benefits and costs of business planning oppose each other or even contradict each other can further contribute to this divide. But while writing a business plan can lead to resource spending (Karlsson and Honig, 2009), at the same time it helps the company gain financing, increasing its resources (Burke et al., 2010). Thus, business planning may be like an investment – it is necessary to spend resources to gain more resources. The choice of the company to plan will ultimately depend on its evaluation of whether the financing obtained will surpass the resources spent. This applies to all benefits and costs of business planning: to engage in business planning the benefits of planning must surpass the costs (Chwolka and Raith, 2012), which means that the value of the business plan must be determined by the entrepreneur before the business development to ascertain whether it will benefit the company or not. Overall, the literature favors to the conclusion that the benefits of the business plan tend to exceed the costs.

References

Ardichvili, A.; Cardozo, R.; Ray, S. (2003) A theory of entrepreneurial opportunity identification and development, *Journal of Business Venturing*, 18(1), 105-123.

Audretsch, D.; Bönte, W.; Keilbach, M. (2008) Entrepreneurship capital and its impact on knowledge diffusion and economic performance, *Journal of Business Venturing*, 23(6), 687-698.

Audretsch, D.; Keilbach, M. (2004) Entrepreneurship capital and economic performance, *Regional Studies*, 38(8), 949-959.

Audretsch, D.; Thurik, R. (2004) A Model of the Entrepreneurial Economy, *International Journal of Entrepreneurship Education*, 2(2), 143-166.

Barringer, B. R.; Bluedorn, A. C. (1999) The relationship between corporate entrepreneurship and strategic management, *Strategic Management Journal*, 20(5), 421-444.

Bewayo, E. D. (2010) Pre-start-up preparations: Why the business plan isn't always written, *Entrepreneurial Executive,* 15, 9-23.

Bianchi, C.; Winch, G.; Grey, C. (1998) The business plan as a learning-orientated tool for small/medium enterprises: A business simulation approach, In *International System Dynamics Conference*, 1-20.

Bracker, J. S.; Pearson, J. N. (1986) Planning and financial performance of small, mature firms, *Strategic Management Journal*, 7(6), 503-522.

Brinckmann, J.; Grichnik, D.; Kapsa, D. (2010) Should entrepreneurs plan or just storm the castle? A meta-analysis on contextual factors impacting the business planning-performance relationship in small firms, *Journal of Business Venturing*, 25(1), 24-40.

Brockner, J.; Higgins, E. T.; Low, M. B. (2004) Regulatory focus theory and the entrepreneurial process, *Journal of Business Venturing*, 19(2), 203-220.

Burke, A.; Fraser, S.; Greene, F. J. (2010) The multiple effects of business planning on new venture performance, *Journal of Management Studies,* 47(3), 391-415.

Cardon, M. S.; Zietsma, C.; Saparito, P.; Matherne, B. P.; Davis, C. (2005) A tale of passion: New insights into entrepreneurship from a parenthood metaphor, *Journal of Business Venturing,* 20(1), 23-45.

Castaño, M. S.; Méndez, M. T.; Galindo, M. Á. (2016) The effect of public policies on entrepreneurial activity and economic growth, *Journal of Business Research*, 69(11), 5280-5285.

Castrogiovanni, G. J. (1996) Pre-startup planning and the survival of new small businesses: Theoretical linkages, *Journal of Management*, 22(6), 801-822.

Choi, Y. R.; Shepherd, D. A. (2004) Entrepreneurs' decisions to exploit opportunities, *Journal of Management*, 30(3), 377-395.

Chrisman, J. J.; McMullan, E.; Hall, J. (2005) The influence of guided preparation on the long-term performance of new ventures, *Journal of Business Venturing*, 20(6), 769-791.

Chwolka, A.; Raith, M. G. (2012) The value of business planning before start-up: A decision-theoretical perspective, *Journal of Business Venturing*, 27(3), 385-399.

Covin, J. G.; Miles, M. P. (1999) Corporate entrepreneurship and the pursuit of competitive advantage, *Entrepreneurship Theory and Practice,* 23(3), 47-63.

Covin, J. G.; Miles, M. P. (2007) Strategic use of corporate venturing, *Entrepreneurship Theory and Practice,* 31(2), 183-207.

Cumming, D.; Johan, S.; Zhang, M. (2014) The Economic Impact of Entrepreneurship: Comparing International Datasets, *Corporate Governance: An International Review*, 22(2), 162-178.

Delmar, F.; Shane, S. (2003) Does business planning facilitate the development of new ventures?, *Strategic Management Journal,* 24(12), 1165-1185.

Delmar, F.; Shane, S. (2004) Legitimating first: Organizing activities and the survival of new ventures, *Journal of Business Venturing*, 19(3), 385-410.

Dess, G. G.; Ireland, R. D.; Zahra, S. A.; Floyd, S. W.; Janney, J. J.; Lane, P. J. (2003) Emerging issues in corporate entrepreneurship, *Journal of Management*, 29(3), 351-378.

Dougherty, D. (1992) A practice-centered model of organizational renewal through product innovation, *Strategic Management Journal,* 13, 77-92.

Fernández-Guerrero, R.; Revuelto-Taboada, L.; Simón-Moya, V. (2012) The business plan as a project: An evaluation of its predictive capability for business success, *Service Industries Journal*, 32(15), 2399--2420.

Freytag, A.; Thurik, R. (2007) Entrepreneurship and its determinants in a cross-country setting, *Journal of Evolutionary Economics*, 17(2), 117-131.

Galindo, M.-Á.; Méndez, M. T. (2014) Entrepreneurship, economic growth, and innovation: Are feedback effects at work?, *Journal of Business Research*, 67(5), 825-829.

Gruber, M. (2007) Uncovering the value of planning in new venture creation: A process and contingency perspective, *Journal of Business Venturing*, 22(6), 782-807.

Guth, W. D.; Ginsberg, A. (1990) Guest editors' introduction: Corporate entrepreneurship, *Strategic Management Journal*, 11(5), 5-15.

Haber, S.; Reichel, A. (2007) The cumulative nature of the entrepreneurial process: The contribution of human capital, planning and environment resources to small venture performance, *Journal of Business Venturing*, 22(1), 119-145.

Honig, B. (2004) Entrepreneurship education: Toward a model of contingency based-business planning, *Academy of Management Learning and Education*, 3(3), 258-273.

Honig, B.; Karlsson, T. (2004) Institutitonal forces and the written business plan, *Journal of Management*, 30(1), 29-48.

Ivanisevic, A.; Katic, I.; Buchmeister, B.; Leber, M. (2016) Business plan feedback for cost effective business processes, *Advances in Production Engineering & Management*, 11(3), 173-182.

Kantur, D. (2016) Strategic entrepreneurship: Mediating the entrepreneurial orientation-performance link, *Management Decision*, 54(1), 24-43.

Karlsson, T.; Honig, B. (2009) Judging a business by its cover: An institutional perspective on new ventures and the business plan, *Journal of Business Venturing*, 24(1), 27-45.

Kuratko, D. F.; Audretsch, D. B. (2009) Strategic entrepreneurship: Exploring different perspectives of an emerging concept, *Entrepreneurship Theory and Practice*, 33(1), 1-17.

Lange, J. E.; Mollov, A.; Pearlmutter, M.; Singh, S.; Bygrave, W. D. (2007) Pre-start-up formal business plans and post-start-up performance: A study of 116 new ventures, *Venture Capital*, 9(4), 237-256.

Lumpkin, G. T.; Dess, G. G. (1996) Clarifying the entrepreneurial orientation construct and linking it to performance, *Academy of Management Review*, 21(1), 135-172.

McGrath, R. G. (2010) Business models: A discovery driven approach, *Long Range Planning*, 43(2–3), 247-261.

Miller, D. (1983) The correlates of entrepreneurship in three types of firms, *Management Science*, 29(7), 770-791.

Minniti, M.; Lévesque, M. (2010) Entrepreneurial types and economic growth, *Journal of Business Venturing*, 25(3), 305-314.

Narayanan, V. K.; Yang, Y.; Zahra, S. A. (2009) Corporate venturing and value creation: A review and proposed framework, *Research Policy*, 38(1), 58-76.

Russell, R. D. (1999) Developing a process model of intrapreneurial systems: A cognitive mapping approach, *Entrepreneurship Theory and Practice*, 23(3), 65-84.

Shane, S.; Delmar, F. (2004) Planning for the market: Business planning before marketing and the continuation of organizing efforts, *Journal of Business Venturing*, 19(6), 767-785.

Shane, S.; Venkataraman, S. (2000) The promise of enterpreneurship as a field of research, *The Academy of Management Review*, 25(1), 217-226.

Sharma, P.; Chrisman, J. J. (1999) Toward a reconciliation of the definitional issues in the field of corporate entrepreneurship, *Entrepreneurship Theory and Practice*, 23(3), 11-27.

Stevenson, H. H.; Gumpert, D. E. (1985) The heart of entrepreneurship, *Harvard Business Review*, 63(2), 85-94.

Thornberry, N. (2001) Corporate entreprenuership: Antidote or oxymoron?, *European Management Journal*, 19(5), 526-533.

Tornikoski, E. T.; Newbert, S. L. (2007) Exploring the determinants of organizational emergence: A legitimacy perspective, *Journal of Business Venturing*, 22(2), 311-335.

Van de Ven, A. H.; Hudson, R.; Schroeder, D. M. (1984) Designing new business startups: Entrepreneurial, organizational, and ecological considerations, *Journal of Management*, 10(1), 87-107.

Vanhaverbeke, W.; Peeters, N. (2005) Embracing innovation as strategy: Corporate venturing, competence building and corporate strategy making, *Creativity and Innovation Management*, 14(3), 246-257.

Wolcott, R. C.; Lippitz, M. J. (2007) The four models of corporate entrepreneurship, *MIT Sloan Management Review*, 49(1), 75-82.

IMPRENSA DA UNIVERSIDADE DE COIMBRA
COIMBRA UNIVERSITY PRESS

· U C ·