Investment choices, intellectual property and workers’ rights

Marco Savioli

February 2014

Joint Ph.D. Program in Economics – University of Tuscany
Department of Economics and Statistics – University of Siena
Dissertation Committee

Prof. Antonio Nicita

Prof. Roberto Renò

Prof. Giulio Zanella
Acknowledgements

I want to thank my very open-minded supervisor, Ugo Pagano. I am also very grateful to Massimiliano Castellani, Martina Cioni, Luca Fanelli and Pierpaolo Pattitoni for helping me in my research journey.

To Simona, Rossella, Giovanni, Manuela and Zoe.
Abstract

Unions and government relations influence markets and determine the degree of marketability of all kinds of asset. This thesis explores four aspects of this complex relationship. In Part I about portfolio theory, suboptimal solutions due to different levels of marketability of assets and entrepreneurs’ perceptions are investigated. In Part II, where intangible assets, imitation and social relations are considered, non-marketability of assets is further characterized. In Part III about work safety, this extreme version of non-marketability is explored. Finally, in Part IV, unionized economies’ performance is studied in terms of interaction between union and government strategies.

Several investment decisions deal with non-marketable assets, available only to one investor and often indivisible. Under these assumptions, taken into consideration by Part I, it is possible to define a new set of conditions under which non-marketable asset represents a good investment, and the efficient frontier may exhibit non-linearities and intervals of discontinuity. Furthermore, Part I measures the excess of weight on the company in investment portfolios of entrepreneurs perceiving their private company to have a lower risk (overconfidence) and a higher return (over optimism) than real.

The importance of intangible assets has greatly increased in last decades. The choice of developing countries (the South) to accept the TRIPs agreement and to stop imitation processes is modeled, in Part II, as the opportunity to participate in the official international market, represented by a multinational. As the South increases its share in the multinational profits, it would also promote to a greater extent the use of IPR. Moreover, the level of participation in the multinational profits needed to support the TRIPs agreement is positively related to the relative value of the quality level achieved by the imitated products with respect to the quality level achieved by the official products. Part II continues with further results concerning social relations and intangible assets. Wellbeing is pursued through individual and social action. Social capital is not always positive, but usually it can solve and reduce both government and market failures. The role of the policy maker is to build positive social capital and exploit it to reach her objectives. An enrichment of the Coase’s theory of the firm explores the links between incomplete contracts, secrecy, IPR and firm dimension. Furthermore, in order
Abstract
to evaluate the overall effect brought by the introduction of IPR on welfare it is necessary
to take in account the problem of asset specificity. The decentralized solution seems to
bring to an excess of the reproduction cost (IPR).

Part III is about work safety, an extreme version of non-marketability. The policy
maker should be aware that among the most significant determinants of Italian accidents
and illnesses occurring at work are bad working conditions. On the contrary, having a
fixed-term contract is not significant.

Finally, in Part IV, a novel policy-game model with perfect information analyzes the
simultaneous interaction between the government and the labor union in a unionized
economy. The model is tested using Italian quarterly data on the period 1960-2009, in
particular fiscal policy (government budget surplus) and strikes (amount of hours not
worked). The empirical results show the existence of at least two cointegrating relationships (the player’s reaction functions), giving support to the Cournot-Nash equilibrium
predicted by the theoretical model. A phase diagram represents the estimated structure
of dynamic adjustment towards equilibrium for Italy in a unified and original way.
## Contents

Dissertation Committee .......................................................... i  
Acknowledgements ................................................................... ii  
Abstract .................................................................................. iii  
1. Introduction ........................................................................ 1  

### I. Investment choices .......................................................... 5  

2. Indivisible non-marketable assets and suboptimal solutions ........ 6  
2.1. Introduction .................................................................... 6  
2.2. Four cases for a model of investment choices ....................... 9  
2.2.1. N risky assets and a risk-free asset (case 1) ..................... 10  
2.2.2. N risky assets, a risk-free asset, and a non-marketable investment 12  
2.3. Solutions to the four cases: a comparison .............................. 16  
2.3.1. Standard deviations .................................................... 16  
2.3.2. Intersection points ..................................................... 20  
2.4. The efficient frontier for different real scenarios and the choice of the investor ............................................. 21  
2.4.1. Four possible scenarios .............................................. 21  
2.4.2. The choice of the investor: entrepreneurs vs. clerks .......... 24  
2.5. Conclusions .................................................................... 25  
2.6. Appendix A ...................................................................... 27  
2.7. Appendix B ...................................................................... 27  

3. Entrepreneurial under-diversification: over optimism and overconfidence .......................................................... 29  
3.1. Introduction .................................................................... 29  
3.2. Theoretical setup ............................................................. 30  
3.2.1. Overconfidence ......................................................... 30  
3.2.2. Over optimism .......................................................... 34
## Contents

3.3. Implicit overconfidence and over optimism levels .......................... 38  
3.4. Conclusions .............................................................................. 41

II. Intangible assets ......................................................................... 42

4. Intellectual property rights and imitation in developing countries ..... 43  
4.1. Introduction .............................................................................. 43  
4.2. Related literature ..................................................................... 45  
4.3. The model ................................................................................ 47  
4.3.1. Company’s and imitated products in S .................................. 47  
4.3.2. The TRIPs agreement ............................................................ 49  
4.3.3. What is better for the South .................................................. 50  
4.4. Empirical validation .................................................................. 52  
4.4.1. Data description ..................................................................... 53  
4.4.2. Empirical results ................................................................... 55  
4.5. Conclusion ................................................................................ 58  
4.6. Appendix: Data ......................................................................... 59

5. Social capital, ideas and relation specific investments .................. 61  
5.1. Social capital, institutions and policy making .............................. 62  
5.1.1. Introduction .......................................................................... 62  
5.1.2. Social capital dimensions ...................................................... 63  
5.1.3. The policy maker ................................................................. 69  
5.1.4. Concluding remarks ............................................................. 73  
5.2. The concept of ideas inside the theory of the firm ...................... 75  
5.2.1. Introduction .......................................................................... 75  
5.2.2. The model ............................................................................ 76  
5.2.3. Concluding remarks ............................................................. 79  
5.3. Relation specific investments to knowledge ............................... 82  
5.3.1. Relation specific investments to private goods ...................... 82  
5.3.2. Relation specific investments to knowledge ......................... 83  
5.3.3. The model ............................................................................ 84
III. Workplace safety  91

6. Safety at the workplace: accidents and illnesses  92
   6.1. Introduction .............................. 92
   6.2. Data ................................... 95
   6.3. Econometric specification .................. 98
   6.4. Results ................................ 100
      6.4.1. The probability of accidents and illnesses at the workplace ..... 100
      6.4.2. The joint probability of accidents and illnesses at the workplace . 107
      6.4.3. The probability of accidents and illnesses at the workplace within
              three years of tenure ........................ 109
   6.5. Conclusion ............................... 112
   6.6. Appendix: Descriptive statistics .......... 113

IV. Government and union  116

7. Government fiscal efforts vs. labour union strikes: it takes two to tango  117
   7.1. Introduction .............................. 117
   7.2. Model .................................. 121
      7.2.1. Reaction functions .................... 123
      7.2.2. Cournot-Nash equilibrium ............. 124
   7.3. Empirical analysis ....................... 125
      7.3.1. Equilibrium .......................... 128
      7.3.2. Phase diagram representation .......... 134
   7.4. Conclusions ............................. 136
   7.5. Appendix: Italian institutional configuration  .......... 137
   7.6. Appendix: Data ........................... 138
## List of Figures

2.1. The frontiers of the four cases in the space \((\mu; \sigma)\) ........................................ 11
2.2. Relevant frontier when \(\alpha = 0\) ................................................................. 23
2.3. Relevant frontier when \(0 < \alpha \leq S\sqrt{\sigma_1^2 - C}\) ........................................ 23
2.4. Relevant frontier when \(\alpha > S\sqrt{\sigma_1^2 - C}\) ............................................. 24
2.5. Indifference curves and investment choices .................................................. 25

3.1. Private company weight in case of overconfidence. ................................. 34
3.2. Frontier shift and overconfidence bias. ....................................................... 35
3.3. Duality in portfolio optimization. ................................................................. 36
3.4. Frontier shift and over optimism bias. ......................................................... 38

4.1. Overall scatterplot and lowess regression curve of IPR and IMITAB over the period 1985-2005 ................................................................. 55

5.1. IPR defense and incomplete contracts: log ............................................... 79
5.2. IPR defense and incomplete contracts: log- ............................................... 80
5.3. IPR defense and incomplete contracts: indifferent firm ............................. 80
5.4. Capacity of individuals 1, 2 and 3 ............................................................... 85
5.5. Profit and indifferent individual ................................................................. 87
5.6. Social planner solution ............................................................................. 89

6.1. Current job tenure effect on the probability of accidents at the workplace – Table 6.2 ................................................................. 104
6.2. Years of education partial effect on the probability of illnesses at the workplace – Table 6.4 ................................................................. 110

7.1. \(effort_t\) and \(strike_t\) over the period 1960.Q1-2009.Q4 ......................................... 126
7.2. Estimated error correction terms ............................................................... 133
7.3. Unified phase diagram ............................................................................. 135
7.4. Estimated phase diagram for Italy under hypotheses (i) and (ii) in the text 136
List of Tables

2.1. Efficient frontiers and scenarios .................................................. 21
3.1. Implicit overconfidence and over optimism levels; parameters from Kerins et al. [2004] ................................................................. 39
3.2. Implicit overconfidence and over optimism levels; our assumptions ...... 40
4.1. short text ...................................................................................... 54
4.2. short text ...................................................................................... 56
4.3. short text ...................................................................................... 59
5.1. R&D in Spain in 2004 ................................................................. 78
6.1. Incidence (%) of work accidents and illnesses by contract type .......... 98
6.2. Probability of accidents and illnesses at the workplace – Probit .... 101
6.3. Probability of accidents/illnesses repetition at the workplace conditional on having had at least one accident/illness – Probit ................. 105
6.4. Joint probability of accidents and illnesses at the workplace – Bivariate probit ................................................................. 107
6.5. Probability of accidents and illnesses at the workplace within three years of tenure – Probit ......................................................... 110
6.6. Descriptive statistics of variables used in the analysis ................. 114
7.1. Diagnostic tests on the estimated VAR system in Eq. (7.24) .......... 127
7.2. LR Trace cointegration rank test .................................................. 129
7.3. Estimated cointegration relationships and adjustment coefficients .... 131
7.4. Cournot-Nash equilibrium properties ......................................... 132
7.5. Descriptive statistics ................................................................. 139
1. Introduction

Politics and corporate governance co-evolve generating complex interactions of labour market and financial institutions. As a result, different forms of workers’ organization and capitalist concentration arise. The latter have important effects, for instance, on under-diversification of investments; the former have effects on unions and government relations that, with different levels of strike activity and social peace, influence markets and the degree of marketability of assets.

Belloc and Pagano [2013] find a significant positive cross-country correlation between the degree of protection of workers’ rights and the degree of corporate ownership concentration. The US and Switzerland are two examples of countries with low level of workers’ protection and ownership concentration. As stated by Roe [1994], dispersed shareholders and concentrated management became the quintessential characteristics of the large American firm. Fearing concentrated private economic power, the public upheld financial intermediary fragmentation and dispersed shareholders with managers in control of firms. The result of the politics on the way financial intermediaries move savings from household to firms were public companies with high levels of marketability.

A correct evaluation of the effects of different levels of asset marketability, like under-diversification of investments, can therefore enhance the understanding of the differences between the American management model and that of continental Europe. Chapter 2 investigates how marketability of assets and entrepreneurs’ perceptions determine sub-optimal solutions in the context of portfolio theory.

Portfolio theory deals with investment decisions and optimal allocation of investor wealth. In the traditional framework, it is generally assumed that investments are marketable, continuously divisible, and investors have an unlimited ability to gather and process information. However, reality often departs in a significant way from these assumptions. In fact, non-marketable investment opportunities often occur. Non-marketable assets are available only to one particular investor and are often indivisible, since they require a fixed amount of investment to be undertaken.

Illegal or black market transactions, real investments of a family-managed business, and human capital investments of a particular individual are all examples of situations
where the assumptions of the portfolio theory are likely to fail. From a personal perspective, these investments generally imply a “take it or leave it” choice for the individual. In all these cases, investors are generally required to commit a non-marginal amount of their wealth to the investment, and they suffer from an under-diversification condition.

Entrepreneurs do not have a systematic higher risk propensity than non-entrepreneurs. Several empirical findings, nevertheless, show that entrepreneurs often invest a large share of their personal wealth in their company, exposing themselves to idiosyncratic risk. A possible explanation for this costly exposure is based on two behavioral biases: overconfidence and over optimism. Both these biases affect the very main variables of the risk-return analysis à la Markowitz. In Chapter 3 is investigated how entrepreneurs overweight the company in their investment portfolios if they perceive their private company to have a lower risk and a higher return than real.

The importance of intangible assets has greatly increased in last decades surpassing 80% of all assets’ value. As a result, intangible assets of companies listed at the stock market have a direct and significant influence on the company (stock) value. One of the biggest values at immaterial assets are intellectual property rights (IPR) like patents or trademarks. Given the increasing importance of intangible assets, heterogeneity in the adoption of IPR is another relevant difference characterizing assets marketability of different economies. In Part II, Chapter 4 models the choice of developing countries to accept the agreement on Trade Related Aspects of Intellectual Property Rights (TRIPs) and to stop imitation processes as the opportunity to participate in the official international market, represented by a multinational.

Innovators, mostly living in rich economies, have vested interests in building IPR institutions. Indeed, innovators seek protection to recoup their investments and appropriate the returns. Since the birth of the WTO and the establishment of the TRIPs agreement, international pressure on Developing Countries (DC) to use Intellectual Property (IP) has noticeably increased. Many DC make a large use of imitation to sell products invented by others at lower prices. Imitations are of lesser quality in comparison to the original products. One argument to convince DC to accept the new setting, with stringent IP, is an increase in imports of the best products. In addition, letting aside imitative process and making use of IP, DC would participate in the official market and they could gather throughout the world the monopoly rents coming from their production of IP.

Skill, effort and competence of management are highly regarded in countries like the US with scattered shareholders [Roe, 1994]. On the contrary, in countries with a high level of workers’ protection and owner concentration, social democracy, social connections and family ties (social capital) are highly valued. Agency problems arising from
1. Introduction

the separation between ownership and control are diminished for families with concentrated ownership that strictly manage their firm. The limited diversification of risks and the poor incentives for managers are balanced against strong connections, lower agency costs, desire for control and power.

Wellbeing is pursued through individual and social action. The social dimension is therefore a decisive economic force, investigated in Chapter 5. Social capital, contributing to the cumulative capacity of individuals and social groups to work together for a common goal, evolves during long periods, or better it co-evolves with all the economic institutions and organizations in which it is embedded. Social capital is not always positive, but usually it can solve and reduce both government and market failures. This is why the policy maker should be aware of the social capital’s existence and the constraints it imposes on the changes and challenges a society can face. The role of the policy maker is to build positive social capital and exploit it to reach her objectives. Human capital investment and a correct education strategy are two viable tools that complement and magnify social capital’s effect on the economy.

Since firms balance off the gains from an internal governance structure against the gains from market trades [Coase, 1937], a low level of asset marketability brings about different kind of firms. Firms and investments’ dimension and characteristics are therefore influenced by many factors, like for instance ideas and knowledge diffusion and specificity of assets, that are explored at the end of Chapter 5.

Part III is about work safety, one of the most important elements of workers’ rights. Accidents and illnesses at the workplace may be measured in economic terms and also priced. However, there is a high level of non-marketability of work safety that has to be taken into account. To enrich the comprehension of workplace safety as a whole by offering a broad view of the phenomenon, this section studies the key determinants of workplace safety in an attempt to document the relationships between work related accidents, illnesses, and types of contracts and then personal, firm, and job characteristics, including working conditions.

The relationship between characteristics of corporate governance, such as the degree of separation between ownership and control, and strength of workers’ organization, such as unions concentration, can be explained in many ways. Belloc and Pagano [2009] point more to a complex relation of co-evolution and co-determination between the two with causality running in both directions. Roe [2003]’s perspective is instead employees’ political rights (social democracy) determining corporate governance forms.

Well established social democracies favored employees over capital, but shareholders resisted building or maintaining private and concentrated their ownership [Roe, 2003].
In continental Europe, firms’ ownership is less dispersed on the market than in the US, with blocks of shares usually owned by families or, in particular in Germany, financial institutions.

Unionized economies’ social democracy generates high pressure on firms for low-risk expansion, in order to avoid risky organizational change. Another relevant issue, taken into account in Part IV, is the interaction between union and government strategies, particularly important to understanding the economic and political performance of unionized economies. In wage bargaining, unions and governments have conflicting aims: unions want to maximize real wages and governments want to minimize unemployment. However, both agents have similar forms of political commitment: unions and governments want to maximize consensus among workers and voters. So, unions call ‘economic’ and ‘political’ strikes to maximize real wages and consensus among workers, and governments enact policies to maximize employment and consensus among voters.

Varying degrees of centralization of wage bargaining [Calmfors and Driffill, 1988] may correspond to same degrees of capital concentration, implying different level of idiosyncratic risk suffered by investors. Furthermore, to have a separation between ownership and control, assets have to be marketable. And that is not always the case. It is important, therefore, correctly evaluating portfolio risks in such a setting where investments are non-marketable. Financial institutions, correctly evaluating risk, can transfer resources to families with concentrated ownership.
Part I.

Investment choices
2. Indivisible non-marketable assets and suboptimal solutions

2.1. Introduction

One of the most relevant topics of economic and financial theory deals with investment decisions and optimal allocation of investor wealth. Theoretical models generally assume that investments are marketable, and investors are fully rational. However, these conditions rarely hold in reality.

In fact, non-marketable investment opportunities often occur. Non-marketable assets are available only to one particular investor and are often indivisible, since they require a fixed amount of investment to be undertaken. It is worth stressing that by saying “available only to one particular investor” we mean that only that investor can directly access it and that other agents can access it only through that investor. In this sense, the investment we consider is by definition out of the market and no market can exist for it.\textsuperscript{1} Furthermore, investors often suffer from bounded rationality [Simon, 1957]: they either have a limited set of information, a limited capacity to process information, or the entire optimization process is too costly. The condition of bounded rationality drives investors toward satisfactory solutions rather than optimal ones. Both these characteristics have relevant consequences on optimal choices by investors, and must be taken into account.

Illegal or black market transactions, real investments of a family-managed business, and human capital investments of a particular individual are all examples of situations where the assumptions of the portfolio theory are likely to fail. From a personal perspective, these investments generally imply a “take it or leave it” choice for the individual. In all these cases, investors are generally required to commit a non-marginal amount of their wealth to the investment, and they suffer from an under-diversification condition.

Even if the situations presented are likely to occur in reality, to the best of our knowledge a contribution that deals with these problems in a general and exhaustive way is missing. Modern portfolio theory is introduced by Markowitz [1952], and Markowitz

\textsuperscript{1}Therefore, the literature on incomplete markets is not directly applicable to this context.
2. Indivisible non-marketable assets and suboptimal solutions

[1959]. These works establish mean-variance approach as the cornerstone of modern portfolio theory. A good review of portfolio theory can be found in Elton and Gruber [1997].

Following the modern portfolio theory framework, Tobin [1958] separation theorem leads to useful implications. According to the separation theorem, each investor can satisfy her investment needs combining the best portfolio of risky assets – identical for all individuals – with a risk-free asset. The optimal allocation between the efficient portfolio of risky assets and the risk-free asset depends on the investor’s preferences. The separation theorem represents the starting point for the construction of the capital market line. In our work, however, the assumptions needed to validate this theorem are not always assured. In fact, we describe a situation where an investor has exclusive access to a non-marketable investment, which is not always infinitely divisible. In this case, the investor is likely to identify a new portfolio of risky assets that differs from that of all other individuals, and define a new personal efficient frontier which includes the non-marketable investment.

Our model takes into account the strongest form of non-marketability: i.e., the investment is completely out of the market since it cannot directly be accessed by the market. Even if non-marketability is generally associated with illiquidity, our interpretation is different. In fact, an asset is illiquid if it is not “easy” or too costly to convert it in cash. Thus, the problem of illiquid assets deal with intertemporal choices, but we do not take this aspect into account. However, in what follows we present several studies that deal with illiquidity.

Mayers [1973] and Mayers [1976] presents the problem of non-marketable assets in relation to the determination of capital asset prices. His formalization considers a portfolio which includes a position in perfectly liquid marketable assets and a position in an illiquid non-marketable asset (e.g., human capital). He shows that the weights of the market portfolio vary across investors, and derives a single-period model of capital asset pricing. Differently from Mayers’ works, we take into account the risk-free asset, we do not explicitly consider a particular utility function for the investor, and we deal with the non-marketable asset giving it an explicit weight. Furthermore, while Mayers mainly proposes a market equilibrium analysis aiming at the correct pricing of marketable and non-marketable assets, our analysis has a micro perspective focused exclusively on the correct evaluation of investment choices by individual investors.

Williams [1978] proposes a model where agents can choose a level of education, an example of non-marketable asset, and allocate their remaining wealth among risky and riskless assets. His results show that when the risk of human capital investment increases,
risk-averse agents reduce their investments in human capital. Williams’ findings stress
the importance of the recognition of the uncertainty specific to the educational process,
and show that the problems of choosing a level of education and the weights of a portfolio
of financial assets are generally inseparable.

Several other contributions focus on the consequences of including non-marketable
assets in portfolio choices. Svensson [1988] deals with the implicit pricing of non-traded
assets. Longstaff [2009] considers a portfolio position in illiquid assets taking its temporal
irreversibility into account. In particular, he finds that in the presence of illiquid assets
agents choose not to hold the market portfolio but they tend to hold highly polarized
portfolios. Our results confirm the departure from the market portfolio.

The problem of non-marketable investment evaluation has also been analyzed by em-
pirical works. Kerins et al. [2004] use market data to estimate the cost of capital for
well-diversified venture capital investors and under-diversified entrepreneurs. The au-
thors support a direct relationship between the cost of capital and the entrepreneur’s
level of commitment to the non-marketable investment she is evaluating. Similar evi-
dence is found in Müller [2008]. In particular, the author shows how idiosyncratic risk
increases the cost of capital, since higher equity returns are required as compensation
for under-diversification. Moskowitz and Vissing-Jorgensen [2002] analyze the problem
of under-diversification for entrepreneurial households. Based on the data of the Survey
of Consumer Finances (SCF), the authors find that in most cases households decide to
invest on a single firm that they manage.

We contribute to the existing literature in more than a way. Considering a set of in-
creasingly realistic assumptions about non-marketable asset divisibility and rationality,
we analyze the problem of non-marketable investments in a systematic way. Furth-
more, applying the original definition of bounded rationality, we show and measure how
it affects investor portfolio choices, and, in particular, their portfolio risk. In addition,
we stress that the presence of a non-marketable asset leads the investor to redefine the
weights of the portfolio of only-risky assets. After that, we define a set of conditions
under which the non-marketable asset represents a good investment. Finally, consider-
ing several scenarios an investor may face in reality, we depict the efficient frontier for
the investor. In particular, we show that, under certain conditions, the efficient frontier
exhibits non-linearities and intervals of discontinuity. This result permits us to divide
investors among risk-tolerant entrepreneurs and risk-averse clerks.

Our analytical formalization is intentionally simple and limited to the single-period
case (myopic portfolio choice). The single-period case is very important for both prac-
titioners and academics: multi-period investment strategies involving hedging demands
are prone to estimation errors that outweigh the gains of a more realistic dynamic specification. Furthermore, the single-period case is optimal when there is a single period horizon, investment opportunities are constant, investment opportunities are stochastic but unhedgeable, and agents utility is logarithmic [Brandt, 2009]. In all other cases, the single-period approach is a reasonable approximation that is useful to give a simplified representation of reality.

The remainder of the paper is structured as follows. Section 2.2 describes the theoretical framework we use in our analysis of investment choices. Varying the assumptions on indivisibility of the non-marketable investment and on rationality, we identify four possible cases an investor may face. Section 2.3 proposes a comparison among the solutions of these four cases. Afterwards, in section 2.4, we describe the efficient frontier for the investor considering plausible scenarios. The final section concludes the paper.

2.2. Four cases for a model of investment choices

In subsequent sections, we present four possible perspectives (cases) an investor faces when evaluating an investment opportunity. Case 1 refers to the traditional problem of a financial market with \(N\) risky assets and a risk-free asset. In this scenario, an investor allocates her wealth by choosing the weights for the risk-free asset and the \(N\) risky assets. The problem is well known in financial literature (see for example Elton et al. [2008]), but we report it to introduce the notation and to compare it with the results we obtain from the other cases.

In cases 2, 3 and 4, we assume that the investor has an exclusive access to a non-marketable investment besides the \(N\) assets and the risk-free asset. In that context, we consider a set of increasingly realistic assumptions. First, in case 2, we analyze the situation when the non-marketable investment is continuously divisible, i.e. its portfolio weight can assume any real value. We are aware that non-marketable investments are often indivisible, but we solve this case in order to obtain lower bound portfolios in terms of risk over all scenarios. Thus, they represent benchmark solutions obtained in ideal conditions.

When conditions are less than ideal, further analysis is necessary. Thus, in cases 3 and 4, we take a new set of more realistic assumptions into account. In case 3, we consider the instance when the non-marketable investment total cost is indivisible, and the investor commits its total amount to the investment. Finally, in case 4, we see how bounded rationality shifts the possible choices of the investor of case 3. In particular, we take into account the situation when she invests in the indivisible non-marketable asset
2. Indivisible non-marketable assets and suboptimal solutions

and the limited set of information and resources forces her to choose only the weights for the risk-free asset and the risky assets portfolio as a whole. In this sense, the investor simplifies the optimization problem she faces, accepting a satisfactory solution rather than an optimal one.

2.2.1. N risky assets and a risk-free asset (case 1)

Consider a portfolio of \( N \) assets with weights \( \omega \). The return of this portfolio can be expressed by \( \mu_P = \omega'\mu \), where \( \mu = (\mu_1, \ldots, \mu_N)' \) is the vector of expected returns of the \( N \) risky assets. The variance of this portfolio is given by \( \sigma^2_P = \omega'\Sigma\omega \), where \( \Sigma \) represents the positive-definite \( N \times N \) variance-covariance matrix of the returns of the risky assets.

Now, suppose the market has a risk-free asset with return \( r_F \) besides \( N \) risky assets, and that lending and borrowing at the interest rate \( r_F \) are allowed. For a given value of portfolio expected return, \( \mu_P = k \), investors prefer the portfolio with the lowest variance. They face the problem

\[
\begin{align*}
\min_{\omega} & \quad \frac{1}{2} \omega'\Sigma\omega \\
\text{subject to} & \quad \omega'\left(\mu - r_F\iota\right) = k - r_F,
\end{align*}
\]

(2.1)

where \( \iota = (1, \ldots, 1)' \) denotes an \( N \)-dimensional vector in which all the elements are equal to 1. The constraints \( \omega'\mu + \omega'F = k \) and \( \omega'\iota + \omega_F = 1 \) are implicit in \( \omega'\left(\mu - r_F\iota\right) = k - r_F \).

Setting up the Lagrangian and solving the problem, the optimal portfolio weights and the portfolio variance are

\[
\omega_1 = \frac{k - r_F}{A} \Sigma^{-1}(\mu - r_F\iota),
\]

(2.2)

\[
\sigma^2_{P1} = \omega_1'\Sigma\omega_1 = \frac{(k - r_F)^2}{A},
\]

(2.3)

where \( A \equiv (\mu - r_F\iota)'\Sigma^{-1}(\mu - r_F\iota) \).

The variance of all portfolios obtained by varying \( k \) represents a parabola in the space \((\mu; \sigma^2)\), and consists of two straight lines with a common intercept at \( r_F \) in the space \((\mu; \sigma)\). However, as you can see in figure 2.1, only portfolios on the right line are relevant in practice \((k > r_F)\), since they are Pareto-optimal portfolios.

A well known result in portfolio theory (see for example Elton et al. [2008]) refers to a particular portfolio, \( M \), that is located on the identified efficient frontier, and is
2. Indivisible non-marketable assets and suboptimal solutions

Figure 2.1.: The frontiers of the four cases in the space $(\mu; \sigma)$ composed only by risky assets ($\omega_F = 0$). The weight vector of this portfolio is

$$\omega_M = \frac{\Sigma^{-1}(\mu - r_F \mathbf{1})}{\mathbf{1}'\Sigma^{-1}(\mu - r_F \mathbf{1})},$$ 

(2.4)

Let denote the expected return and the variance of $M$ respectively by $\mu_M = \omega_M'\mu$, and $\sigma_M^2 = \omega_M'\Sigma\omega_M$.

If we assume that the $N$ assets are all the risky assets in the universe of investments, and that all investors have the same beliefs about $\mu$ and $\Sigma$, the locus of combinations of expected returns and variance for efficient portfolios is the same for all investors and $M$ is the market portfolio. It is worth noting that, by the separation theorem [Tobin, 1958], all the efficient portfolios can be obtained by a linear combination of the risk-free asset and $M$. In addition, it can be shown that

$$A = \frac{(\mu_M - r_F)^2}{\sigma_M^2} \equiv S^2,$$ 

(2.5)

where $S$ corresponds to the Sharpe ratio. The Sharpe ratio, or the market price of risk, measures the excess return per unit of risk. By equations (2.3) and (2.5) we obtain the well known Capital Market Line $\mu_P = r_F + \frac{\sigma_P}{\sigma_M}(\mu_M - r_F)$. 


2. Indivisible non-marketable assets and suboptimal solutions

2.2.2. N risky assets, a risk-free asset, and a non-marketable investment

In this section, we assume that an investor has exclusive access to a non-marketable investment besides the N assets and the risk-free asset. Let $\mu_I$, $\sigma^2_I$, and $\omega_I$ denote respectively the expected return, the variance, and the portfolio weight on this new investment. Also let $\sigma = (\sigma_{I,1}, \ldots, \sigma_{I,N})'$ be the vector of the N covariances among the risky assets and the non-marketable investment returns. We need to clarify that, coherently to our definition of non-marketable asset, $\mu_I$, $\sigma^2_I$, and $\sigma$ are free, in principle, to take any possible value. In particular, they are not influenced by the market pressures that drive $\mu$ and $\Sigma$.

When the approach presented in case 1 is applied to investment choices that include a non-marketable investment available only to one investor, some differences must be taken into account. In particular, we need to specify a new set of constraints for the optimization problem. In cases 2, 3, and 4, we present the solution to the optimization problem considering several scenarios. In case 2, the non-marketable investment is continuously divisible. In case 3, we take a more realistic scenario into account considering the non-marketable investment as indivisible: we keep the amount of wealth spent in the non-marketable investment fixed. In case 4, other things being equal to case 3, we relax the assumption of complete rationality of the investor, so that, not being able to pursue the general minimization problem, she is limited to a sub-optimal choice. In particular, in the latter scenario, the investor can only choose the weights on the market portfolio $M$ as a whole and the risk-free asset.

**Continuously divisible non-marketable investment (case 2)**

When the non-marketable investment is continuously divisible, the problem is

$$
\begin{align*}
\min_{\omega, \omega_I} & \frac{1}{2} (\omega' \Sigma \omega + \omega_I^2 \sigma^2_I + 2 \omega_I \omega' \sigma) \\
\omega' (\mu - r_F \mathbf{1}) & = k - r_F - \omega_I (\mu_I - r_F),
\end{align*}
$$

(2.6)

where $\omega$ and $\omega_I$ are determined simultaneously. Note that the constraints $\omega' \mu + \omega_F r_F + \omega_I \mu_I = k$ and $\omega' \mathbf{1} + \omega_F + \omega_I = 1$ are implicit in the form $\omega' (\mu - r_F \mathbf{1}) = k - r_F - \omega_I (\mu_I - r_F)$. From an analytical standpoint this problem is simply the problem faced in case 1 with an additional investment and is part of the more general discussion on added assets (for a recent contribution on this topic see Zhang et al. [2010].) However, it is worth presenting it explicitly to measure the effects of the non-marketable asset on the investor’s personal portfolio.
Solving this problem, we get the weights and the variance

\[
\omega_2 = \frac{(k - r_F)}{A + \frac{(\mu_I - r_F - B)^2}{\sigma_I^2 - C}} \Sigma^{-1} \left( \mu - r_F 1 - \frac{\mu_I - r_F - B}{\sigma_I^2 - C} \sigma \right),
\]

\[
\omega_I^2 = \frac{(k - r_F)}{A + \frac{(\mu_I - r_F - B)^2}{\sigma_I^2 - C}} \left( \frac{\mu_I - r_F - B}{\sigma_I^2 - C} \right),
\]

(2.7)

\[
\sigma_{P2}^2 = \frac{(k - r_F)^2}{A + \frac{(\mu_I - r_F - B)^2}{\sigma_I^2 - C}},
\]

(2.8)

where \( B \equiv (\mu - r_F 1) \Sigma^{-1} \sigma \) and \( C \equiv \sigma' \Sigma^{-1} \sigma \).

As in case 1, the variance describes a parabola in the space \((\mu; \sigma^2)\), and consists of two straight lines with a common intercept in \( r_F \) in the space \((\mu; \sigma)\). It is worth noting that, in this latter space, the relevant part of the frontier coincide with the line on the right. (See figure 2.1.)

Let \( \sigma_{IM} \) be the covariance between the returns of the non-marketable investment and the market portfolio. \( \sigma_{IM} \) is obtained analytically by

\[
x_M' \Omega x_I, \quad x_M = (\omega_M, 0)',
\]

\[
\Omega = \begin{bmatrix} \Sigma & \sigma \\ \sigma' & \sigma_I^2 \end{bmatrix}, \quad x_I = (0, \ldots, 0, 1)'.
\]

Developing the expression, we get \( \sigma_{IM} = \omega_M^I \sigma \). By the last expression, it is easy to recognize that

\[
B = (\mu_M - r_F) \frac{\sigma_{IM}}{\sigma_M^2}.
\]

(2.9)

Using equation (2.5), we are able to see easily that

\[
\sigma_{P2}^2 = \frac{(k - r_F)^2}{S^2 + \frac{\alpha^2}{\sigma_I^2 - C}},
\]

(2.10)

where \( \alpha \) is the Jensen’s alpha, i.e. \( \alpha = \mu_I - r_F - (\mu_M - r_F) \frac{\sigma_{IM}}{\sigma_M^2} \). Jensen’s alpha measures the excess return of the investment over its theoretical expected return determined by the CAPM. Note that, all things being equal, the portfolio total risk is inversely related to the square of the Sharpe ratio and the square of Jensen’s alpha. Thus, whenever \( \alpha \neq 0 \), the portfolio risk in case 2 is smaller than that in case 1. In particular, when \( \alpha < 0 \), \( \omega_I^2 \) will be negative.
2. Indivisible non-marketable assets and suboptimal solutions

Indivisible non-marketable investment (case 3)

In this section, we consider again an individual facing the portfolio problem, when she is the only one who can access the non-marketable investment. However, we make an additional assumption: the investment total cost $I$ is indivisible and it has to be exactly covered. Therefore $\omega_I \equiv \frac{I}{W}$ is the fraction of the total wealth used for the non-marketable investment. This is a more realistic assumption than the one presented in the previous case, since non-marketable investments are often indivisible.

Therefore the problem is

$$
\begin{align*}
\min_{\omega} & \quad \frac{1}{2} \omega' \Sigma \omega + \frac{I^2}{W^2} \sigma_I^2 + 2 \frac{I}{W} \omega' \sigma \\
\text{s.t.} & \quad \omega' (\mu - r_F t) = k - r_F - \frac{I}{W} (\mu_I - r_F).
\end{align*}
$$

(2.11)

Note that the value of $\omega_I$ is known, since it is fixed at $\omega_I \equiv \frac{I}{W}$.

Solving this problem, the weights and the variance are

$$
\omega_3 = \omega_1 + \frac{I}{AW} (B - \mu_I + r_F) \Sigma^{-1} (\mu - r_F t) - \frac{I}{W} \Sigma^{-1} \sigma,
$$

(2.12)

$$
\sigma_{P3}^2 = \sigma_{P1}^2 - 2 \frac{k - r_F}{A} \frac{I}{W} (\mu_I - r_F - B) + \frac{I^2}{W^2} \left[ \frac{(\mu_I - r_F - B)^2}{A} + \sigma_I^2 - C \right].
$$

(2.13)

$\sigma_{P3}^2$ describes a parabola in the space $(\mu; \sigma^2)$, and an hyperbola in the space $(\mu; \sigma)$. It is worth noting that, as it will be clearly stated in a subsequent observation, because of the constraint $\omega_I \equiv \frac{I}{W}$, neither the parabola nor the hyperbola intersect the $\mu$-axis. As you can see in figure 2.1, efficient portfolios are those on the right side of the hyperbola.

After substituting $A$ and $B$, we get the following expression:

$$
\sigma_{P3}^2 = \sigma_{P1}^2 - 2 \frac{k - r_F}{S^2} \frac{I}{W} \alpha + \frac{I^2}{W^2} \left[ \frac{\alpha^2}{S^2} + \sigma_I^2 - C \right].
$$

(2.14)

Note that the total risk described in case 3 corresponds to the total risk in case 1 plus two terms, the latter always positive and the former negative if $I, \alpha > 0$.

Indivisible non-marketable investment and bounded rationality (case 4)

In this section, we consider the case when bounded rationality forces the investor to sub-optimal choices. Solving the problem described in the previous section requires an elaborate information processing strategy. However, a single investor may face some
limitations on her ability to obtain and process the information required to monitor $N + 1$ assets.

To take this consideration into account, we assume that the investor simply holds a three-asset portfolio, consisting of a fixed position in the non-marketable investment, a position ($\omega_M$) in the market portfolio (described in case 1)\(^2\), and a position in the risk-free asset. In this sense, the investor exploits the knowledge of all investors in the market, since their solution to case 1 determines the weights of the market portfolio, $\mu_M$ and $\sigma_M^2$.

Note that, assuming that case 1 is implicitly solved, the only additional information our investor needs to know is $\sigma_{IM}$, $\sigma_I$, and $\mu_I$. Furthermore, bounded rationality avoids the optimization problem to the investor. In fact, given $\omega_I \equiv \frac{I}{W}$, the solution for $\omega_M$ comes out by the simple combination of the two usual constraints: $\omega_M \mu_M + \omega_F r_F + \omega_I \mu_I = k$ and $\omega_M + \omega_F + \omega_I = 1$.

$$\omega_M = \frac{1}{\mu_M - r_F} \left[ (k - r_F) - \frac{I}{W} (\mu_I - r_F) \right].$$

(2.15)

As we can see, in this case the share of wealth invested in the market is equal to the pursued excess return over the risk-free asset return minus $\frac{I}{W}$ times the non-marketable investment excess return over the risk-free asset return, all expressed as a percentage of the market excess return over the risk-free asset return.

The variance of this portfolio is

$$\sigma_{P4}^2 = \frac{\sigma_M^2 \left[ (k - r_F) - (\mu_I - r_F) \frac{I}{W} \right]^2}{(\mu_M - r_F)^2} + \frac{I^2 \sigma_{M}^2 \left[ (k - r_F) - (\mu_I - r_F) \frac{I}{W} \right]}{(\mu_M - r_F)^2} + \frac{I^2}{W^2} \sigma_I^2,$$

(2.16)

which, as in case 3, describes a parabola in the space ($\mu; \sigma^2$), and an hyperbola in the space ($\mu; \sigma$). Because of the bounded rationality, the hyperbola in case 4 is located above the hyperbola in case 3. As you can see in figure 2.1, efficient portfolios are those on the right side of the hyperbola.

\(^2\)Since we are dealing with bounded rationality, in this subsection we can consider the market portfolio as the average solution of all the agents on the market. If these agents have rational expectations, the average solution will tend to be the optimal one.
2.3. Solutions to the four cases: a comparison

In this section, we propose a comparison among the solutions to the four cases. This comparison allows us to understand when the exclusive access to the non-marketable investment under consideration may improve the risk-return trade-off for the investor, i.e. the investor faces a smaller risk to pursue the same expected return.

In each case we can express the standard deviation as function of the portfolio expected return. In this way, we are able to evaluate the relative convenience of the four solutions in terms of risk. That is the approach we follow in section 2.3.1. As we will see, there is no clear-cut order of dominance along all the range of variation of $k$ between case 1 and 3, and between case 1 and 4. We focus on the comparison among these cases in section 2.3.2.

2.3.1. Standard deviations

- **Case 1**

  By equation (2.3), we are able to get the standard deviation for case 1

  $$\sigma_{p_1} = \frac{k - r_F}{S}.$$  \hspace{1cm} (2.17)

- **Case 2**

  Similarly to case 1, by equation (2.10), we can obtain the standard deviation for case 2.

  $$\sigma_{p_2} = \frac{k - r_F}{\sqrt{S^2 + \frac{\alpha^2}{\sigma_I^2} C}}.$$ \hspace{1cm} (2.18)

  By comparing equations (2.17) and (2.18) the following observation is straightforward.

  **Observation 1** When $\alpha = 0$ the frontier in case 2 corresponds exactly to that of case 1. When $\alpha \neq 0$, the frontier of case 2 has a slope smaller than that of case 1. In any case, the two frontiers start at the same point on the $\mu$-axis, i.e. $k = r_F$.

- **Case 3**

  Case 3 differs from the previous ones, since the equation for $\sigma_{p_3}$ describes an hyperbola in the space $(\mu; \sigma)$. That is due to the extra constraint we impose on the fraction of
2. Indivisible non-marketable assets and suboptimal solutions

wealth committed by the investor to the non-marketable investment. In appendix A, we obtain the standard form

\[
\left( \frac{k - r_F - \frac{I}{W^\alpha}}{S^2 \frac{I^2}{W^2} (\sigma_I^2 - C)} \right)^2 - \frac{\sigma_{P3}^2}{\frac{I^2}{W^2} (\sigma_I^2 - C)} = -1, \tag{2.19}
\]

which is an hyperbola in the variables \((k, \sigma_{P3})\). This hyperbola is centered at a point on the \(\mu\)-axis with coordinates \((T, 0)\), where

\[
T = r_F + \frac{I}{W} \alpha. \tag{2.20}
\]

When \(k = T\), the portfolio standard deviation, \(\sigma_{P3}\) reaches its minimum. Thus, it follows the subsequent observation.

**Observation 2** If \(I, \alpha > 0\), \(\sigma_{P3}\) is at its minimum for a value of \(k > r_F\).

The upward sloping asymptote\(^3\) of this hyperbola is

\[
\sigma_{P3}^{\text{asym}} = \frac{k - T}{S} = \frac{k - r_F}{S} - \frac{I}{W} \frac{\alpha}{S}. \tag{2.21}
\]

Note that this asymptote is simply the frontier of case 1 translated by \(-\frac{I}{W} \alpha\).\(^4\) Thus, it is easy to prove the following observation.

**Observation 3** Whenever the conditions of observation 2 hold, that is \(I, \alpha > 0\), an intersection point between the frontier of case 1 and case 3, \(k_{13}\), exists. The frontier of case 3 dominates that of case 1 to the right of \(k_{13}\), while the opposite applies to the left of \(k_{13}\). If \(\alpha = 0\), the intersection point does not exist, and the frontier of case 1 dominates along all the interval of variation of \(k\).

Observation 3 states that committing a fixed amount of one’s wealth to a non-marketable investment can be convenient (there is an interval where the frontier of case 3 dominates that of case 1) if \(\alpha > 0\) – i.e. whenever the non-marketable investment has a better risk-return trade-off compared to market assets with the same level of systematic risk.

\(^{3}\)An hyperbola of the form \(\frac{(x - T)^2}{b^2} - \frac{y^2}{a^2} = -1\) has an asymptote equal to \(y = \frac{a}{b} (x - T)\).

\(^{4}\)It is worth noting that, in the space \((\sigma; \mu)\), the asymptote becomes \(\mu = r_f + \frac{(\mu_M - r_F) \sigma_{P3}^{\text{asym}} + I}{\frac{\alpha}{W}}\).

Henceforth, this asymptote is simply a translation of the Capital Market Line along the \(\mu\)-axis of the amount \(\frac{I}{W} \alpha\). The investor’s individual frontier lies between the Capital Market Line and this asymptote.
2. Indivisible non-marketable assets and suboptimal solutions

The vertex of this branch of the hyperbola is located at \((T; \sigma_P^{\min})\). Using equation (2.19) we find

\[
\sigma_P^{\min} = \frac{I}{W} \sqrt{\sigma_I^2 - C}.
\]

(2.22)

Then it follows

**Observation 4** In case 3, the minimum standard deviation is greater than zero whenever \(\sigma_I^2 > C\).

Thus, even if the non-marketable investment has an advantageous risk-return trade-off, the investor can never eliminate risk if \(\sigma_I^2 > C\).

As it is clear from figure 2.1, there is a tangency point between the frontier of case 2 and that of case 3. This point is located on the right of the vertex of the hyperbola, and corresponds to the point where the \(\omega_I\) we get from the optimization procedure of case 2 coincides with the \(\frac{I}{W}\) we assumed fixed in case 3.

- Case 4

Analogously to case 3, \(\sigma_P^4\) describes an hyperbola in the space \((\mu; \sigma)\). In appendix B, we obtain the standard form

\[
\frac{(k - r_F - \frac{I}{W} \alpha)^2}{S^2 I^2 W^2 \left( \sigma_I^2 - \frac{\sigma_{I,M}^2}{\sigma_M^2} \right)} - \frac{\sigma_P^4}{I^2 W^2 \left( \sigma_I^2 - \frac{\sigma_{I,M}^2}{\sigma_M^2} \right)} = -1,
\]

(2.23)

which is an hyperbola in the variables \((k, \sigma_P^4)\). Also this hyperbola, as the one in case 3, is centered at the point on the \(\mu\)-axis with coordinates \((T,0)\) and has the asymptote \(\sigma_{P}^{\text{asym}}\).

When \(k = T\), the portfolio standard deviation, \(\sigma_P^4\), reaches its minimum. Therefore, it can be easily proved the next observation.

**Observation 5** If \(I, \alpha > 0\), \(\sigma_P^4\) is at its minimum for a value of \(k > r_F\).

We recall by case 3 that the asymptote \(\sigma_{P}^{\text{asym}}\) is simply the frontier of case 1 translated by \(-\frac{I}{W} \frac{\alpha}{S}\). Therefore, it follows the next observation.

**Observation 6** Whenever the conditions of observation 5 hold, that is \(I, \alpha > 0\), an intersection point between the frontier of case 1 and case 4, \(k_{14}\), exists. The frontier of case 4 dominates that of case 1 to the right of \(k_{14}\), while the opposite applies to the left of
2. Indivisible non-marketable assets and suboptimal solutions

$k_{14}$. If $\alpha = 0$, the intersection point does not exist, and the frontier of case 1 dominates along all the interval of variation of $k$.

Even in the case of bounded rationality, we find that committing a fixed amount of one’s wealth to a non-marketable investment can be convenient if $\alpha > 0$.

The vertex of this branch of the hyperbola is located at $(T;\sigma_{P4}^{\min})$. Using equation (2.23) we find

$$\sigma_{P4}^{\min} = \left| I \right| W \left( \sqrt{\sigma_I^2 - \frac{\sigma_{I,M}^2}{\sigma_M^2}} \right).$$

Then it follows

**Observation 7** In case 4, the minimum standard deviation is greater than zero whenever $\sigma_I^2 > \frac{\sigma_{I,M}^2}{\sigma_M^2}$.

Thus, even if the non-marketable investment has an advantageous risk-return tradeoff, the investor can never eliminate risk if $\sigma_I^2 > \frac{\sigma_{I,M}^2}{\sigma_M^2}$. The condition $\sigma_I^2 > \frac{\sigma_{I,M}^2}{\sigma_M^2}$ has a significant financial meaning. Let us define the non-marketable return as $r_I = \gamma + \beta r_M + u$, where $\beta = \frac{\sigma_{I,M}}{\sigma_M^2}$. Suppose that the market return $r_M$ has expected value $E(r_M) \equiv \mu_M$ and variance $\text{Var}(r_M) \equiv \sigma_M^2$; the idiosyncratic component of the return, $u$, has expected value $E(u) \equiv 0$ and variance $\text{Var}(u) \equiv \varpi^2$; and the covariance between $r_M$ and $u$ is $\text{Cov}(r_M, u) = 0$. It follows that the expected value of $r_I$ is $\mu_I = \gamma + \beta \mu_M$, while its variance is $\sigma_I^2 = \left( \frac{\sigma_{I,M}}{\sigma_M^2} \right)^2 \sigma_M^2 + \varpi^2 = \frac{\sigma_{I,M}^2}{\sigma_M^2} + \varpi^2$. Therefore, $\frac{\sigma_{I,M}^2}{\sigma_M^2}$ measures the systematic risk, while $\varpi^2$ the idiosyncratic risk. In other words, by observation 7 we can say that the $\sigma_{P4}^{\min}$ is greater than zero whenever the non-marketable investment has a component of idiosyncratic risk. Since a single investment is likely to have an idiosyncratic risk component, the $\sigma_{P4}^{\min}$ will be virtually always greater than zero.

Since the only difference between case 4 and case 3 is the bounded rationality of the investor, we are able to measure its effect in terms of risk for any given $k$. The difference between $\sigma_{P3}$ and $\sigma_{P4}$ is at its maximum in correspondence of the two vertexes, and it declines the more we move to the right along the $\mu$-axis, as the two hyperbolas converge to the same asymptote. Therefore, defining $BR(k)$ as the bounded rationality distortion, we get

$$BR(k) \leq \sigma_{P4}^{\min} - \sigma_{P3}^{\min} = \left| I \right| \frac{W}{\sqrt{\sigma_I^2 - \frac{\sigma_{I,M}^2}{\sigma_M^2}} - \sqrt{\sigma_I^2 - C}}.$$  

5It is evident here as bounded rationality depends on the values $C$ and $\frac{\sigma_{I,M}^2}{\sigma_M^2}$. In order to compare
2. Indivisible non-marketable assets and suboptimal solutions

2.3.2. Intersection points

Observations 3 and 6 show that one can determine where the solution to case 1 dominates the solutions to case 3 and 4 (or vice versa), only after one ascertains the existence of an intersection point between the frontiers. Thus, in this section, we look for the intersection points between the frontier of case 1 and the frontiers of cases 3 and 4. The intersection point is defined as the value of \( k \) that equates the standard deviations of two cases.

On one side, for cases 1 and 3, we have by equation (2.3)

\[
A\sigma_p^2 = (k - r_F)^2. \tag{2.26}
\]

On the other side, by equation (2.13), we have

\[
A\sigma_p^2 = (k - r_F)^2 - 2(k - r_F)\frac{I}{W} (\mu_I - r_F - B) + \frac{I^2}{W^2} \left( (\mu_I - r_F - B)^2 - AC + A\sigma_I^2 \right). \tag{2.27}
\]

Therefore

\[
k_{13} = r_F + \frac{I}{2W} \frac{S^2(\sigma_I^2 - C) + \alpha^2}{\alpha}. \tag{2.28}
\]

It is straightforward by equation (2.28) to prove the following observation.

**Observation 8** The intersection point between the frontier of case 1 and the frontier of case 3, \( k_{13} \), is inversely related to the value of \( \alpha \) and goes to infinity when \( \alpha \to 0 \).

For cases 1 and 4, on one side, we have equation (2.3); on the other side, by equation (2.16), we have

\[
\sigma_p^2 A = \left[ k - r_F - (\mu_I - r_F) \frac{I}{W} \right]^2 + 2 \frac{I}{W} \frac{\sigma_{1,M}}{\sigma_M^2} \left[ k - r_F - (\mu_I - r_F) \frac{I}{W} \right] (\mu_M - r_F) + \frac{\sigma_{1,M}^2 (\mu_M - r_F)^2}{\sigma_M^2 W^2}. \tag{2.29}
\]

_better the values, we can develop the expression

\[
\frac{\sigma_{2,M}}{\sigma_{2,M}} = \frac{\omega_M \omega_M'}{\sigma_M^2} \sigma = \frac{\Sigma^{-1}(\mu - r_F) \Sigma^{-1}(\mu - r_F)}{(\mu - r_F)^2}\sigma = \frac{B^2}{A}. \]

Therefore, the greater the difference between \( C \) and \( \frac{B^2}{A} \), the greater the distortion led by the bounded rationality.
After some algebraic manipulation, we derive the expression

\[ k_{14} = r_F + \frac{I}{2W} \left( \sigma_I^2 - \frac{\sigma_{I,M}^2}{\sigma_M^2} \right) + \alpha^2. \] (2.30)

It is then straightforward the following observation.

Observation 9 The intersection point between the frontier of case 1 and the frontier of case 4, \( k_{14} \), is inversely related to the value of \( \alpha \) and goes to infinity when \( \alpha \to 0 \).

2.4. The efficient frontier for different real scenarios and the choice of the investor

In the previous sections, we described the frontier for each of the four theoretical cases. In this section, we relate those theoretical cases to several scenarios an investor may face in reality. First, in reality, not all individuals have access to non marketable investments. Second, even if an investor identifies a viable non-marketable investment, she may choose not to undertake it accordingly to her risk-return preferences. Finally, investor’s rationality may be bounded. Thus, the objective of this section is to depict the effective frontier that emerges from each possible scenario. Table 2.1 summarizes all scenarios we consider.

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Full rationality</th>
<th>Bounded rationality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investment not available</td>
<td>Frontier of case 1</td>
<td>-</td>
</tr>
<tr>
<td>Continuously divisible investment</td>
<td>Frontier of case 2</td>
<td>-</td>
</tr>
<tr>
<td>Indivisible Investment</td>
<td>Frontier of case 1 or 3</td>
<td>Frontier of case 1 or 4</td>
</tr>
</tbody>
</table>

Table 2.1.: Efficient frontiers and scenarios

2.4.1. Four possible scenarios

- The investor does not identify the non-marketable investment

This scenario coincides with case 1. The relevant frontier is the standard capital market line, which describes the locus of efficient portfolios when all investors have access to the same set of assets. Figure 2.1 (case 1) presents the frontier in the space \((\mu; \sigma)\). It is worth noting that, when \( I = 0 \), this frontier is the limit result for all the other scenarios.

- The investor identifies the non-marketable investment and can choose to commit any amount to it
In this scenario, which corresponds to case 2, the investor has exclusive access to a continuously divisible non-marketable investment. In this sense, the investment is available only to her. Being $\omega_I$ free to vary, this scenario is the best situation the investor may face among all the scenarios described, i.e. the solutions to case 2 are always at least as good as the solutions to all the other cases. To see this, note that cases 1, 3 and 4 may be interpreted as restricted versions of case 2, where additional constraints are imposed. In case 1, $\omega_I = 0$; in case 3, $\omega_I \equiv \frac{I}{W}$; and in case 4, $\omega_I \equiv \frac{I}{W}$ and $\omega \equiv \omega_M$. As previously mentioned, there is a tangency point between the frontier of case 2 and that of case 3, where $\omega_I = \frac{I}{W}$ (see figure 2.1).

- The investor identifies the non-marketable investment, but it is just a “take it or leave it” choice

In this scenario the investor has exclusive access to an indivisible non-marketable investment, therefore the amount she can commit to the investment is fixed and equal to $I$. It is important to note that the investor is not compelled to take on the non-marketable investment. Thus, for any given $k$, the frontier will be that of case 1 or that of case 3 according to the level of the standard deviation. In particular, solutions to case 1 may be Pareto-preferred for some intervals of $k$, but solutions to case 3 may be superior for other intervals. Depending on the value assumed by $\alpha$, which measures the excess return of the non-marketable investment over the theoretical expected return at that level of risk, three frontier shapes emerge.

The first frontier (see figure 2.2) corresponds to the case when $\alpha = 0$. In this situation, the investor would disregard the opportunity to undertake the non-marketable investment, since the solutions to case 1 dominate those of case 3 for all possible values of $k$. In practice, since the non-marketable investment has an expected return in line with that of other market assets at the same level of risk, it is not worth taking on the non-marketable investment.

The second frontier occurs when $\alpha$ assumes a value between zero (excluded) and the value corresponding to the situation when the frontier of case 1 intersects the hyperbola of case 3 exactly on its vertex. Analytically, we get this value by substituting $T$ for $k$ and $\sigma_{P3}^{\text{min}}$ for $\sigma_{P1}$ in equation (2.17). The interval we obtain is $0 < \alpha \leq S \sqrt{\sigma_I^2 - C}$. The resulting frontier is shown in figure 2.3. The frontier is that of case 1 for low values of $k$, and that of case 3 for higher values of $k$. Thus, the investor’s choice will ultimately depend on her risk-return preferences.

Finally, in figure 2.4, we can see the shape of the relevant frontier when $\alpha > S \sqrt{\sigma_I^2 - C}$. It is worth noting that between $k = S \sigma_{P3}^{\text{min}} + r_F$ and $k = T$ the frontier exhibits an
2. Indivisible non-marketable assets and suboptimal solutions

Figure 2.2.: Relevant frontier when $\alpha = 0$

Figure 2.3.: Relevant frontier when $0 < \alpha \leq S\sqrt{\sigma_I^2 - C}$
2. Indivisible non-marketable assets and suboptimal solutions

Figure 2.4.: Relevant frontier when $\alpha > S \sqrt{\sigma^2_I - C}$

interval of discontinuity. It would not be rational for the investor to choose $k$ so that $S \sigma_{P3}^{\min} + \tau_F < k < T$, since she can get $k = T$ for a lower level of risk. Besides the interval of discontinuity, the frontier is that of case 1 for low values of $k$, and that of case 3 for high values of $k$.

- The investor identifies the “take it or leave it” non-marketable investment, but her rationality is bounded

This last scenario is similar to the preceding one. The only difference is due to the bounded rationality of the investor. Now the relevant frontier is obtained combining the frontiers of case 1 and 4. Also in this scenario, there are three different shapes for the frontier that are very similar to those illustrated in figures 2.2, 2.3, and 2.4. The intervals of $\alpha$ which generate the three shapes are $\alpha = 0, 0 < \alpha \leq S \sqrt{\sigma^2_I - \sigma^2_{M} / \sigma^2_M}$, and $\alpha > S \sqrt{\sigma^2_I - \sigma^2_{LM} / \sigma^2_M}$.

2.4.2. The choice of the investor: entrepreneurs vs. clerks

In reality individuals are heterogeneous in their risk-return preferences. Each frontier presented in the previous subsection leads to a possible optimum choice for the investor. The optimum choice is generally identified by the tangency point between her indifference
2. Indivisible non-marketable assets and suboptimal solutions

Figure 2.5.: Indifference curves and investment choices

curve and the frontier. This tangency point ultimately relates to the investor’s risk-return profile: higher values of \( k \) are chosen by risk prone individuals, lower values of \( k \) are chosen by risk-averse individuals.

In the first place, the investor may or may not identify a non-marketable alternative. However, even when she identifies the opportunity of a non-marketable investment, this does not guarantee that she will undertake it. Her choice will depend on her attitude towards risk. In particular, scenarios 3 and 4 allow us to describe the entrepreneurial attitude. If she is a risk-averse individual, a clerk, she will not take on the investment. Conversely, if she is a risk-prone individual, an entrepreneur, she will take on the investment. As it is evident from figure 2.5, a clerk’s indifference curve is tangent to the first part of the frontier (solution to case 1), an entrepreneur’s indifference curve is tangent to the second part of the frontier (solution to case 3 or case 4).

2.5. Conclusions

Portfolio theory deals with investment decisions and optimal allocation of investor wealth. In the traditional framework, it is generally assumed that investments are marketable, continuously divisible, and investors are fully rational. However, reality often departs in a significant way from these assumptions. In order to give a theoretical basis to invest-
2. Indivisible non-marketable assets and suboptimal solutions

ment choices under all circumstances, in this paper we define what happens when we deviate from these ideal-school assumptions. To the best of our knowledge, there is yet to be a contribution that deals in a general and exhaustive way with investments that are by definition out of the market.\(^6\)

We construct four cases describing different situations an investor may face, varying the assumptions on indivisibility of the non-marketable investment and on rationality of the individual. By the subsequent comparison of the results, we highlight several interesting implications as the value of the parameters \(\alpha, S, \frac{I}{W}, \sigma_I^2, C\) and \(\frac{\sigma_{I,M}^2}{\sigma_M^2}\) vary. First, we determine the shape of the efficient frontiers for each case we consider and then we identify the intersection points among these efficient frontiers. We also measure the effects of bounded rationality in terms of additional risk of the investor portfolio. We show that an important role is played by Jensen’s alpha and Sharpe ratio. In particular, when \(\alpha = 0\), there is no convenience for the investor to include the non-marketable investment in her portfolio. Furthermore, we highlight that, when the investment needs a fixed amount to be undertaken, the portfolio risk is always greater than zero, i.e. it cannot be eliminated.

In last sections, we compare the solutions to the four cases by taking into account several scenarios that may occur in reality. This allows us to identify new shapes of the efficient frontier, some of which are non-linear and exhibit intervals of discontinuity. In particular, when the investor identifies the non-marketable investment, but it is just a “take it or leave it” choice, there are some values of \(\alpha\) that allow us to split investors between clerks and entrepreneurs according to their attitude towards risk. This may suggest why some individuals engage in personal enterprises and others do not.

Most of the shortcomings of our work are consistent with those of traditional portfolio theory. For example, the mean-variance representation of the space of investments disregards higher moments of asset return distributions. Furthermore, we consider only one period in our analysis. Considering higher moments or a multi-period analysis could shed more light on the dynamic effects of non-marketability and bounded rationality (at the price of a greater analytical complexity). We intend bounded rationality as a limiting factor of the investor’s processing ability in her optimization problem, or as a lack in her availability of information on the risky assets market. We know that other more complex development of bounded rationality may be taken into account. However, we intentionally use a simple definition of bounded rationality, which is enough for our purpose. Lastly, we consider only one non-marketable asset. Considering more

\(^6\)We recall that our definition of non-marketable differs from illiquidity.
than one non-marketable asset could enrich our understanding of the interaction among marketable and non-marketable investments. Another direction for further research may address in depth how different kinds of utility functions interact with the frontiers we identify.

2.6. Appendix A

By equation (2.13)

\[ \sigma^2_{P3} = \frac{(k - r_F)^2}{A} - 2\frac{(k - r_F)}{A} \frac{I}{W} (\mu_I - r_F - B) + \frac{I^2}{AW^2} (\mu_I - r_F - B)^2 + \frac{I^2}{W^2} (\sigma^2_I - C) \]

\[ \sigma^2_{P3} = \frac{1}{A} \left[ k - r_F - \frac{I}{W} (\mu_I - r_F - B) \right]^2 + \frac{I^2}{W^2} (\sigma^2_I - C) \]

\[ \sigma^2_{P3} = \frac{1}{A} \left[ k - r_F - \frac{I}{W} (\mu_I - r_F - B) \right]^2 = \frac{I^2}{W^2} (\sigma^2_I - C) \]

\[ \frac{I^2}{W^2} (\sigma^2_I - C) - \frac{k - r_F - \frac{I}{W} (\mu_I - r_F - B)^2}{A \frac{I^2}{W^2} (\sigma^2_I - C)} = 1 \]

2.7. Appendix B

By equation (2.16)

\[ \sigma^2_{P4} \frac{(\mu_M - r_F)^2}{\sigma^2_M} = \left[ k - r_F - (\mu_I - r_F) \frac{I}{W} \right]^2 + \]

\[ + 2 \frac{I}{W} \frac{\sigma_{I,M}}{\sigma^2_M} \left[ k - r_F - (\mu_I - r_F) \frac{I}{W} \right] (\mu_M - r_F) + \frac{\sigma^2_{I,M}}{\sigma^2_M} (\mu_M - r_F)^2 \frac{I^2}{W^2} \]

\[ \sigma^2_{P4} \frac{(\mu_M - r_F)^2}{\sigma^2_M} = \left[ k - r_F - (\mu_I - r_F) \frac{I}{W} + \frac{\sigma_{I,M}}{\sigma^2_M} (\mu_M - r_F) \frac{I}{W} \right]^2 - \]

\[ - \frac{\sigma_{I,M}}{\sigma^2_M} (\mu_M - r_F)^2 \frac{I^2}{W^2} + \sigma^2_{M} (\mu_M - r_F)^2 \frac{I^2}{W^2} \]

\[ \sigma^2_{P4} \frac{(\mu_M - r_F)^2}{\sigma^2_M} = \left\{ k - r_F - \frac{I}{W} \left[ \mu_I - r_F - (\mu_M - r_F) \frac{\sigma_{I,M}}{\sigma^2_M} \right] \right\}^2 + \]
2. Indivisible non-marketable assets and suboptimal solutions

\[
\begin{align*}
&\quad + \frac{(\mu_M - r_F)^2}{\sigma_M^2} \frac{I^2}{W^2} \left( \sigma_I^2 - \frac{\sigma_{I,M}^2}{\sigma_M^2} \right) \\
&\quad - \frac{\sigma_{P4}^2}{\sigma_M^2} \frac{I^2}{W^2} \left\{ k - r_F - \frac{I}{W} \left[ \mu_I - r_F - (\mu_M - r_F) \frac{\sigma_{I,M}^2}{\sigma_M^2} \right] \right\}^2 = \\
&\quad - \frac{\sigma_{P4}^2}{\sigma_M^2} \frac{I^2}{W^2} \left\{ k - r_F - \frac{I}{W} \left[ \mu_I - r_F - (\mu_M - r_F) \frac{\sigma_{I,M}^2}{\sigma_M^2} \right] \right\}^2 = 1
\end{align*}
\]
3. Entrepreneurial under-diversification: over optimism and overconfidence

3.1. Introduction

Risk propensity is a central theme of the entrepreneurial finance literature (Busenitz [1999]; Smith and Smith [2004]). Even if risk propensity seems to be the common characteristic in most definitions of entrepreneurship, past research has found that entrepreneurs do not have a systematic higher risk propensity than non-entrepreneurs (Brockhaus [1980]; Brockhaus and Horwitz [1986]; Palich and Ray Bagby [1995]; Simon et al. [2000]).

Several empirical findings, nevertheless, show that entrepreneurs often invest a large share of their personal wealth in one company, exposing themselves to idiosyncratic risk: their stake in the company is generally higher than the stake that a risk-return analysis would suggest (Heaton and Lucas [2000]; Moskowitz and Vissing-Jorgensen [2002]; Müller [2008]; Müller [2011]; Yazdipour [2011]). This exposure to idiosyncratic risk is very costly (Kerins et al. [2004]; Pattitoni et al. [2013]).

Several studies point out that indeed entrepreneurs demand compensation for their exposure to idiosyncratic risk (Müller [2008]; Müller [2011]). Thus, one possible explanation for this puzzling evidence - i.e., that entrepreneurs do not understand idiosyncratic risk - can be ruled out. Other justifications mostly rely on non-pecuniary benefits as benefits of control: entrepreneurs obtain substantial rewards from being their own boss and, thus, they are willing to accept a suboptimal risk-return trade-off (Moskowitz and Vissing-Jorgensen [2002]; Müller [2008]; Müller [2011]; Shefrin [2011]). It is still, however, debated why entrepreneurs overinvest in their private companies given the suboptimal risk-return trade-off.

In this paper, we propose a complementary story by suggesting that two behavioral biases may help explain this phenomenon. In particular, we focus on overconfidence and over optimism (Shefrin [2007]; Shefrin [2008]). Overconfidence may lead the entrepreneur to undervalue the actual risk on the investment on her private company. Over optimism
may cause her to overvalue its actual return. Thus, both these biases, which we include in our model as parameters, affect the very main variables of the risk-return analysis à la Markowitz [1952] and Markowitz [1959]: i.e., standard deviation and expected return. With a biased perception of standard deviation and expected return, an entrepreneur may choose suboptimal portfolio weights on her private company. In particular, if the entrepreneur perceives the private company to have a lower risk and a higher return than real, she may overweight the company in her portfolio. Through this parameterization, we are able to measure the potential bias in the portfolio weights of over optimist and/or overconfident entrepreneurs, and thus to explain, at least in part, entrepreneurial underdiversification.

Through a simulation analysis, we calculate how distinct parameters of overconfidence and over optimism affect entrepreneurial portfolio choices. More interestingly, our simulation analysis is useful to calculate the overconfidence and over optimism levels that, given a set of assumptions on the model parameters, are implicit in the entrepreneurs’ observable portfolio choices.

The structure of the paper is the following. In Section 3.2, we present our theoretical setup. In Section 3.3, we perform a numerical analysis. In section 3.4, we conclude.

3.2. Theoretical setup

We consider an entrepreneur that has to choose her portfolio allocation, i.e., which part of her wealth (and human capital) to invest in her private company and which to invest in the stock market. The portfolio optimization problem is dual: either the entrepreneur minimizes the risk for a given portfolio expected return, or she maximizes the return for a given portfolio risk. In Section 3.2.1, we discuss the impact of overconfidence on risk minimization. In Section 3.2.2, we deal with the effect of over optimism on return maximization.

3.2.1. Overconfidence

Risk minimization

Consider an entrepreneur who holds a portfolio of two risky assets with weights \( \omega = (\omega_I, \omega_M)' \) and a risk free asset with weight \( \omega_F \). The asset \( I \) is the entrepreneur investment in her private company and the asset \( M \) is the entrepreneur investment in a well diversified market portfolio. The excess return of the entrepreneur’s portfolio can be expressed by \( \mu_P = \omega'\mu \), where \( \mu = (\mu_I, \mu_M)' \) is the vector of the excess returns over
the risk free rate $r_F$. The variance of this portfolio is given by $\sigma^2_P = \omega' \Sigma \omega$, where

$$\Sigma = \begin{bmatrix} \sigma_I^2 & \sigma_{IM} \\ \sigma_{IM} & \sigma_M^2 \end{bmatrix}$$

represents the positive-definite variance-covariance matrix of the returns of the risky assets with $\det \Sigma = \sigma_I^2 \sigma_M^2 - \sigma_{IM}^2 > 0$.

For a given value of portfolio expected excess return, $\mu_P = k$, the entrepreneur prefers the portfolio with the lowest variance. She faces the problem

$$\begin{cases} \min \omega' \Sigma \omega \\ \omega' \mu = k \end{cases}$$

(3.1)

Note that the constraint $\omega_I + \omega_M + \omega_F = 1$ is implicit in $\omega' \mu = k$.

Setting up the Lagrangian and solving the problem [Pattitoni and Savioli, 2011], the optimal portfolio weights are

$$\omega(k) = \frac{k \Sigma^{-1} \mu}{\mu' \Sigma^{-1} \mu}.$$  

(3.2)

The first element of $\omega(k)$ represents the weight on the private company, namely

$$\omega_I(k) = \frac{k \left( \frac{\sigma_I^2 \mu_I - \rho_{IM} \sigma_I \sigma_M \mu_I}{\sigma_M^2 \mu_I^2 - 2 \rho_{IM} \sigma_I \sigma_M \mu_I \mu_M + \sigma_{IM}^2 \mu_M^2} \right)}{k \sigma_M^2 \alpha} = \frac{\sigma_I^2 \mu_I^2 - 2 \rho_{IM} \sigma_I \sigma_M \mu_I \mu_M + \sigma_{IM}^2 \mu_M^2}{k \sigma_M^2 \alpha},$$

(3.3)

where $\rho_{IM} = \frac{\sigma_{IM}}{\sigma_I \sigma_M}$ and $\alpha$ is the Jensen’s alpha, i.e., $\alpha = \mu_I - \rho_{IM} \frac{\sigma_I}{\sigma_M} \mu_M$.

**Overconfidence-driven under-diversification**

Overconfidence causes the entrepreneur to undervalue the actual risk of the investment in her private company. In this case, the biased standard deviation of the company returns, indicated by $\tilde{\sigma}_I$,\footnote{From now on, the hat over a symbol (e.g., $\tilde{\sigma}_I$) indicates a biased parameter or variable.} is lower than the actual standard deviation, i.e., $\tilde{\sigma}_I < \sigma_I$.

We model $\tilde{\sigma}_I$ as

$$\tilde{\sigma}_I = \sigma_I (1 - \delta_C), \quad \delta_C \in [0, 1),$$

(3.4)

where $\delta_C$ is the overconfidence parameter, which ranges from 0 (no overconfidence) to 1 (maximum overconfidence).\footnote{Since the covariance between the private company returns and the market returns is given by $\sigma_{1M} = \rho_{1M} \sigma_I \sigma_M$, if the perceived standard deviation of the company returns, $\tilde{\sigma}_I$, differs from the actual one, $\sigma_I$, then overconfidence leads to a biased perception of the covariance, $\tilde{\sigma}_{1M}$.} When $\delta_C$ tends to 1, then $\tilde{\sigma}_I$ tends to zero.

To analyze the impact on the portfolio weight in the private company caused by

\footnote{Choosing $\delta_C \in (-\infty, 1)$, we would allow for underconfidence.}
overconfidence, we define $\tilde{\omega}_I(k)$ as the $\omega_I(k)$ of Equation (3.3) with $\tilde{\sigma}_I$ in place of $\sigma_I$. We consider two cases. In the first case, the private company returns and the market portfolio returns are uncorrelated. In the second case, they are (positively or negatively) correlated.

**Case 1) Uncorrelated returns: $\rho_{IM} = 0$**

In this case, $\frac{\partial \tilde{\omega}_I(k)}{\partial \delta C} > 0$. Thus, when the private company returns and the market returns are uncorrelated, then the overconfident entrepreneur tends to overinvest in her private company and to be under-diversified. The overconfidence bias is thus positive and equal to

$$b_c = \tilde{\omega}_I(k) - \omega_I(k) > 0.$$  \hspace{2cm} (3.5)

**Case 2) Correlated returns: $\rho_{IM} \neq 0$**

Using the definition of $\tilde{\omega}_I(k)$, we get the partial derivative

$$\frac{\partial \tilde{\omega}_I(k)}{\partial \delta C} = \frac{k \sigma_I \sigma_M \mu_I \mu_M \left[2 \sigma_I (1 - \delta_C) \sigma_M \mu_I \mu_M + \rho_{IM} \sigma^2_M \mu^2_I - \rho_{IM} \sigma^2_I (1 - \delta_C)^2 \mu^2_M \right]}{[\sigma^2_M \mu^2_I - 2 \rho_{IM} \sigma_I (1 - \delta_C) \sigma_M \mu_I \mu_M + \sigma^2_I (1 - \delta_C)^2 \mu^2_M]^2}. \hspace{2cm} (3.6)$$

Looking at Equation (3.6), we now conveniently divide our analysis in two subcases.

**Case 2.1) Negatively correlated returns: $\rho_{IM} < 0$**

When the private company returns and the market portfolio returns are negatively correlated, then it is easy to show that $\frac{\partial \tilde{\omega}_I(k)}{\partial \delta C} > 0$. Therefore, the result in Equation (3.5) continues to hold, and the entrepreneur will underestimate the risk of her private company and thus overinvest in it.

**Case 2.2) Positively correlated returns: $\rho_{IM} > 0$**

When the private company returns and the market portfolio returns are positively correlated, the sign of $\frac{\partial \tilde{\omega}_I(k)}{\partial \delta C}$ is not straightforward. Imposing the condition $\frac{\partial \tilde{\omega}_I(k)}{\partial \delta C} = 0$, we find two stationary points. In the space $(\delta_C, \tilde{\omega}_I)$, the coordinates of these two points are

$$\left(\delta_C^-, \tilde{\omega}_I^+\right) = \left[1 - \frac{\mu_I}{\mu_I - \alpha} \left(1 + \sqrt{1 - \rho^2_{IM}}\right), \frac{k}{2 \mu_I} \frac{\rho^2_{IM}}{\sqrt{1 - \rho^2_{IM} - (1 - \rho^2_M)}}\right]$$

$$\left(\delta_C^+, \tilde{\omega}_I^-\right) = \left[1 - \frac{\mu_I}{\mu_I - \alpha} \left(1 - \sqrt{1 - \rho^2_{IM}}\right), \frac{k}{2 \mu_I} \frac{\rho^2_{IM}}{\sqrt{1 - \rho^2_{IM} - (1 - \rho^2_M)}}\right]. \hspace{2cm} (3.7)$$
3. Entrepreneurial under-diversification: over optimism and overconfidence

Since \((\delta_C^-, \omega_I^-)\) is a minimum and \((\delta_C^+, \omega_I^+)\) is a maximum, \(\frac{\partial \hat{\omega}_I(k)}{\partial \delta_C} > 0\) in the interval \((\delta_C^-, \delta_C^+)\), \(\delta_C^+ \leq 0\) outside.

We notice that \(\alpha > 0\) is a sufficient condition for \(\delta_C < 0\) to hold.\(^4\) However, we assume \(\delta_C \in [0, 1]\). Thus, \(\hat{\omega}_I\) reaches its minimum when \(\delta_C = 0\) and \(\hat{\omega}_I = \omega_I\).

When overconfidence approaches its limiting value, we find a particular weight:

\[
\lim_{\delta_C \to 1} \hat{\omega}_I(k) = \frac{k}{\mu_I}. \tag{3.8}
\]

When \(\delta_C \in (\delta_C^+, 1)\), \(\frac{\partial \hat{\omega}_I(k)}{\partial \delta_C} < 0\). In this case, the entrepreneur invests so much in her private company that, to meet the constraints of the portfolio selection problem, the weight in the well-diversified market portfolio needs to be negative.\(^5\) This case, when the entrepreneur sells the market portfolio short, is of little interest from an economic point of view. If we exclude this extreme case,\(^6\) there is no ambiguity on the sign of the derivative. Thus, we conclude that, in general, overconfidence leads to overinvestment in the entrepreneur’s private company and to portfolio under-diversification.

Figure 3.1 offers a graphical representation of all the aforementioned results.

Overconfidence implies suboptimal portfolio weights and a biased perception of portfolio risk. Since the perceived private company risk decreases with the level of overconfidence (i.e., \(\frac{\partial \hat{\omega}_I(k)}{\partial \delta_C} = -\sigma_I < 0\), whenever the perceived portfolio risk increases with the perceived private company risk (i.e., \(\frac{\partial \hat{\omega}_I(k)}{\partial \delta_C} > 0\) and overconfidence leads to overinvest in the company (i.e., \(\frac{\partial \hat{\omega}_I(k)}{\partial \delta_C} > 0\)),\(^7\) then it follows that \(\frac{\partial \hat{\omega}_I(k)}{\partial \delta_C} = \frac{\partial \hat{\omega}_I(k)}{\partial \delta_C} \frac{\partial \delta_C}{\partial \delta_I} \frac{\partial \delta_I}{\partial \delta_C} < 0\). Figure

\(^4\)\(\alpha > 0\) is a prerequisite to justify investments in private companies because it can be interpreted as a positive Net Present Value (NPV > 0). If financial markets are efficient, then positive NPV investments are feasible only for real investment projects (as the one that we are considering), and not for financial investment projects, for which the NPV is zero.

\(^5\)Using the constraint in the problem (3.1), we get \(\hat{\omega}_M = \frac{k - \frac{\hat{\omega}_I}{\mu_I}}{\mu_M}\). Therefore, when overconfidence reaches its limiting value and \(\hat{\omega}_I(k) = \frac{\Delta}{\mu_I}\), then \(\hat{\omega}_M = 0\). When \(\hat{\omega}_I(k) > \frac{\Delta}{\mu_I}\), as happens for \(\delta_C \in (\delta_C^+, 1)\), then \(\hat{\omega}_M < 0\). Imposing \(\hat{\omega}_M > 0\) means that only corner solutions are possible, where the maximum weight that the overconfident entrepreneur may invest in her company is given by \(\hat{\omega}_I(k) = \frac{\Delta}{\mu_I}\). Thus, the overconfident entrepreneur may even be frustrated by not being able, without short selling the market portfolio, to invest the desired amount of wealth in her private company.

\(^6\)In this extreme case, \(\hat{\omega}_I(k) > 0\), \(\hat{\omega}_M < 0\), and \(\rho_{1M} > 0\). Since the portfolio variance is \(\hat{\sigma}_P^2 = (1 - \delta_C)^2 + 2\hat{\omega}_I\hat{\omega}_M\rho_{1M}\sigma_I(1 - \delta_C)\sigma_M + \hat{\sigma}_I^2\sigma_M^2\), following an increase in \(\delta_C\), the change in the contribution of the variance term (the second term in the expression) to the portfolio variance is positive and dominates the change in the contribution of the variance term (the first term in the expression) which is negative. Therefore, the higher the overconfidence, the higher the perceived portfolio risk for a given \(\hat{\omega}_I\) and, thus, the lower the \(\omega_I\) chosen by the entrepreneur. Of course, this case is in opposition with what previous studies have shown, i.e., that overconfidence leads to underestimation of risk.

\(^7\)That is the most common case. Conversely, in case 2.2, when \(\delta_C \in (\delta_C^-, 1)\) (i.e., when the entrepreneur should sell the market portfolio short), \(\frac{\partial \hat{\omega}_I(k)}{\partial \delta_I} < 0\) but also \(\frac{\partial \hat{\omega}_I(k)}{\partial \delta_C} < 0\). Therefore, the inequality
3. Entrepreneurial under-diversification: over optimism and overconfidence

3.2 presents this result, showing the link between the perceived frontier of investments (the dashed line) and the weight in the entrepreneur’s private company. The first plot in Figure 3.2 shows the shift in the frontier caused by overconfidence, while the second one projects this shift in the private company weight. Note that the slope of the curve in the second plot is determined by \( \frac{\partial \tilde{\omega}_I(k)}{\partial \tilde{\sigma}_P} \).

All the aforementioned results describe the effects of underevaluating the actual risk due to overconfidence and are summed up in the following observation.

**Observation 10** Typically, overconfidence leads to overinvestment in the entrepreneur’s private company, \( \frac{\partial \tilde{\omega}_I(k)}{\partial \tilde{\sigma}_P} > 0 \), and to portfolio under-diversification. The only situation in which \( \frac{\partial \tilde{\omega}_I(k)}{\partial \delta_C} < 0 \) occurs is when \( \rho_{IM} > 0 \) and the level of overconfidence is particularly high, \( \delta_C \in (\delta_C^+, 1) \).

### 3.2.2. Over optimism

It is well known that the portfolio optimization problem is dual: either the entrepreneur minimizes the risk for a given portfolio expected return, or she maximizes the return for a given portfolio risk. Since overconfidence affects risk perception, in Section 3.2.1 
\( \frac{\partial \tilde{\omega}_I(k)}{\partial \tilde{\sigma}_P} < 0 \) still holds.
3. Entrepreneurial under-diversification: over optimism and overconfidence

Figure 3.2.: Frontier shift and overconfidence bias.
3. Entrepreneurial under-diversification: over optimism and overconfidence

we studied its effect on portfolio risk using a risk minimization approach, which keeps the expected return level fixed. Conversely, in this section we analyze the effect of over optimism on portfolio return using a return maximization approach, which holds the objective risk constant. Figure 3.3 shows the duality of the problem by representing the tangency conditions that identify the lower iso-risk (left plot) and the upper iso-return (right plot).

**Return maximization**

The duality of the problem allows us to consider return maximization as the solution for the entrepreneur optimization problem for a given value of portfolio risk, $\sigma_P^2 = s^2$. In such a setting, she faces the problem

$$
\begin{align*}
\max_{\omega} & \quad \omega^{\prime} \mu + r_F \\
\text{s.t.} & \quad \omega^{\prime} \Sigma \omega = s^2.
\end{align*}
$$

(3.9)

Setting up the Lagrangian and solving the problem, the optimal portfolio weights are\(^8\)

$$
\omega(s) = \frac{s \Sigma^{-1} \mu}{(\mu^{\prime} \Sigma^{-1} \mu)^{\frac{1}{2}}}.
$$

(3.10)

\(^8\)As the problem is quadratic, we obtain also a second solution with weights equal to minus those of Equation (3.10). We discard them since they are dominated ($\alpha > 0 \Rightarrow \omega_I(s) > 0$).
3. Entrepreneurial under-diversification: over optimism and overconfidence

The weight on the private company is

\[
\omega_I(s) = \frac{s \left( \sigma_M^2 \mu_I - \sigma_{IM} \mu_M \right)}{\left[ (\sigma_I^2 \sigma_M^2 - \sigma_{IM}^2) \left( \sigma_M^2 \mu_I^2 - 2 \sigma_{IM} \mu_I \mu_M + \sigma_M^2 \mu_M^2 \right) \right]^{\frac{1}{2}}} \]

\[
= \frac{s \sigma_M}{\sigma_{IM}^2} \left[ (\sigma_I^2 \sigma_M^2 - \sigma_{IM}^2) \left( \sigma_M^2 \mu_I^2 - 2 \sigma_{IM} \mu_I \mu_M + \sigma_M^2 \mu_M^2 \right) \right]^{\frac{1}{2}}.
\]

(3.11)

Over optimism-driven under-diversification

Over optimism causes the entrepreneur to overestimate the actual return of the investment in her private company. In this case, the biased expected return, indicated by \( \tilde{\mu}_I \), is larger than the actual expected return, i.e., \( \tilde{\mu}_I > \mu_I \). We model \( \tilde{\mu}_I \) as

\[
\tilde{\mu}_I = \frac{\mu_I}{1 - \delta_O}, \quad \delta_O \in [0, 1),
\]

(3.12)

where \( \delta_O \) is the over optimism parameter, which ranges from 0 (no over optimism) to 1 (maximum over optimism). When \( \delta_O \) tends to 1, then \( \tilde{\mu}_I \) tends to infinity.

In order to see the variation of the portfolio weight in her private company in case of over optimism, we define \( \tilde{\omega}_I(s) \) as the \( \omega_I(s) \) of Equation (3.11) with \( \tilde{\mu}_I \) in place of \( \mu_I \).

Using this definition, we get the partial derivative

\[
\frac{\partial \tilde{\omega}_I(s)}{\partial \delta_O} = \frac{s \mu_I^2 \sigma_M^2 \left( \sigma_I^2 - \sigma_{IM}^2 \right)^{\frac{1}{2}}}{\left( \sigma_M^2 \mu_I^2 - 2 \sigma_{IM} \mu_I \mu_M + \sigma_M^2 \mu_M^2 \right)^{\frac{1}{2}}}.
\]

(3.13)

Since \( \frac{\partial \tilde{\omega}_I(s)}{\partial \delta_O} > 0 \), the over optimist entrepreneur tends to overinvest in her private company and to be under-diversified. The over optimism bias is

\[
b_o = \tilde{\omega}_I(s) - \omega_I(s) > 0.
\]

(3.14)

The limit case for over optimism identifies a particular weight:

\[
\lim_{\delta_O \to 1} \tilde{\omega}_I(s) = \frac{s \sigma_M}{\left( \sigma_I^2 \sigma_M^2 - \sigma_{IM}^2 \right)^{\frac{1}{2}}}.
\]

(3.15)

Over optimism implies suboptimal portfolio weights and a biased perception of portfolio

\[\footnote{The justifications of entrepreneur’s under-diversification based on non-pecuniary benefits as the desire for control can be modeled by varying \( \mu_I \) as well. In that case, the “biased” \( \mu_I \) would incorporate the value of non-pecuniary benefits.}

\[\footnote{Choosing \( \delta_O \in (-\infty, 1) \), we would allow for under optimism.}

37
3. Entrepreneurial under-diversification: over optimism and overconfidence

Figure 3.4.: Frontier shift and over optimism bias.

return. Since the perceived portfolio return increases with the perceived private company return (i.e., $\frac{\partial \tilde{\mu}_P}{\partial \mu_I} > 0$), the perceived private company return increases with the level of over optimism (i.e., $\frac{\partial \tilde{\mu}_I}{\partial \delta_O} = \frac{\mu_I}{(1-\delta_O)^2} > 0$) and the weight in the private company increases with the level of over optimism (i.e., $\frac{\partial \tilde{\omega}_I(s)}{\partial \delta_O} > 0$), then it follows that $\frac{\partial \tilde{\omega}_I(s)}{\partial \tilde{\mu}_P} = \frac{\partial \tilde{\omega}_I(s)}{\partial \delta_O} \frac{\partial \delta_O}{\partial \delta_O} \frac{\partial \mu_I}{\partial \mu_I} > 0$. This result is presented in Figure 3.4, which shows the link between the perceived frontier of investments (dashed) and the weight in the private company.

The plot on the right of Figure 3.4 shows the shift in the frontier caused by over optimism; the plot on the left projects this shift on the private company weight. Note that the slope of the curve in the plot on the left is determined, as explained above, by $\frac{\partial \tilde{\omega}_I(s)}{\partial \tilde{\mu}_P}$.

All the results above describe the effects of the overestimation of the actual return due to over optimism and can be summarized in the following observation.

**Observation 11** Over optimism always leads to overinvestment in the entrepreneur’s private company, $\frac{\partial \tilde{\omega}_I(s)}{\partial \delta_O} > 0$, and to portfolio under-diversification.

3.3. Implicit overconfidence and over optimism levels

To understand the effects of overconfidence and over optimism on the weight invested by the entrepreneur in her private company, we consider a couple of numerical examples.

Using the estimates in tables 2 and 4 of Kerins et al. [2004] who analyzed a sample of IPOs in technologically-oriented industries, we choose the following set of parameters:
3. Entrepreneurial under-diversification: over optimism and overconfidence

\[ \sigma_I = 1.204, \sigma_M = 0.162, \rho_{IM} = 0.195, \mu_I = 0.535 \text{ and } \mu_M = 0.06. \]

Furthermore, we set \( k = 0.3 \) and \( s = 0.566. \) Based on this set of parameters, in Table 3.1 we calculate \( \tilde{\omega}_I, \tilde{\omega}_M, \text{ and } \tilde{\omega}_F, \) by varying the level of overconfidence (Panel A) or the level of over optimism (Panel B).

Table 3.1.: Implicit overconfidence and over optimism levels; parameters from Kerins et al. [2004]

<table>
<thead>
<tr>
<th>Panel A: Overconfidence effects on risk minimization</th>
<th>( \delta_C )</th>
<th>0.0</th>
<th>0.2</th>
<th>0.4</th>
<th>0.6</th>
<th>0.8</th>
<th>( \approx 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tilde{\omega}_I )</td>
<td>0.343</td>
<td>0.412</td>
<td>0.480</td>
<td>0.534</td>
<td>0.563</td>
<td>0.561</td>
<td></td>
</tr>
<tr>
<td>( \tilde{\omega}_M )</td>
<td>1.943</td>
<td>1.328</td>
<td>0.723</td>
<td>0.235</td>
<td>-0.024</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>( \tilde{\omega}_F )</td>
<td>-1.286</td>
<td>-0.740</td>
<td>-0.203</td>
<td>0.230</td>
<td>0.461</td>
<td>0.439</td>
<td></td>
</tr>
<tr>
<td>( \tilde{\mu}_P )</td>
<td>0.300</td>
<td>0.300</td>
<td>0.300</td>
<td>0.300</td>
<td>0.300</td>
<td>0.300</td>
<td></td>
</tr>
<tr>
<td>( \tilde{\sigma}_P )</td>
<td>0.566</td>
<td>0.487</td>
<td>0.387</td>
<td>0.267</td>
<td>0.135</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Over optimism effects on return maximization</th>
<th>( \delta_O )</th>
<th>0.0</th>
<th>0.2</th>
<th>0.4</th>
<th>0.6</th>
<th>0.8</th>
<th>( \approx 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tilde{\omega}_I )</td>
<td>0.343</td>
<td>0.383</td>
<td>0.421</td>
<td>0.452</td>
<td>0.472</td>
<td>0.479</td>
<td></td>
</tr>
<tr>
<td>( \tilde{\omega}_M )</td>
<td>1.943</td>
<td>1.544</td>
<td>1.058</td>
<td>0.498</td>
<td>-0.103</td>
<td>-0.694</td>
<td></td>
</tr>
<tr>
<td>( \tilde{\omega}_F )</td>
<td>-1.286</td>
<td>-0.927</td>
<td>-0.479</td>
<td>0.050</td>
<td>0.630</td>
<td>1.215</td>
<td></td>
</tr>
<tr>
<td>( \tilde{\mu}_P )</td>
<td>0.300</td>
<td>0.298</td>
<td>0.289</td>
<td>0.272</td>
<td>0.247</td>
<td>0.215</td>
<td></td>
</tr>
<tr>
<td>( \tilde{\sigma}_P )</td>
<td>0.566</td>
<td>0.566</td>
<td>0.566</td>
<td>0.566</td>
<td>0.566</td>
<td>0.566</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>( k )</th>
<th>( s )</th>
<th>( \mu_I )</th>
<th>( \mu_M )</th>
<th>( \sigma_I )</th>
<th>( \sigma_M )</th>
<th>( \rho_{IM} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.300</td>
<td>0.566</td>
<td>0.535</td>
<td>0.060</td>
<td>1.204</td>
<td>0.162</td>
<td>0.195</td>
</tr>
</tbody>
</table>

For moderate levels of overconfidence, an increase in \( \delta_C \) leads to an increase in \( \tilde{\omega}_I \). Note that even low levels of overconfidence cause a severe overinvestment in the private company (e.g., when \( \delta_C = 0.2 \), the weight in the private company is about 20% larger than it should be if the entrepreneur was not overconfident). \(^{12}\)

When the level of overconfidence tends to its limiting value (i.e., \( \delta_C \to 1 \)), \( \tilde{\omega}_I = 0.561 \) and \( \tilde{\omega}_M = 0 \). When \( \tilde{\omega}_I > 0.561 \) (i.e., when \( \delta_C \) is about 0.8), \( \tilde{\omega}_M < 0 \), implying that the entrepreneur is short selling the market portfolio. When selling short is not allowed, \( \tilde{\omega}_I \) is thus equal to 0.561.

Considering the over optimism bias, we note that, as with the case of overconfidence, for particularly high levels of over optimism (e.g., \( \delta_O \approx 0.8 \) or higher), \( \tilde{\omega}_M < 0 \). Instead, differently from what happens in case of overconfidence, an increase in \( \delta_O \) always causes

\(^{11}\)Choosing \( k \), we implicitly determine \( s \).

\(^{12}\)This result comes from the comparison between the values assumed by \( \tilde{\omega}_I \) when \( \delta_C \) varies. When \( \delta_C = 0 \), \( \tilde{\omega}_I = 0.343 \). When \( \delta_C = 0.2 \), \( \tilde{\omega}_I = 0.412 \). Then, the percentage change is \( \frac{0.412 - 0.343}{0.343} \times 100 = 20.12\% \).
3. Entrepreneurial under-diversification: over optimism and overconfidence

an increase in $\tilde{\omega}_I$. We also underline that, in this numerical example, the consequences of over optimism on overinvestment are less salient than those of overconfidence.\textsuperscript{13}

Since the value of $\sigma_I$ in the previous numerical example is particularly high,\textsuperscript{14} in Table 3.2 we present another numerical example using a new set of parameters that we may consider as an “average case”, in contrast with the “extreme case” presented above. In particular, we choose $\sigma_I = 0.4$, $\sigma_M = 0.2$, $\rho_{IM} = 0.2$, $\mu_I = 0.15$, $\mu_M = 0.06$, $k = 0.12$, and $s = 0.273$. Even if the latter results quantitatively differ from the ones presented in Table 3.1, we notice that their qualitative pattern does not change using this new set of parameters.

Table 3.2.: Implicit overconfidence and over optimism levels; our assumptions

| Panel A: Overconfidence effects on risk minimization |
|----------------|-----------|-----------|-----------|-----------|-----------|
| $\delta_C$     | 0.0       | 0.2       | 0.4       | 0.6       | 0.8       |
| $\tilde{\omega}_I$ | 0.509     | 0.605     | 0.696     | 0.768     | 0.805     | 0.800     |
| $\tilde{\omega}_M$ | 0.727     | 0.488     | 0.259     | 0.079     | -0.013    | 0.000     |
| $\tilde{\omega}_F$ | -0.236    | -0.093    | 0.045     | 0.153     | 0.208     | 0.200     |
| $\bar{\mu}_p$   | 0.120     | 0.120     | 0.120     | 0.120     | 0.120     | 0.120     |
| $\bar{\sigma}_p$ | 0.273     | 0.234     | 0.185     | 0.127     | 0.064     | 0.000     |

| Panel B: Over optimism effects on return maximization |
|----------------|-----------|-----------|-----------|-----------|
| $\delta_O$     | 0.0       | 0.2       | 0.4       | 0.6       | 0.8       |
| $\tilde{\omega}_I$ | 0.509     | 0.565     | 0.618     | 0.660     | 0.687     | 0.696     |
| $\tilde{\omega}_M$ | 0.727     | 0.571     | 0.383     | 0.169     | -0.057    | -0.278    |
| $\tilde{\omega}_F$ | -0.236    | -0.136    | 0.000     | 0.170     | 0.369     | 0.582     |
| $\bar{\mu}_p$   | 0.120     | 0.119     | 0.116     | 0.109     | 0.100     | 0.088     |
| $\bar{\sigma}_p$ | 0.273     | 0.273     | 0.273     | 0.273     | 0.273     | 0.273     |

Parameters

<table>
<thead>
<tr>
<th>$k$</th>
<th>$s$</th>
<th>$\mu_I$</th>
<th>$\mu_M$</th>
<th>$\sigma_I$</th>
<th>$\sigma_M$</th>
<th>$\rho_{IM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.120</td>
<td>0.273</td>
<td>0.150</td>
<td>0.060</td>
<td>0.400</td>
<td>0.200</td>
<td>0.200</td>
</tr>
</tbody>
</table>

Note that, if we assume that our model is sufficiently adequate to describe entrepreneurs’ portfolio choices under overconfidence and/or over optimism, Tables 3.1 and 3.2 can also be read bottom-up to calculate the “implicit” overconfidence and over optimism levels given a set of parameters and an observed $\tilde{\omega}_I$. For example, based on Table 3.1 – Panel A (i.e., using a risk minimization perspective with $k = 0.3$), after observing $\tilde{\omega}_I = 0.48$, we can conclude that $\delta_C = 0.4$.

Calculating the implicit overconfidence and over optimism levels is useful to determine

\textsuperscript{13}This claim can be proved calculating the percentage change in $\tilde{\omega}_I$ when $\delta_O$ varies, and showing that is lower than the ones calculated above for $\delta_C$.

\textsuperscript{14}Kerins et al. [2004] referred to companies going public before and during the Internet bubble.
3. Entrepreneurial under-diversification: over optimism and overconfidence

how these behavioral biases may affect entrepreneurial decisions. In other words, not only the presence of these biases, but also their magnitude impacts the entrepreneurs’ portfolio weight in their private companies and, consequently, the level of idiosyncratic risk they bear due to sub-optimal decisions in portfolio formation.

3.4. Conclusions

Previous findings in the literature show that entrepreneurs tend to overinvest in their private company, bearing higher levels of idiosyncratic risk, with respect to what would be optimal. We propose a possible explanation for this sub-optimal behavior, based on overconfidence and over optimism. Our proposal complements other explanations which rely on non-pecuniary benefits, such as the entrepreneurs’ desire of control.

We present a model that allows us not only to show, but also to measure how behavioral biases affect the fundamental variables of the risk-return analysis à la Markowitz and lead entrepreneurs to choose sub-optimal portfolio weights in their private company and hold under-diversified portfolios. Overconfidence leads to underestimation of the risk associated with the entrepreneur’s private company, while over optimism to the overestimation of its expected return.

Our theoretical model contributes to the literature on risk perception, but it could also be modified to consider other motivations for the entrepreneurial behavior, such as the desire for control, since the latter, like over optimism, leads to overestimation of the expected return on the private company.

Since we insert both overconfidence and over optimism as parameters in our model, we perform a simulation analysis to ascertain how much the entrepreneur’s decision for investing in her private company is affected by variations in the two behavioral biases. As our theoretical model predicts, the entrepreneur invests more in her company as overconfidence and over optimism increase. The simulation can also be used, given other parameters and the weight in the private company, to calculate the implicit levels of entrepreneurial overconfidence and over optimism.
Part II.

Intangible assets
4. Intellectual property rights and imitation in developing countries

4.1. Introduction

Innovators, mostly living in rich economies, have vested interests in building IPR (Intellectual Property Rights) institutions. Indeed, innovators seek protection to recoup their investments and appropriate the returns. Since the birth of the WTO and the establishment of the TRIPs (Trade Related Aspects on Intellectual Property Rights) agreement, international pressure on DC (Developing Countries) to use IP (Intellectual Property) has noticeably increased. Many DC make a large use of imitation to sell products invented by others at lower prices. Imitations are of lesser quality in comparison to the original products. One argument to convince DC to accept the new setting, with stringent IP, is an increase in imports of the best products. In addition, letting aside imitative process and making use of IP, DC would participate in the official market and they could gather throughout the world the monopoly rents coming from their production of IP.

A policymaker’s decision on the amount of IPR protection depends on weighing the benefits and costs. The intuition behind this paper is that there is a moment during the development process when a country is ready to enter this official market. Entering beforehand could be not convenient for a poor country and a wrong decision can be detrimental for the development process. At the beginning of this development process, indeed, competition is mainly supported by imitation and adaptive innovation. As a country reaches higher technological levels and as its demand targets high-quality and differentiated products, a larger share of domestic firms advocates IPR. Ultimately, whenever a country’s GDP per capita approaches the highest levels, IPR compliance steps up considerably.

As always, history is a good teacher to consult. The experience of the USA and other advanced countries about economic development is that they initially did not
4. Intellectual property rights and imitation in developing countries

respect IP.¹ At the beginning of the American industry, markets were functioning with the infringement of foreign IP. Only in a subsequent moment, when the US economy became a mature one, IP was utilised as a pervasive instrument, and this happened long after UK started to adopt IP.² The Chinese experience seems also to confirm this intuition, as suggested in Yang and Clarke [2005].³ Today and not before, having left the DC club, China is ready to enter into the TRIPs agreement. Only now the biggest country in the world seems ready to give up the wild usage of imitation of occidental products and ready to shape the incentives of the R&D sector in the official way.

A point stemming from usual economic reasoning can be made easily to enrich the perspective over these observed economic phenomena. Following a Coasian bargaining process and assuming that DC are not less rational than developed ones, there are not economic arguments to support the thesis that DC are not choosing what is best for them. If the adoption of IPR, fostering FDI (Foreign Direct Investment) and licensing, would better the technological transfer [Yang and Maskus, 2009] compared to imitative process there is no reason why DC would not choose unilaterally to adopt IPR from the beginning of their development process. Globerman [1988] pointed out that allowing imitations and piracy might discourage FDI but at the same time “encourages the creation of indigenous industry by supplying local producers with the financial capital, experience and trained labor force required to become legitimate producers in the future” (emphasis added).

Having these intuitions in mind, I develop a model where the choice to accept the TRIPs agreement is based on the incentive to participating into the official international market. The model represents an attempt to explore and develop this intuition. Since the focus is to understand the incentives of DC, I make a few assumptions that, keeping simple and sufficiently realistic the technical formulation, allow to sharpen the division of profits on which the choice of TRIPs agreement participation should be made. The empirical analysis that follows, taking into account panel structure of the data, relates the adopted level of IPR to the amount of patenting and the imitative ability, taking

¹Chang [2001], at page 293, underlines that “Historical evidence shows that ... in the early days of industrial development in the now-advanced countries, IPR, especially other countries’ IPR, were not well respected. Compared with the developed countries of yesteryears, the contemporary developing countries seem to be behaving much better in many ways. And if that is the case, it seems unfair to ask the modern-day developing countries to behave to a standard that was not even remotely observed when the now-advanced countries were at the similar, or even more advanced, stages of development.”

²The adoption of IP laws in US started with the Patent Act of 1790, approximately one century after UK, and extended progressively to more and more areas of business.

³After a thorough overview of the development of the IP system in China, in the conclusion, Yang and Clarke [2005] explicitly state: “There can be no doubt that China has made more significant progress on the protection of IP in recent years. During the past decade or so, a new legal mechanism has been implemented.”
4. Intellectual property rights and imitation in developing countries

into account GDP per capita, population, WTO membership, openness to trade, FDI, and economic freedom.

The remainder of the paper is organised as follows. In Section 4.2, I briefly review the most relevant literature. In Section 4.3, I introduce the model and discuss its main implications for the South. In Section 4.4, I empirically validate the theoretical results by means of panel data. Section 4.5 concludes the paper.

4.2. Related literature

There are many different contributions that have dwelled on the foremost question of the paper. Whether and when it is profitable for DC to implement a system of IP has been addressed directly and indirectly by economic literature since the eighties.

Krugman [1979], starting from the concept of the “product cycle” [Vernon, 1966], shows how the trade between the North and the South can stem from a temporal lag between the innovation in the North and the diffusion of the technology in the South. The result is the North exporting new products and importing old products. Grossman and Helpman [1991a,b] work in the same direction featuring endogenous innovation and endogenous technology transfer. Diwan and Rodrik [1991] state that patents should be protected in the South in order to promote the development of technologies more appropriate to the South. The paper by Helpman [1993], developing a dynamic general equilibrium framework, finds that the South does not benefit from tight IP. This finding is robust with respect to examined variations such as the presence of FDI.

Starting from the framework developed in these seminal contributions, many papers investigate some related issues. Lai [1998] discusses the importance of the channel of production transfer from the North to the South. On the same theoretical ground, Glass and Wu [2007] notice that when innovations consist of new varieties, stronger Southern IPR protection encourages FDI and innovation, but when innovations consist of higher quality levels, FDI and innovation can fall. Developing a similar framework, Dinopoulos and Segerstrom [2010] find that a more stringent IPR protection in the South raises permanently the rate of international technology transfer within multinational firms and generates a temporary increase in the Northern innovation rate. However, finding an optimal level of IPR enforcement in the South is a matter of balance between the attraction of multinationals’ technology and a limitation of rent transfer to the high-income developed countries [Markusen, 2001]. Currie et al. [1999] delineate three phases of Southern development in which the very same IP policy can generate different outcomes. Yang and Maskus [2001] study the effects of stronger IP in the South on the incentives of firms.
in the North to innovate and to license advanced technologies. Glass and Saggi [2002] study a setting where, in addition to innovation and imitation, also FDI is endogenous. Their work shows that the results can be very different depending on whether IPR protection consists of an increase in the cost of imitation or of an exogenous decrease in the imitation intensity. On the same line, Parello [2008] stresses that strengthening IPR hurts the rate of technology transfer to the South via a permanent fall in the long-run rate of imitation. Lai and Qiu [2003] find that the Nash equilibrium IPR protection standard of the South is naturally weaker than that of the North. Grossman and Lai [2004] consider the simultaneous choice of IPR protection by trade partners finding that the optimal index of patent protection may be independent of, or even decreasing in, the size of the economy; in addition, they find that harmonization of patent policies is neither necessary nor sufficient for global efficiency. Chen and Puttitanun [2005] conduct a theoretical and empirical analysis on the optimal level of Southern IPR that balances the trade-off between facilitating the imitation of Northern advanced technologies and providing incentives for domestic innovations. They find a U-shaped curve of a DC’s IPR with respect to its economic development. Whether the U-shaped is a good picture of reality or not, a large part of these contributions seem to pinpoint that only after a stage in the development process it is profitable to line up the IPR to the level of the most developed countries.

Many authors focus on the empirical side of this issue. Smith [1999, 2001] finds a significant empirical relation among imitative ability, patent protection, and trade. Evidence to confirm this view is presented also in the paper by Co [2004]. The importance to consider imitative ability with respect to IPR protection is finally stressed by Falvey et al. [2006, 2009]. In their empirical works, following the intuition by Thompson and Rushing [1996], they strongly support the existence of a threshold level in development after which patent protection changes its impact on economies. Finally, also Ginarte and Park [1997] suggest a threshold effect, a critical size of an innovating sector before a country has an incentive to provide patent rights.

Perhaps, the work to which the theoretic model of this paper is the closest is that by Chin and Grossman [1991]. They propose a North-South model finding that the interests of the North and the South generally conflict in the matter of protection of IP. The channel through which patent protection can enhance the welfare of the South is a reduction of production costs due to innovation. The counterbalancing force is that an

---

4Even though patents have complex effects on imitation costs [Mansfield et al., 1981], I will simplify this issue in the model, where respecting IPR means no possibility of imitation whatsoever for the South.
enforcement of patent protection mitigates oligopolistic competition, to the detriment of welfare in both countries. On the contrary, in the current work, the products of the North and the South are of different quality, consumers are heterogeneous, and imitative ability, explicitly modeled and empirically proxied, directly affects the quality level of the product of the South not enforcing IPR. By explicitly considering the division of profits, assumed to be connected to R&D capacity, I am finally able to investigate how the choice of the South to respect IP is influenced by its imitative ability.

4.3. The model

There are two countries, $N$ (North) and $S$ (South), and a company that has no capacity constraint, $M$. The company represents the official way of world production. In $S$, it can be purchased either the official product of the company with quality $Q_M$ or a product obtained locally through imitation with quality $Q_I \equiv \rho Q_M$; where $0 < \rho < 1$ is the imitative ability of the country $S$. Production costs are assumed to be equal to 0.\footnote{I am deeply indebted to Bordoy [2002] as I started to build my model from her Ph.D. work.}

The consumers’ utility functions in the two countries are as follows:

$$U^S(Q,p) \equiv \theta Q - p, \quad U^N(Q,p) \equiv 2\theta Q - p;$$

where $\theta$ is the taste parameter uniformly distributed in the interval [0,1]. This formulation of utility functions gives us the intuition that, for the same level of quality and the same position in the taste dimension inside the segment [0,1], a consumer in $N$ is willing to pay the double than a consumer in $S$. It is like saying that in $S$ people are less exigent about quality level or that their opportunity cost is smaller in absolute terms.

In order to investigate the incentives of $S$ to join the TRIPs agreement, respecting IPR and giving up the possibility to imitate ($\rho = 0, Q_I = 0$), let us contrate now on $S$.

4.3.1. Company’s and imitated products in $S$

The local industry is competitive, so $p_I = 0$.

Let’s call the indifferent consumer $\tilde{\theta}$.

$$\tilde{\theta} Q_M - p^S_M = \tilde{\theta} Q_I - p_I, \quad \tilde{\theta} = \frac{p^S_M}{Q_M - Q_I}.\footnote{By having put the production costs at the same level in $N$ and in $S$, $Q$, can be interpreted as an efficiency index of production, for example it can represent the ratio between the quality of the product and the production cost.}$$

5
4. Intellectual property rights and imitation in developing countries

The quality differential is $\Delta \equiv Q_M - Q_I = Q_M(1 - \rho)$.

The demands the company and local producers face are

$$D^S_M = \max \left\{ 1 - \frac{p^S_M}{\Delta}, 0 \right\}, \quad D_I = \min \left\{ \frac{p^S_M}{\Delta}, 1 \right\}. $$

The government of $S$ has at its disposal the tariff $\lambda \geq 0$ as instrument to protect local producers and to maximise its total welfare. Thus, $\lambda$ is the degree of protection of $S$, and $D^S_M \lambda$ is the total government revenue in $S$.

The timing of the game is: (1) the government of $S$ fixes $\lambda$; (2) the company fixes $p^S_M$; (3) the consumers in $S$ choose, so $D^S_M$ and $D_I$ are obtained. In order to find the equilibrium values I proceed by backward induction, hence addressing (3), (2), (1).

Choosing $p^S_M$, the company maximises its profit, that is

$$\Pi^S_M = (p^S_M - \lambda) \left[ \max \left\{ 1 - \frac{p^S_M}{\Delta}, 0 \right\} \right].$$

Neither $p^S_M = \Delta$ (for which $D^S_M = 0$) nor $p^S_M = 0$ (for which $D^S_M = 1$) are optimal strategies. The equilibrium value is $p^S_M \in (0, \Delta)$.

Solving the maximization problem, I find

$$p^S_M = \frac{\Delta + \lambda}{2}, \quad D^S_M = \frac{\Delta - \lambda}{2 \Delta}, \quad D_I = \frac{\Delta + \lambda}{2 \Delta}. $$

Now, I am able to obtain the consumer’s surplus.

$$CS^S(\lambda) = \int_0^{\frac{\Delta + \lambda}{2 \Delta}} \theta Q_I d\theta + \int_{\frac{\Delta + \lambda}{2 \Delta}}^1 \left( \theta Q_M - \frac{\Delta + \lambda}{2} \right) d\theta = \frac{Q_M}{2} + \frac{\lambda^2 - 2 \lambda \Delta - 3 \Delta^2}{8 \Delta}. $$

The government’s revenue is equal to $\lambda \frac{\Delta + \lambda}{2 \Delta}$, so the total welfare of $S$ is

$$TW^S(\lambda) = \frac{Q_M}{2} - \frac{3 \lambda^2 - 2 \lambda \Delta + 3 \Delta^2}{8 \Delta}. $$

The maximization of $TW^S(\lambda)$ brings to the optimal level of industry protection in $S$, that is $\lambda^* = \frac{\Delta}{3}$.

**Observation 12** The level of industry protection (optimal tariff) is negatively related to the country’s imitative ability. The higher $\rho$, the higher $Q_I$, the smaller $\Delta$, and the smaller $\lambda^*$.

Observation 12 is coherent with the intuition that stronger economies need less protection.
4. Intellectual property rights and imitation in developing countries

than weaker ones. The value of $\lambda^*$ is such that $D_M^S$, $D_I$ do not depend on $\Delta$: $D_M^S = \frac{1}{3}$, $D_I = \frac{2}{3}$. The government uses the tariff to obtain always the same division of the demand between the company and the local sector.

Now, the precise value of the total welfare, useful in the following, is computed.

$$TW^S = \frac{Q_M}{2} - \frac{3\Delta^2 - \frac{2}{3}\Delta^2 + \frac{1}{3}\Delta^2}{8\Delta} = \frac{Q_M}{6} + \frac{Q_I}{3} = \frac{Q_M}{6} + \rho\frac{Q_M}{3}. \quad (4.1)$$

4.3.2. The TRIPs agreement

When $S$ signs the TRIPs agreement, it gives up the possibility to imitate the product of $M$. Thus, imitative ability goes to 0 ($\rho = 0$, $Q_I = 0$).

Respecting the IPR international setting is often coupled with the fulfillment of the WTO conditions about free trade. The two international agreements are deeply connected and the discussion of one issue is usually paired to the discussion of the other one. As a result, it seems likely that the governments that respect the TRIPs agreement also refrain from using tariffs. Thus, I make here the additional assumption of $\lambda = 0$.

In this new setting the company has not to compete with local producers. The demand the company serves is composed by all the consumers that get a positive utility choosing to purchase the product. The new indifferent consumer is easily obtained: $\tilde{\theta}Q_M - p_M^S = 0$.

By one side, the price chosen by the company is such that $p \leq \theta_{\max}Q_M = Q_M$, because otherwise there would be no consumer with $U^S \geq 0$, and so $D_M^S = 0$.\footnote{This constraint was not binding in the preceding subsection because $U^S = \theta Q - p$, $U^S \geq 0$, $\theta Q \geq p$, and there is always the possibility of buying the local product at $p_I = 0$, so $\theta Q_I \geq 0$, $\theta_{\min} = 0$, $0 = 0$.} By the other side, in order to sell at least one product the company has to choose the price such that $p \geq \theta_{\min}Q_M = 0$. Thus, $p_M^{S*} \in (0, Q_M)$, the demand of the company is $D_M^S = 1 - \frac{p_M^S}{Q_M}$, and its profit is $\Pi_M^S = (p_M^S - 0)(1 - \frac{p_M^S}{Q_M})$.

The solution to the company’s maximization problem is:

$$\frac{d\Pi_M^S}{dp_M^S} = 0, \quad 1 - \frac{p_M^S}{Q_M} = \frac{p_M^S}{Q_M}, \quad p_M^{S*} = \frac{Q_M}{2}.$$ 

Thus,

$$D_M^S = 1 - \frac{Q_M}{2Q_M} = \frac{1}{2}, \quad \tilde{\theta} = \frac{1}{2}.$$
4. Intellectual property rights and imitation in developing countries

The consumer surplus is

\[ CS^S = \int_{\frac{1}{2}}^{1} \left( \theta Q_M - \frac{Q_M}{2} \right) d\theta = \frac{Q_M}{8}. \]

The company’s equilibrium profit for the products sold in \( S \) is

\[ \Pi^S_M = \frac{Q_M}{2} \cdot \frac{1}{2} = \frac{Q_M}{4}. \]

The government’s revenue is zero because there is no tariff.

In \( N \), the company gets the profit \( \Pi^N_M = p^N_M D^N_M \). Since the utility function is different, \( U^N(Q, p) = 2\theta Q - p \), the indifferent consumer is \( 2\theta Q_M - p^N_M = 0 \), \( \tilde{\theta} = \frac{p^N_M}{2Q_M} \).

The boundaries of the company’s price in \( N \) are obtained as usual: \( U^N \geq 0 \), \( 2\theta Q_M \geq p^N_M \), \( \theta_{\text{max}} = 1 \), \( p^N_M \leq 2Q_M \); \( p^N_M \geq 0 \), so \( p^N_M \in (0, 2Q_M) \).

The maximization plan of the company follows.

\[ \frac{d\Pi^N_M}{dp^N_M} = 0, \quad \left[ 1 - \frac{p^N_M}{2Q_M} \right] = \frac{1}{2Q_M} p^N_M, \quad p^N_M = Q_M. \]

Consequently, the equilibrium values in \( N \) are

\[ D^N_M = 1 - \frac{Q_M}{2Q_M} = \frac{1}{2}, \quad \tilde{\theta} = \frac{1}{2}, \]

\[ CS^N = \int_{\frac{1}{2}}^{1} (2\theta Q_M - Q_M) d\theta = \frac{Q_M}{4}, \quad \Pi^N_M = \frac{Q_M}{2}. \]

4.3.3. What is better for the South

In the model, the company represents the official way of world production. Nowadays, companies are increasingly delocalised and the ownership structure is scattered among different countries. Thus, I assume that the company’s profit is divided between the two countries.

The question I am interested in is when it is profitable, hence advisable, for a DC to enter into the WTO respecting the TRIPs agreement, at what level of the profit gained by the official system, in my case \( M \), it is convenient to give up the system based on imitation. In order to answer this question, I need to find \( \tilde{x} \), namely the value of \( x = \frac{\Pi_M}{\Pi_M} \) such that \( S \) prefers to respect the TRIPs agreement. Evidently, the total profit of the company is the sum of what it earns in \( N \) and what it earns in \( S \), \( \Pi_M = \Pi^N_M + \Pi^S_M \). On the contrary, \( \Pi_S \) is how much \( \Pi_M \) gets \( S \).
After the TRIPs agreement, the total welfare in $S$ is

$$TW^S_T = CS^S + \Pi_S = CS^S + x (\Pi^N_M + \Pi^S_M) = \frac{Q_M}{8} + x \left( \frac{Q_M}{2} + \frac{Q_M}{4} \right) = \frac{Q_M}{8} + x \frac{3}{4} Q_M. \quad (4.2)$$

By (4.1) and (4.2), the value $\hat{x}$ that respects the condition $TW^S = TW^S_T$, is obtained and the next observation follows.

**Observation 13** $S$ prefers to sign and respect the TRIPs agreement when its share of the company’s profits is above the threshold level

$$\hat{x} = \frac{1}{18} + \frac{4}{9} \rho. \quad (4.3)$$

In order to better understand the incentives to stop imitation, it is possible to think in terms of the development of a country. At the beginning of the development process, having no technological expertise for products of high quality, $x = 0$. In this situation, looking at (4.1) and (4.2), $TW^S > TW^S_T$; without the possibility to share the profits obtained through $M$, it is strongly preferred to stick to the imitation regime. As $x$ increases, the incentives to adhere to the TRIPs agreement grow and, when $x$ reaches the value of $\hat{x}$, the government of $S$ is indifferent. When $x = 1$ (a situation similar to $(1 - x) = 1$ for $N$ at the beginning of the development process of $S$), it holds that $TW^S < TW^S_T$; it is strongly preferred to adhere to the TRIPs agreement.

Looking at (4.3), one can see that as imitative ability $\rho$ decreases $\hat{x}$ decreases too.

**Observation 14** The less a country is able to imitate, the smaller is the share in the profits of the official system that makes the country willing to accept to give up the imitation of the company’s product.

The last two observations will be empirically tested in the next section. Specifically, it will be tested if there is a positive relation between a country’s share in the official profits and IPR adoption and if this relation depends negatively on imitative ability.

A possible additional assumption would be to divide the company’s profit according to the country’s participation to the construction of the quality level of the product. That could be the amount of IP produced in any country, the IP on the knowledge necessary to obtain the product with that quality level. That is,

$$xQ_M = Q_S, \quad (1 - x)Q_M = Q_M - Q_S = Q_N.$$
Thus, in addition to being the share of the company’s profit, $x$ would be the share of the creation of the quality $Q_M$ through the R&D performed by $S$. Symmetrically, $(1 - x)$ would be the share of $N$.

By this perspective, there are new insights about the consequences of the passage to the TRIPs agreement for a country originally belonging to the South. At the beginning $Q_N = Q_M$ and $Q_S = 0$. $N$ is the only country that participates to the R&D world process, $N$ is the only one that contributes to the creation of the quality $Q_M$; consequently, $N$ is the only country getting the company’s profit ($x = 0$). If $S$ has the opportunity to change its imitative system to an R&D one, transforming $Q_I$ in $Q_S$, with a $x = \frac{Q_S}{Q_M}$ at least equal to the values of $\tilde{x}$ seen above, the TRIPs agreement becomes convenient for $S$. As countries develop and switch from imitative to innovative R&D, they are more likely to be interested in promoting stronger IPR protection.

Technology may be spread out through many different formal and informal channels. A country’s quality of production depends upon its level of development and whether it is able to carry out technical innovation or imitate existing technology. Perhaps, the value of $Q_S$ is connected to the value of $\rho$, so it can be that at the beginning of a development process, despite the rise of $\tilde{x}$, $\rho$ increases at the same time that $Q_S$ is increasing. At the beginning of a development process, countries are not likely to be significant innovators, but they may well have the adaptive capabilities to engage in imitation activities [Falvey et al., 2006]. Then, in a subsequent moment, $Q_S$ could achieve the sufficient value to catch $\tilde{x}$ and make a country willing to accept the TRIPs agreement. This intuition, exploring the linkages between $Q_S$ and $\rho$, can suggest an interesting direction for further research.

If we agree that a country’s share in the official profits is determined by the amount of its contribution to the creation of the quality $Q_M$, and ultimately by the share in the R&D world process, we can proxy this dimension with patents. The theoretical model, therefore, yields the testable predictions that a country’s IPR increase with its patents and that this relation depends negatively on imitative ability.

### 4.4. Empirical validation

In this section, I use the Blundell and Bond [1998] methodology to empirically validate the theoretical results. A panel dataset, described in the following subsection, allows

---

9 Data on R&D expenditures, unfortunately, are scarce (especially before 1996) and not reliable for developing countries.

10 Blundell and Bond [1998], building on the work of Arellano and Bover [1995], developed a system estimator of linear dynamic panel-data models that uses additional moment conditions.
4. Intellectual property rights and imitation in developing countries

to determine whether and how the adopted level of IPR is affected by the amount of patenting and imitative ability, taking into account the level of development, market size, and other relevant economic variables of a country. The implicit assumption in the empirical validation is that the data on patents encompass both a measure of the profits resulting from innovation and the innovative capacity of a country.

4.4.1. Data description

To empirically test the predictions of the model, I use observations on a hundred countries (all countries for which data are available) relative to the period 1980-2005. Detailed descriptions of the data sources may be found in Appendix 4.6.

The index of IPR protection is provided by Ginarte and Park [1997] and Park [2008], and it is the most commonly used indicator of IPR protection.

Imitative ability, IMITAB, refers to a country’s capacity to copy and produce technology and goods of other countries. Following Falvey et al. [2009] and Smith [1999], I use schooling as a proxy for imitative ability. In particular, imitative ability is measured by the average number of years of education received by people aged 25 and older.\(^{11}\) The schooling data are taken from the International Human Development Indicators database and they are an update of the data by Barro and Lee [2012].

A widely used measure of innovation is the number of patent applications filed by residents. PATENTPC, the per capita number of residents’ patent applications, proxies a country’s share both in the official profits and in the R&D world process.

A country’s level of development is measured by its GDP per capita, GDPPC. Population, POP, is a measure of market size. WTO is a dummy for world trade organization membership.

As a proxy for a country’s integration to the world economy, I use its openness to international trade, TRADE, measured by the sum of exports and imports as a share of GDP. FDI stands for foreign direct investments or, better, net inflows of investment from abroad (new investment inflows less disinvestment). Good institutions ensure free and competitive environments. In order to control for differences in institutional quality, I use the Fraser Institute’s economic freedom index,\(^{12}\) ECOFRE.\(^{13}\)

\(^{11}\) Similar qualitative results were obtained using other proxies for imitative ability, like average years and percentage of completion of secondary and tertiary schooling in the population over 15 and 25 years (data source: Barro and Lee [2012]). These results are available upon request.

\(^{12}\) In alternative to the Fraser Institute’s economic freedom index, I run estimations also with the Heritage Foundation’s index of economic freedom, reaching similar qualitative results. Since data on the latter start only in 1995 reducing hence the estimation sample, I proceeded with the former. Even so, the results are available upon request.

\(^{13}\) I also examined tariff rate and R&D expenditures, that have problems regarding the availability of
4. Intellectual property rights and imitation in developing countries

Table 4.1.: Descriptive statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPR</td>
<td>2.765</td>
<td>1.162</td>
<td>.2</td>
<td>4.88</td>
</tr>
<tr>
<td>IMITAB</td>
<td>6.865</td>
<td>2.910</td>
<td>.5</td>
<td>13.1</td>
</tr>
<tr>
<td>PATENTPC</td>
<td>1.2e-04</td>
<td>3.2e-04</td>
<td>2.1e-08</td>
<td>.003</td>
</tr>
<tr>
<td>GDPPC</td>
<td>8406.6</td>
<td>9810.8</td>
<td>101.5</td>
<td>49246.8</td>
</tr>
<tr>
<td>POP</td>
<td>5.6e+07</td>
<td>1.6e+08</td>
<td>224000</td>
<td>1.3e+09</td>
</tr>
<tr>
<td>WTO</td>
<td>.450</td>
<td>.498</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>TRADE</td>
<td>70.383</td>
<td>46.129</td>
<td>13.54</td>
<td>397.13</td>
</tr>
<tr>
<td>ECOFRE</td>
<td>6.210</td>
<td>1.357</td>
<td>1.78</td>
<td>9.14</td>
</tr>
</tbody>
</table>

The sample period is 1980-2005.

Since IPR evolve slowly during time, the index of IPR protection is surveyed quinquennially. Therefore, to be consistent with the patent rights index, data at 5-year intervals are used. The schooling data proxying IMITAB are also surveyed on a quinquennial basis [Barro and Lee, 2012]. With the exception of WTO, observed precisely in the sample years, all the remaining variables are averaged over 5-year periods, with the last year of the period identifying the time of the variable value. In other words, for instance, 2000 values of IPR, IMITAB, and WTO correspond to 1996-2000 averaged data of all the remaining variables. Averaging observations over 5-year periods, 4 of which precede the year of IPR, diminishes potential simultaneity without deleting most of the covariance between IPR and the independent variables. Furthermore, robustness checks performed with no averaged lagged values do not change significantly the results, and the dynamic specifications of the estimated model deal with strict endogeneity.

Table 4.1 reports some descriptive statistics for all the variables. Furthermore, a full breakdown of our dataset by 5-year interval is provided in Table 4.3 in Appendix 4.6.

With a positive correlation of .74 between IMITAB and IPR, the effect of imitative ability on IPR protection seems to be straightforward. Indeed, Figure 4.1 plots the positive relationship between imitative ability and IPR protection. This relationship, however, should be investigated more accurately and it could also be reversed once we take into account profits by innovation and innovative capacity, as I discuss in the next subsection.

---

data, and government consumption expenditures. For brevity, I do not report these results. The results reported in the paper are robust with respect to omission of the above variables.
4. Intellectual property rights and imitation in developing countries

4.4.2. Empirical results

The data presented above allow to explain the level of IPR adoption in terms of imitative ability, patenting, level of development, market size, and other relevant variables. In particular, I investigate whether the response of IPR to patents varies with imitative ability.

In order to estimate the model, before averaging over 5-year periods, I transform some variables in natural logarithms (lnPATENTPC, lnGDPPC, lnPOP). In addition, I interact the patent variable with imitative ability. The coefficient on the interaction term (IMITAB*lnPATENTPC), determining how imitative ability influences the relation between patent applications and IPR adoption, allows me to test the predictions of Observation 13 and 14. The four estimated specifications are presented in Table 4.2.

Specifications (1) and (2) are obtained by pooled OLS. Since innovation investments take time to pay off, changes in IPR should not affect the contemporaneous explanatory variables. However, to control for potential endogeneity, variables are averaged on past years (see data description). I cluster errors at the country level to allow for unrestricted correlation between annual observations within the same country. Both specifications include year dummies and constant. Specification (1) controls for more variables; on the
Table 4.2.: OLS and Blundell-Bond linear dynamic panel estimations

<table>
<thead>
<tr>
<th>Dependent variable: IPR</th>
<th>(1) Pooled OLS</th>
<th>(2) Pooled OLS</th>
<th>(3) Dynamic panel</th>
<th>(4) Dynamic panel</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPR lag 1</td>
<td>-</td>
<td>-</td>
<td>.674***</td>
<td>.811***</td>
</tr>
<tr>
<td>IMITAB*lnPATENTPC</td>
<td>-.007**</td>
<td>-.005*</td>
<td>-.012***</td>
<td>-.011***</td>
</tr>
<tr>
<td></td>
<td>(-2.19)</td>
<td>(-1.77)</td>
<td>(-3.64)</td>
<td>(-1.98)</td>
</tr>
<tr>
<td>lnPATENTPC</td>
<td>.167***</td>
<td>.176***</td>
<td>.036</td>
<td>.065</td>
</tr>
<tr>
<td></td>
<td>(3.66)</td>
<td>(5.03)</td>
<td>(.63)</td>
<td>(.91)</td>
</tr>
<tr>
<td>lnGDPPC</td>
<td>.172*</td>
<td>.190***</td>
<td>.189</td>
<td>.025</td>
</tr>
<tr>
<td></td>
<td>(1.73)</td>
<td>(2.84)</td>
<td>(1.56)</td>
<td>(.15)</td>
</tr>
<tr>
<td>lnPOP</td>
<td>.118**</td>
<td>.106***</td>
<td>.044</td>
<td>.014</td>
</tr>
<tr>
<td></td>
<td>(2.61)</td>
<td>(2.83)</td>
<td>(.83)</td>
<td>(.19)</td>
</tr>
<tr>
<td>WTO</td>
<td>.256</td>
<td>.222</td>
<td>.357**</td>
<td>.299</td>
</tr>
<tr>
<td></td>
<td>(1.18)</td>
<td>(1.20)</td>
<td>(2.11)</td>
<td>(1.41)</td>
</tr>
<tr>
<td>TRADE</td>
<td>.0002</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(.19)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FDI</td>
<td>.0005</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(.57)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECOFRE</td>
<td>.072</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(1.06)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Observations: 423 453 441 441
Groups: - 100 100
Instruments: - 90 44
Year dummies F (1-2) or χ² (3-4): 7.72*** 11.97*** 21.59*** 22.54***
Regression F (1-2) or χ² (3-4): 50*** 58*** 1332*** 724***

The estimation sample is 1980-2005 (5-year intervals), including initial lags; robust t (1-2, errors clustered by country) and z-statistics (3-4) in parenthesis; significance levels: *10%, **5%, ***1%; all model specifications include constant; specifications (3-4) treat lnPATENTPC as endogenous; “ln” refers to natural logarithm of the variable.

Contrary, Specification (2) is more parsimonious.

To assess robustness of results, I estimate also two dynamic specifications. Verbeek [2008] shows that the estimated coefficients of the dynamic and static panel data models are interpretable as short- and long-run effects, respectively. Dynamic specifications (3) and (4) take into account potential residual endogeneity between patents and IPR. I fit these specifications with the two-step GMM system estimator [Blundell and Bond, 1998] and use the Windmeijer’s finite-sample correction for standard errors [Windmeijer, 2005]. Both specifications include year dummies, constant, and one lag of the dependent variable. lnPATENTPC is treated as endogenous and instrumented by lags of the variables. To avoid the problem of instrument proliferation, contrary to Specification (3) that uses all available lags of the dependent and the endogenous variables as instrument, Specification (4), to economize in instruments, sets the maximum number of lags to two. However, given the coherence among the signs of the estimated coefficients, instrument
proliferation does not seem to exert a large influence in the estimated model.

The estimation results on the interaction term in Table 4.2 indicate a lower positive effect of patents on IPR for countries with higher imitative ability. The overall impact of an increase in lnPATENTPC on IPR is dependent on IMITAB; further, the negative impact of an increase in IMITAB is dependent on lnPATENTPC.\footnote{Therefore, the positive relation between IMITAB and IPR presented in Figure 4.1 is reversed once PATENTS are taken into account.}

An increase in GDP per capita implies that the quality level of products to purchase is higher. The significantly positive effect of GDP per capita in Specification (1) and (2) confirms that more developed countries have stronger patent regimes. On the contrary, whenever the dynamic structure of the data is accounted for in Specification (3) and (4), it is not the level of development \textit{per se} that influences the adoption of IPR, but rather the patent profits. Once patents and imitative ability are controlled for, the GDP per capita variable is no longer important.

The coefficients for population, representing market size, are positive consistent with the scale effect on IPR architecture costs. Since the TRIPs agreement requires WTO members to increase their IPR standards, the coefficient estimates for WTO are found positive. Trade, FDI, and economic freedom have positive influences on IPR but are not statistically significant at conventional levels. Therefore, they are dropped after Specification (1) to gain estimation efficiency.

All models have significant (increasing) year dummies, pointing to a time trend in IPR. It seems, therefore, that countries institute patent regimes in response to global pressure on the level of patenting. The Arellano-Bond tests for serial correlation in the first-differenced errors at order one\footnote{Since first-differenced errors are negatively serially correlated by construction, rejecting the null hypothesis of no serial correlation at order one is expected.} and two are, for Specification (3), respectively $z = -3.70^{***}$ and $z = -2.47^{**}$, and, for Specification (4), respectively $z = -3.67^{***}$ and $z = -2.42^{**}$. Finally, the Sargan test of overidentifying restrictions, that tests the null that the overidentifying moment conditions of the models are valid, is, for the homoscedastic version of Specification (3), $\chi^2(78) = 86.73$, and of Specification (4), $\chi^2(32) = 56.11^{***}$. Even if the diagnostic tests suggest that the estimated dynamic specifications are not fully satisfactory, the model is supported by the fact that the estimated coefficients are largely in line with the theoretical economic hypotheses.

Consistent with the theoretical predictions, the empirical analysis finds that patents and imitative ability are strong determinants of patent protection levels across countries and time. In particular, a comparison of all the results shows that the qualitative nature of the effect of imitative ability on the relation between residents’ patents and IPR
adoption is quite robust, giving strong support for Observation 13 and 14. Representing profits by innovation and innovative capacity, patents positively impact on IPR, but this impact decreases with imitative ability.

4.5. Conclusion

While IPR may be considered an important determinant of rich economies’ innovation and growth rates, other instruments, like imitation, may be more relevant for DC (the South). Under imitation, a higher share of newly discovered goods, services, and processes would be available at cheaper prices. An IPR infrastructure, in addition to stimulating research, has costs to operate and imposes burdens on consumers and on producers who use innovations in their work. The cost of drafting patent legislations, training skilled personnel, and building the necessary IPR institutions can be too high for DC. In their pursue of the maximum social welfare, DC have to thoroughly choose whether and how much to comply with international commitments on IPR. This paper develops the intuition that the advisability of tight IPR depends on variables like the amount of research performed, the capacity of imitating, and the quality level of domestic products. Only after having developed a full fledged R&D capacity, a country has incentives to change its legislation in favour of stricter IPR.

At low levels of economic development, an increase in a country’s technology has a lower impact on domestic innovations than on the ability of imitating Northern technologies, which makes IPR less desirable. Once the country’s technology is above a certain threshold, since the effect of domestic innovations on social welfare dominates that of imitative ability, the optimal protection of IPR increases. The main result of this paper is that the South finds profitable to adhere to the IPR international architecture once it has the capacity to generate the IP that guarantees a sufficient participation in the multinational’s profits. This threshold level of participation, $\tilde{x}$, is positively related to the imitative ability of the South. Or equivalently, given the profits share, there is some level of imitative ability after which participation is not optimal. An interaction term in the empirical validation is included to account for this.

The empirical results support the implications of the theoretical model. They confirm both the positive impact of patents on adopted level of IPR protection and the presence of an imitation effect on this impact. The significant negative coefficient on the interaction between residents’ patents and imitative ability is evidence that a higher level of imitative ability requires a higher level of patents to sustain a given level of IPR. I observe also for countries with higher levels of development and a larger market size, a higher adoption
4. Intellectual property rights and imitation in developing countries

Table 4.3.: Sample breakdown by 5-year interval

<table>
<thead>
<tr>
<th>Year</th>
<th>N. of countries</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>70</td>
<td>15.45</td>
</tr>
<tr>
<td>1985</td>
<td>74</td>
<td>16.34</td>
</tr>
<tr>
<td>1990</td>
<td>70</td>
<td>15.45</td>
</tr>
<tr>
<td>1995</td>
<td>83</td>
<td>18.32</td>
</tr>
<tr>
<td>2000</td>
<td>83</td>
<td>18.32</td>
</tr>
<tr>
<td>2005</td>
<td>73</td>
<td>16.11</td>
</tr>
<tr>
<td>Total</td>
<td>453</td>
<td>100</td>
</tr>
</tbody>
</table>

of IPR. However, whenever the dynamic structure of the data is accounted for, the proxy variables GDP per capita and population are no longer important and it is WTO participation that positively affects IPR.

This work illuminates the conflict opposing countries over TRIPs agreement and, more generally, strong IPR enforcement. At the core of this conflict are profits by innovation, innovative capacity, and imitative ability. Given that no variable among these is exogenously given, a promising direction for further research should be to better explore the possibility and the mechanism by which a country could convert imitative ability in R&D capacity.

4.6. Appendix: Data

The empirical analysis is based on a macroeconomic panel dataset relative to a hundred countries’ economies for which comparable data are available. The data cover the period 1980-2005. A sample breakdown by 5-year interval is presented in Table 4.3. The variables used in the analysis are obtained as follows:

- **IPR** is the index of intellectual property rights protection as measured by Ginarte and Park [1997] and Park [2008]. It is an unweighted sum of scores for coverage, membership in international treaties, duration of protection, enforcement mechanism, and restrictions.

- **IMITAB** is the imitative ability proxied by the mean years of schooling of adults, that is the average number of years of education received by people aged 25 and older, converted from education attainment levels using official durations of each level (source: International Human Development Indicators and Barro and Lee [2012]).
4. Intellectual property rights and imitation in developing countries

- PATENTPC is the per capita number of residents’ worldwide patent applications filed through the Patent Cooperation Treaty procedure or with a national patent office for exclusive rights for an invention (source: World Intellectual Property Organization).

- GDPPC is the real gross domestic product divided by midyear population (source: World Bank and OECD). Data are in constant 2000 U.S. dollars.

- POP is the de facto total population, which counts all midyear estimates of residents regardless of legal status or citizenship (source: United Nations Population Division).

- WTO is a dummy for world trade organization membership (source: WTO).

- TRADE is the sum of exports and imports of goods and services measured as a share of gross domestic product (source: World Bank and OECD).

- FDI is foreign direct investment composed by net inflows of investment (new investment inflows less disinvestment) to acquire a lasting management interest (10 percent or more of voting stock) in an enterprise operating in the economy from foreign investors as a share of gross domestic product (source: IMF, World Bank and OECD).

- ECOFRE is an economic freedom index composed by size of government and taxation, private property and the rule of law, soundness of money, trade regulation and tariffs, regulation of business, and labour and capital markets (source: Fraser Institute).
5. Social capital, ideas and relation specific investments

In this chapter, further results concerning social relations and intangible assets are presented.

- Firstly, the policy maker is investigated in her role to build positive social capital and to exploit it to reach her objectives.

- Secondly, an enrichment of the Coase’s theory of the firm explores the links between incomplete contracts, secrecy, IPR and firm dimension.

- Thirdly, it is taken in account the problem of asset specificity. The decentralized solution seems to bring to an excess of the reproduction cost (IPR).
5. Social capital, ideas and relation specific investments

5.1. Social capital, institutions and policy making

“Political Economy or Economics is a study of mankind in the ordinary business of life; it examines that part of individual and social action which is most closely connected with the attainment and with the use of the material requisites of wellbeing. Thus it is on the one side a study of wealth; and on the other, and more important side, a part of the study of man. For man’s character has been moulded by his every-day work, and the material resources which he thereby procures, more than by any other influence unless it be that of his religious ideals.”

Alfred Marshall, Principles of Economics, 1920

Economic processes, consisting of interactions between human beings, exploit the social capital of persons endowed with specific culture and identities. The role of institutions and policy makers is to build positive social capital and exploit it to reach their objectives. However, social capital is elusive and has several dimensions by which to interpret its multifaceted functions in economics and society. We cannot forget, furthermore, that social capital sometimes is undesirable for society, for instance when used for unethical uses. Even so, it is widely accepted that social capital has stable and positive effects.

5.1.1. Introduction

Wellbeing is pursued through individual and social action. The social dimension is therefore a decisive economic force. Social capital, contributing to the cumulative capacity of individuals and social groups to work together for a common goal, is however often overlooked by economic theory. Neoclassical economics models are sometimes too harsh in depicting human behaviours, choices and dynamics. This usually results, at best, in useless economic theories and, at worst, in wrong policy prescriptions and forecasts. A feeling of incongruity may arise from the application to the human behaviour of the homo oeconomicus’ logic. Indeed, economic processes consist of interactions between human beings. These processes exploit the social capital of persons endowed with specific identities.

Social capital evolves during long periods, or better it co-evolves with all the economic institutions and organizations in which it is embedded. Social capital is not always positive, but usually it can solve and reduce both government and market failures. This is why the policy maker should be aware of the social capital’s existence and the constraints it imposes on the changes and challenges a society can face. The role of the policy maker is to build positive social capital and exploit it to reach her objectives. Human capital investment and a correct education strategy are two viable tools that complement and magnify the social capital’s effect on the economy.
5. Social capital, ideas and relation specific investments

To correctly address the relations among social capital, institutions and policy making, it is important to stress that, notwithstanding it can be enhanced, social capital is also highly elusive. On the one hand, even that social capital should be used by institutions to achieve their goals, it cannot be disposed as it were physical capital. On the other hand, even that social capital is out of control and not easy to influence by politicians, it can be raised in the long run by farsighted policies.

In the next sections, we explore social capital dimensions and we consider the intertwinements of social capital, institutions and organizations. A role for social capital in theorizing is then investigated. In the second part of the paper, we introduce social capital among the constraints and aims of the policy maker. The active consideration of social capital as a policy resource gives advice on the proper course of action to reach the policy maker’s objectives, among which we stress education as co-objective to pursue.

5.1.2. Social capital dimensions

Given its multidimensionality, social capital can be defined in many ways. Therefore, since this work does not want to skip any fundamental interpretation about social capital, a clear and precise definition of social capital lacks all along this study. It is interesting, however, to circumscribe at best this concept starting from its characteristic dimensions.

Social capital has different dimensions by which to interpret its multifaceted functions in economics and society. The structural domain of social capital is important to understand organizations, institutions and leadership shaping the structure of social relations. Also the bonding/bridging structure of social capital relates to this dimension. The relational aspect of social capital pertains to trust, norms and identity. This dimension of social capital stresses that relational ties are inimitable. Indeed, ethnic communities are socially complex and have strictly idiosyncratic characteristics. Another dimension of social capital is the cognitive domain that involves mental processes, concepts and ideas. Social groups have seated and shared mental processes embedded in their language, stories and culture. The resulting articulated value system, strictly related to religious principles, has a profound impact on economic development [Weber, 1930, Guiso et al., 2003].

Social capital concerns, for instance, tangible and virtual resources which are collected by parties through social structure and which promote their achievements of goals. The set of resources relevant for social capital are those embedded in relationships. Other concepts often paired to social capital are trust and norms. These concepts are self-reinforcing and cumulative, in the sense that there is a self-sustained process of accumulation of all the resources. Social capital can be built in childhood, without sacrifice and
calculation, and also in an optimizing framework as individual decision making process, with conscious calculation taking in account opportunity costs [Rupasingha et al., 2006].

A fundamental dimension of social capital is trust. Trust, a lubricant of the social system [Arrow, 1974], can be defined as the willingness to permit the decisions of others to influence our welfare [Sobel, 2002]. Notwithstanding its importance in the composition of social capital, trust is different from social capital particularly at the individual level [Glaeser et al., 2002]. Trusting other individuals brings positive externalities to them, therefore, the level of social capital of individuals prone to trust others a lot can be less if they are not targeted of trust at the same level. Trust can be in competence and in intentions [Nooteboom, 2007]. Others can have good intensions towards us but not the competence to help us effectively. On the other hand, others can have the competence and the instruments to affect positively our conditions but lacking the intentions to do it. Trust can change and be raised in different ways, not always idyllic. For instance, in a situation in which individuals have different opinions and diverge on the proper course of events, if trouble is solved by voice and not exit, trust tends to deepen. Trust is also in a significant relation to control. Usually, we think about substitutes between trust and control, meaning that once trust is built control is not necessary anymore. However, some form of control can also complement trust activating a self-reinforcement of these behaviours [Nooteboom, 2007].

Two types of social capital are commonly differentiated [Burt, 1982, Coleman, 1988, Putnam, 2000]. First, bonding social capital refers to ties internal to the social group, ties that are stronger and common in denser networks. Second, bridging social capital refers to external ties, ties that are weaker and are common in larger networks. The former kind of social capital is not always to be preferred to the latter. Too much bonding social capital stifles innovation and adaptation, generates monopolies, collusion and cartels. In addition, weaker ties leading to larger networks are better to share more information than stronger ties of denser networks having a higher degree of overlapping information. The differentiation of these two types of ties is a useful tool also in the field of theory of the firm. In order to develop a well-functioning firm, entrepreneurs should maintain a dual network of both weak ties and strong ties. With a right balance of ties, employees and other actors share common goals and maximize information diffusion.

Many kinds of capital have to evolve to guarantee a proper development of a community. In addition to physical capital, all the forms of capital have usually a positive relation with social capital and they are accessed through it. Community membership gives access to pooled resources, but usually high quality ties are necessary. Also human capital in a community is accessed by its participants in a preferential way. Ethnic or
cultural capital is a resource formed by language, accent, manner and religion rituals. Usually it is positive but there are also cases in which cultural capital has negative influence. For instance, cultural capital can be an obstacle in the process by which social capital is converted to economic wealth when it is not supportive to entrepreneurship [Light and Dana, 2013].

Identity capital or motivational capital is another important form of capital. From identity comes a sense of belonging. When individuals experience an identity, they feel emotionally and cognitively tied to their organization. Identity is important in economics because it can solve the problem of incomplete contracts whenever effort is hard to observe or to reward [Akerlof and Kranton, 2005]. There is also a substitutability to take into account between monitoring, to achieve completeness of information, and identity, to exert effort by intrinsic values. Indeed, since monetary incentives (extrinsic motives) crowd out identity, they are substitutes. Economists’ language is the language of incentives: we offer something valuable, in order to influence and change the choice of the subject. However, as Grant [2002] states, an incentive scheme hides always a relation of power. Finally, also identity is not always positive and can pose problems. Some problems arise, for instance, because identity can be experienced with the working group and not with the organization as a whole.

Institutions matter

In the economy, the same stimulation does not always produce the same effect. In social contexts, institutions affect and modify the outcome of a policy, and more in general a behaviour. Institutions substitute and complement social capital, they co-evolve with social capital [Aoki, 2007] in a positive but also negative way. Whenever agents recognize fair institutions, they stimulate positive conformations of social capital.¹ Trust and good institutions self-reinforce themselves. In addition to understanding such intricate feedback-type interconnections existing in a given period, policy makers should consider their evolutionary dynamics. Indeed, institutions are the result of a historical process; they are, together with social capital, the legacy of history. Social capital is usually paired to the concept of informal institutions. Douglass C. North, in his Nobel lecture, defined institutions as a combination “of formal constraints (e.g. rules, laws, constitutions), informal constraints (e.g. norms of behaviour, conventions, self-imposed codes of conduct) and their enforcement characteristics”.

Impersonal exchanges of contemporary industrial economies are characterized by sev-

¹A typical example of the effect of good institutions on social capital is the raise in entrepreneurship [Percoco, 2012].
eral players, with little information about other parties in transactions and with strategic interactions or games often not repeated or with a precise end. These conditions do not allow cooperation to be sustained. On the contrary, in small-scale societies with personalized exchange, cooperation is pursued more times than competition. The creation of effective economic and political institutions in contemporary economies could involve the alteration of the benefit-cost ratios in favor of cooperative solutions. If this is the case, the predetermination of an institutional setting cannot be omitted in analytical models.

Many holists acquiesce that only individuals, not institutions, can be agents of change, while most individualists endorse that society profoundly affects the individual. Douglass C. North, during the 1980s, explored long-run institutional change. Separating institutions from organizations, he defined the former as the rules of the game and the latter as the players. Groups of individuals with the same objective function compose organizations.\(^2\) And organizations, while pursuing their objectives, are the agents of change. Indeed, while pushed by competition, they try to change the institutional framework to boost their competitive position. The vehemence of competition will affect the rate of institutional change and the perceptions of the actors involved will influence the direction of change. Ultimately, the opportunities resulting from the institutional framework determine the same existence of the organizations that compete to survive.

Individuals create, out of their actions, society. But, to some significant extent, the social situation creates the individual. Individuals develop their habits and routines and accept their social conventions and norms. If we believe that individuals, acting in their self-interest, unintentionally give rise to social rules, we would prefer an invisible-hand approach to institutions rather than a collectivist one. The belief whether spontaneous institutional developments and changes would be consistent with the economic and social system has a deep impact in the position in the debate on interventionism versus non-interventionism. Nevertheless, economic efficiency and social benefit, that should guide interventionists, are difficult to define in the context of an evolving institutional system.

**A role for social capital**

There are several market and government failures for which social capital gains relevance and is a useful instrument. Issues like free-riding, public good, externalities and incomplete information are better solved at the community level. In labour markets, for instance, information is incomplete and therefore recruiting often happens from within

\(^2\)For instance, firms, unions, cooperatives, political parties, regulatory agencies, churches, sport associations and clubs.
homogeneous social categories. Since these same homogeneous social categories spread social norms and monitor members, employers can obtain loyalty and social control within employees’ network.

A community is a group of people who interact directly, frequently and in multi-faceted ways [Bowles and Gintis, 2002]. In communities, there is strong reciprocity, meaning that people are willing to bear the costs of punishments to enforce norms and equity [Fehr and Gächter, 2000]. Other relevant experiments dealing with social capital confirm that trust and reciprocity are basic elements of human behaviour [Berg et al., 1995] and find that the degree of social connection positively correlates with trustworthiness [Glaeser et al., 2000]. However, there are also community failures like the necessary small scale and homogeneity that impede diversity. Public intervention should take into account all these aspects, understanding the limits of communities but also that the community level can solve public and market failures.

Communities are social networks, for which embeddedness of economic actions permits access to external or network resources. The denser a social network the clearer and more firmly held the norms governing behaviours. The representation by means of networks is useful also to understand the kind of links connecting nodes. Whenever two distant part of the network are connected we speak about bridging. In this case, key individuals deal with structural holes and sometimes these individuals are entrepreneurs that generate profit motivated by their position and connections. Long lasting relationships, represented through network participation, are based on reciprocity. The economic consequences of social networks are many, as, for instance, when price changes depending on network position, or in other words social capital [Granovetter, 2005].

Network externalities arise in these contexts. If an individual enters a network, that will bring positive externalities to individuals connected to her but not yet connected to that network. Externalities can also be negative, as for all the outsiders of the network whenever the network absorbs common resources. The link of a network represents a tie, but ties are of different kind/importance. Granovetter [1973] highlights that the strength of ties in social networks depends on factors such as investment of time, emotional involvement, intimacy and reciprocity. Status, social influence that enables to reward and punish others, is also a form of social capital.

---

3Social capital differs depending on the number and the kind of links. For the particular position of entrepreneurs near to structural holes, in which there are by definition no many links, we could think, contrary to the predominant view, of low social capital for them.

4In presence of network externalities, the more widespread and compatible a good the higher its value. In this situation, scale economies and problems of lock-in, path dependence and multiple equilibria emerge.
Horizontal networks are better for knowledge sharing. Instead, vertical networks produce a clear hierarchy among individuals. Hierarchy is often negatively considered, in particular for issues like knowledge diffusion, trust and reciprocity. For instance, trust is lower in countries with dominant hierarchical religions [Porta et al., 1996]. However, we can think also of situations in which a clear hierarchy of competent and influential individuals can have a positive impact on the social capital embedded in an organized and effective social network.

Because of the limits of man’s abilities to comprehend and compute in the face of complexity and uncertainty, in other words bounded rationality, heuristics and suboptimal equilibria arise. Neoclassical instruments are not able to cope with this kind of problems. On the other hand, individuals’ social capital can improve the economic situation and sustain a “good” equilibrium. All the situations plagued with coordination failure may improve with higher levels of social capital. Individuals’ opportunism can be very detrimental with specific investment. Also in this situation social capital fits well in dealing with the problem. The same relationships at the base of social capital are, by definition, specific. Indeed, social capital cannot be traded like every asset that is entirely specific.

The decision maker, like every actor in the economy, faces information and decision-making costs, psychological constraints and existing social norms. Individuals are often rational in an adaptive sense. If conditions change or superior alternatives become available, individuals will adapt behaviours establishing new satisfactory habits and routines. Social norms, that specify many of the goals of action, would be accommodated accordingly. Bicchieri [2006] states that social norms motivate action, but only indirectly. The direct, underlying motives are the beliefs and desires that support the norm. Indeed, the particular form of bounded rationality that characterizes individuals depends on the social context in which they live.

Therefore, if we accept that decision making costs and psychological constraints restrict the role of rational evaluation and promote habits and rules of thumb, rationality is more a matter of gradual adaption over time than a matter of instantaneous optimization. In addition, social norms constrain individual action and shape preferences and goals. Endogeneity of preferences to social capital [Sobel, 2002] has paramount implications with respect to assumptions in our models. We can say that social capital changes what individuals prefer, but also that preferences determine the amount of social capital accumulation.

We cannot forget, finally, that social capital sometimes is undesirable for society. The most common problem is for outsiders that undergo negative effects for the action of
But also insiders lose individual freedom of action because of the rigid enforcement of social norms. Social capital itself may be used for unethical uses by criminals, terrorists and gangs. Social capital can protect mediocrities, reduce objectivity, impose mental conformity and inhibit escape from failing partners.

5.1.3. The policy maker

Welfare functions require explicit value judgements and interpersonal comparisons make many economists uncomfortable. There is no trivial solution to the problem of social welfare. We cannot be satisfied with an overoptimistic picture of government always acting in the social interest [Putnam et al., 1994]. However, also the market does not guarantee efficiency, nor it is free of coercion. The actual markets and governments’ performance has to be continually supervised, analyzed, and regulated to meet changing economic and ethical claims. There is no prior guarantee that the adjustments actually made or not made are in the best interest of society.

Whenever we agree on a shared set of objectives to pursue through public action, we ask the policy maker to act in accordance to them in the most efficient and proper way. It is, therefore, conspicuous the urge of defining better the role of the policy maker. The importance of proper policy intervention is stressed by economic theory. Whenever cooperation and complementarities are common, multiple equilibria exist, some of which with low levels of social capital.

Social contracts are valuable since they affect productivity of individuals and organizations. Social connections yield rules of conduct. When social networks are associated to reciprocity norms, social capital is formed because the reciprocal obligations lessen the incentives for deceitful and opportunistic behaviour. Transaction costs are, indeed, low in instances of mutual trust [Fukuyama, 1995]. As Putnam [2000] points out “a society characterized by generalized reciprocity is more efficient than a distrustful society, for the same reason that money is more efficient than barter”. Civil commitment and social capital, implying reciprocal obligation and responsibility for action, are particularly important regarding productivity for complex tasks, since tacit knowledge, informal rules and cooperation are pervasive wherever these tasks are performed.

Bernheim and Bagwell [1988] take into consideration the effects of public policies in relation to social capital. Family linkages produce complex networks, and individuals pertaining to families belong to many dynastic groupings. Therefore, equity issues of...
public policies are strongly affected by the presence of side transfers all along these complex networks. Many neutrality results, like irrelevance of public redistributions, distortionary taxes and prices, are implied by the complex proliferation of linkages between families. All policy-related results based on the dynastic framework, including the Ricardian equivalence, are affected by the actual structure of family linkages.

Policy makers, as with all the forms of capital, should try to enforce a policy supportive to social capital formation, and also try to exploit social capital in order to achieve their goals. Many works on social capital take it as given and fixed in its amount. However, social capital can change in the long run. Its evolution is slow because social capital is based on norms, on views as to how people should behave that depend upon the particular situation.

Social capital is, therefore, indispensable in the policy maker’s actions. No economic policy that wants to succeed in a specific territory can forget the amount and quality of social capital embedded in its population. On top of that, social capital should be exploited in a smart way to sustain and facilitate the policy maker’s objectives.

**Social capital as a policy resource**

Among the many definitions of social capital, the World Bank’s definition is that of enduring social relationships of trust and reciprocity that enhance a group’s capacity to coordinate actions of its members as they work toward a collective good. Since the aim of economic policy is exactly to coordinate better the actions of individuals toward a collective good, social capital is a natural policy resource. The positive spillovers of social capital are several. A culture of trustfulness brings benefits for everybody. Empirical research shows that trustfulness is positively related to infantile survival, to education and even to efficiency, to the success of many firms, to GDP growth.

The pursuit of economic growth is a well-established part of the political agenda for many countries. However, some concerns about simple economic growth arise when we focus on the forms of social participation. As a matter of fact, developed economies undergo a risk of social impoverishment [Putnam, 2000]. In addition to material needs, individuals have substantial relational needs. The satisfaction of the latter requires the implication of other individuals, that is, some form of social participation. Nonetheless, developed economies have reached a high level of satisfaction with material needs, while relational needs are sometimes overlooked.

Economic growth increases pressure on disposable time, generating substitution of time-intensive activities with time-saving activities. Since most of the activities involving social participation are time-intensive, economic growth generates an impulse toward
5. Social capital, ideas and relation specific investments

private (non-social) activities. As a result of the shrinking of social participation, a social poverty trap (namely a Pareto-dominated equilibrium of the economy) can arise [Sacco et al., 2006].

Several factors may contribute to a further decrease in the level of social participation. Alesina and La Ferrara [2002] identify some factors that explain low trust levels: recent traumatic experiences, belonging to a discriminated minority, failure in terms of education and income, and mixed community (racial and/or income level). Therefore, population heterogeneity, in terms of income or ethnicity, may act as a deterrent to social interaction, as social networks based on homogeneity and shared interests are reduced or broken. Additionally, the increased mobility levels of the working population and rising in- and out-migration trends can bring further momentum to social segmentation. This trend can be observed through the falling rate of homeownership. Indeed, because of high transaction costs in the real estate market, low mobility is paired to homeownership.

Montgomery [2000] studies how political leaders may use social capital originally formed for some purposes to accomplish totally different objectives. Among these objectives, social capital is a potential source of improved voting levels, a multiplicity of active memberships, macro-level improvements in public health associated with community activity and micro-level successes in administering development projects.

Conditional cooperation, experimentally investigated by Fischbacher et al. [2001], supports the fact that people condition their contribution level to the contribution of others. Therefore, there is mounting evidence that cooperation is not merely an altruistic act, but can serve selfish purposes as well [Axelrod, 1984]. People cooperate also because they obtain benefits. For instance, cooperative people encourage reciprocity from others, have a reputation of dutiful members of society with a virtuous behaviour, develop a social status and lay claim to leadership. A frequent and sizeable virtuous cycle of cooperation allows to consider social capital predictable and as a useful basis for policy initiatives.

Even though family network ties seem to affect individual migration [Palloni et al., 2001], when people choose to emigrate they break a large part of their old group ties and affiliations. In such a way, individual behaviours are less driven by inherited communal obligations and more by personal considerations deriving from market (for example wage levels) and non-market value preferences (for example culture charm) [Heffron, 2000]. Masses of migrants change and co-determine in such a way the evolving structure of human and social capital in a given region. However, there are significant costs to moving from one community to another. Therefore migration generates a deep impact on social capital of individuals, affecting, for instance, their propensity to become entrepreneurs.
5. Social capital, ideas and relation specific investments

[Wahba and Zenou, 2012].

Social capital is heterogeneous and somewhat volatile as a policy resource. Public access to energies enclosed in social capital to policymakers depends on the kind of social capital. Indeed, the drivers of a fifty-year period of rapid social and economic development, in addition to the materialist values such as individualism and achievement, were more traditional social norms like obedience and religious faith than postmaterialist values of trust and cooperation one normally associates with social capital [Inglehart, 1997]. In addition, social capital is a neutral resource. If social capital is accessed, its uses may raise the efficiency of a policy. However, the same uses strengthen the commitment to the organizations’ values and benefits that have either positive or negative side effects for the society as a whole.

Education as a co-objective of the policy maker’s action

The role of the policy maker requires to develop the capital of the nation. In the past, the main focus was on investment in physical capital, whereas, today, the emphasis has shifted towards investment in human capital. Increasing human capital is of paramount importance to realize well-being. Knowledge disseminated by education positively affects non-market aspects of the quality of life. Additionally, the externality-type benefits that reach others are even more important to human welfare than pure economic growth.

Raising human capital, through more educated and healthier population, leads to higher productivity that translates into future higher firms’ profits. In knowledge and competence-based economies, people have become extremely important. A more educated population is attained through better educational facilities[^6], meaning better physical and immaterial assets. Investment in education should be prioritized by policy makers for its externality-type benefits on all the population, like reduced crime and enhanced social cohesion.

Social capital is a peculiar kind of capital also for its ability to be infinitely transferred without exhaustion. This characteristic, that it shares with human capital and knowledge, enables individuals to share something that can be given to others without diminishing its availability to the original individuals. Social capital appreciates with use: the more often it is accessed, the higher its level of utility to individuals. Social capital has a strong complementarity with other capitals. Social entrepreneurship, bettering the reputation of a region, produces social capital that can be exploited also by commercial entrepreneurs [Estrin et al., 2013]. Therefore, a widespread non-for-profit

[^6]: There are three parts in installing an educational facility: fabricating a school, operating a school and building human capital.

72
sector and a high level of civil development is supportive to a broad diffusion of social capital and capital in general.

Education can sustain and yield higher profits and wages, improved health, lower fertility rates, political stability and freedom, lower poverty and inequality rates, better environment and lower crime rates. Indeed, education, being critical to the broader process that sustain economic growth, is the building block of economic development generating non-market returns and externalities that are vital to human welfare.

Social capital accumulation occurs via inter- and intra-generational transmission mechanisms. Individuals acquire good culture mainly from their parents and update their values intra-generationally through their life contacts. The relevancy of inter-generational networks is attested by many scholars, as done for example by Magruder [2010] finding that fathers serve as useful network connections to their sons. However, intra-generational networks are less investigated.

Schools and social “spaces” where interactions take place are decisive in creating and shaping social capital. Therefore, spatial segregation should be responsible for the actual dissemination of trust and beliefs. The life in a certain neighbourhood usually brings to certain occupations, socio-economic positions and particularly to shared values. Industrial clusters, being sensitive to incentives not completely encompassed in contracts, witness different levels and typologies of social capital.

Formal and informal institutions contribute to the construction of social capital. In addition to their peculiar characteristics, their spatial localization is a fundamental factor to take in account. Interestingly, Tabellini [2010] argues that informal institutions may help explaining why even under the same formal institutions, regional differences may resist the force of history.

5.1.4. Concluding remarks

Social capital co-evolves with institutions and individuals. Policy makers, not obstructed by a reductionist vision of economics populated only by perfect rational individuals, have to take into account and also to exploit social capital. Indeed, policy making has to be fine-tuned with the proper balance of instruments like market-oriented incentives, public coercion and social and civil enhanced participation. The interactions among these instruments are many and they can puzzle policy makers. Crowding out may occur sometimes with more precise property rights, with attempts to induce higher levels of work efforts, compliance to norms or environmental conservation by fines and sanctions. For such changes, some time may be needed to interiorize the new norms of conduct.

Human relations support economic growth, sustainable development and social progress.
People exist in a relational contest and have relational needs, in addition to biological needs. Therefore, increasing efficiency could be useless if it neglects such needs. But if “no man is an island” [Merton, 2005], Heffron [2000] notes that, given the recent explosion of voluntary group activity, no group is an island either. Individuals use their group membership to access goods and services of other groups of which they are not members.

The amount of social capital within a group is inversely related to the level of the group’s insulation from the other groups and associations in society. Nowadays, group ties and affiliations are often individually chosen driven more by market forces, non-market value preferences and duty to society and the state than by community obligations.

However, social capital and trust binding members of a group together can cause hostility against other groups. Durable relationships help keeping groups together, but they do not necessarily imply trust. Individuals can take advantage of a possible free ride and relax their own efforts believing that the others will carry on. Even so, it is widely accepted that social capital has stable and positive effects.
5. Social capital, ideas and relation specific investments

5.2. The concept of ideas inside the theory of the firm

What is a firm? Why do firms exist? What is the optimal dimension of a firm? These questions were answered by Coase making use of transaction costs. The following work presents an enrichment of this theory exploring the links between incomplete contracts, secrecy, intellectual property rights and firms’ dimension. The principal claim is that the answers to such questions seem to be more complex than the simple transaction cost argument and also that the innovative-idea factor is growing in importance in the new “knowledge era”.

5.2.1. Introduction

A new firm always starts up as a new challenge, as a new idea that tries to enter the market. The analysis that the entrepreneur does is what kind of new product-service consumers will be willing to pay for. In terms which are familiar to economists, a firm tries to raise the utility of consumers selling them something new and different from the others. To do this, the entrepreneur, with her original and new idea, dream or vision, organizes the factors of production in a new mix, sometimes adding a new one.

What I want to discuss is that at the heart of the answers to questions like what firms are and why firms do exist there is the motivation for which a firm is created: an idea. In addition to the well known theory of the firm born with Coase [1937], I would add that in a dynamic setting a firm can not only be seen as a way to minimize transaction costs, but also as a way to create new value for consumers and new profits for who works on the supply side of the market.

As Schumpeter [1934] recognized, entrepreneurship and R&D are at the “core” of the process of economic innovation. Profit is seen as the premium for innovation, it has a temporary nature and it converges to zero in the long run. This is the engine of the capitalist development. The most of this engine is developed inside firms: entrepreneurship and R&D are deeply connected to the identity that differentiates each firm on the market.

By understanding better the “soul” of a firm, we can also better address the problem of the dimension of a firm. My feeling is that, in some way, a firm’s dimension is linked to the exploitation of its original idea. Surely minimization of transaction costs is a driving force shaping the form of the market. However, other forces are playing at the same time: rent seeking behaviour, spontaneous creation of new products like in the open source case, etc.

The purpose of this paper is to concentrate on ideas as motivation for the existence
5. Social capital, ideas and relation specific investments

of firms; and bound to ideas as cause of firms and profits it is straightforward to think about secrecy. Big multinationals like Microsoft and Coca Cola exist also in order to defend their secret formulas as source of profit. By this view, firms’ configurations are chosen as the best way to keep the secret and to continue to make profits.

5.2.2. The model

The model that will be constructed has to be as simple as possible in the aspects that are not important to discuss, letting us to concentrate on the issues of the paper: ideas, secrecy and firms’ dimension.

There are 1,..,i,..,I ideas and every idea is converted in a good sold by a different firm labeled with the same i of the idea. The idea i has demand \( D_i = r^2 \); the markets for ideas are heterogeneous, this simulates different levels of importance of these. Costs of production are = 0 and prices are = 1. There are \( L \) workers in the economy that are split among all firms, thus \( L = \sum L_i \) and \( L_i \) are the workers in the firm i. Production = \( L_i^2 \). There are increasing returns to scale but demands and prices are fixed so that every firm produces and sells all the units of the good-idea that are demanded. In every firm the profit \( \pi_i = L_i^2 \) is divided equally between the workers \( L_i \). Thus, every worker in firm i gets \( L_i \).

The intuition behind this formalization is that the dimension of a firm is given by the demand it is able to create with its product-service sold, in other words, with the implementation of its entrepreneurial idea.

Secrecy or IPR?

Concentrating on the story about the birth and growth of an enterprise could better clarify my argumentation. Initially, there is an entrepreneur that has an innovative idea. However, she needs help from others to implement her idea. The work to do is usually considerable and she needs some help to divide it with other persons. Unsurprisingly, she does not address her job offer to an unknown guy. Usually, as the family style firm so diffused in Italy exemplifies very well, she shares her innovative idea, and the tough work to be done to implement it, with some relative and friend in which she can trust. In this respect, we can see as the neoclassical paradigm is not able to explain this behaviour. In such a setting, the entrepreneur would have chosen someone by chance because of the symmetry assumption of individuals, in other words there is no private information to defend.

In this way, getting rid of the indivisibility of the first sunk cost to be made to imple-
ment the innovative idea in a product, it is installed the first capacity to serve the market
and the first profits emerge. Boldrin and Levine [2008] stress and explain this point very
well, finding that indivisibility may lead monopoly to innovate less than competition.

When and if the market gets larger, the firm has to expand production employing
more workers that in some way will enter in contact with the innovative and productive
idea. Again, secrecy and trust will play a major role in the process of selection of new
workers. All of this story is meant to give rationale to the intuition modeled in this
paper, that is the existence and the dimension of a firm are significantly linked to the
secrecy of the first innovative idea and the way in which it can generate profits.

Coming back to the model, I make now the assumption that: if in some way the
innovative idea is copied (leakage of the industrial secret) another firm identical to the
original enters the same market and competes à la Bertrand, pushing the price and the
profit in sector i to zero. This fear induces firms to defend their profit in two ways:
either they are able to keep secret their ideas or they explicitly extract profit by means
of the intellectual property rights (IPR) instrument. Boldrin and Levine [2004] study
this kind of choices investigating their possible consequences on welfare. My concern
is however different, I am looking to the origin, the motivation and the dimension of a
firm.

**Contractual incompleteness**

Contractual incompleteness is thought responsible for the creation of firms and for their
add this feature to this model, $0 < \theta < 1$ parameterizes contract incompleteness: the
higher $\theta$ the more incomplete contracts. The consequence of more incomplete contracts
for firms is that the probability to lose the profit coming from the exploitation of the
innovative idea is higher.

Thus, the probability to be copied for the firm $i$ is assumed equal to

$$Pr_i \equiv \theta^{1/L_i} \quad (5.1)$$

The probability to be copied has the following characteristics:

- If $\theta = 1$, $Pr_i = 1^{1/L_i}$;
- If $\theta = 0$, $Pr_i = 0$;
- For a given $\theta$, $Pr_i$ is increasing in firm’s dimension $L_i$ (since $\theta < 1$), the bigger the
  firm the higher the probability that someone will bring out the idea from the firm.
5. Social capital, ideas and relation specific investments

i. If $L_i = 1$, that is its minimum, also the probability will be at its minimum (for a given value of $\theta$).

Another plausible feature of the model is that the bigger the firm the more vulnerable it is because of contractual incompleteness. This seems to be in contradiction with the Coase’s theory implication that incomplete contracts render big firms stronger than small firms because contracts “outside” of a firm are more incomplete than contracts “inside” of a firm. On the contrary, this model, by focusing on valuable innovative ideas inside firms, implies that bigger firms, involving more workers and more incomplete contracts among them, are more vulnerable.

The institution of IPR is a way to reestablish the usual property rights setting with the rival property of goods. In this setting, bigger firms usually overcome better a situation of incomplete contracts and score better on the market. Thus, with the increasing dimension of a firm there will be an increasing use of IPR to defend the innovative idea of the firm. When $P r_i$ gets higher the secrecy instrument to defend the innovative idea is losing its efficacy. It then becomes natural to try to defend the profit of the firm in an alternative way. Empirically, this positive relation is usually found. For instance, in the example about Spain presented in Table 5.1, the share of firms that demanded a patent over the firms that made R&D increases with the dimension of these.

<table>
<thead>
<tr>
<th>Table 5.1.: R&amp;D in Spain in 2004</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>$#$ firms that made R&amp;D</td>
</tr>
<tr>
<td>&lt; 250 employees</td>
</tr>
<tr>
<td>$\geq$ 250 employees</td>
</tr>
<tr>
<td>$#$ firms that demanded patents in 2002-04</td>
</tr>
<tr>
<td>&lt; 250 employees</td>
</tr>
<tr>
<td>$\geq$ 250 employees</td>
</tr>
<tr>
<td>Share of firms making R&amp;D that demanded patents</td>
</tr>
<tr>
<td>&lt; 250 employees</td>
</tr>
<tr>
<td>$\geq$ 250 employees</td>
</tr>
</tbody>
</table>

Source: Instituto Nacional de Estadística, www.ine.es/infoine

If the parameter $0 < r < 1$ is assumed to be the measure of IPR defense by institutions, $r * \pi_i$ is a good proxy of the profit share accruing to a firm that chooses to use the IPR instrument selling the license of an idea able to produce a profit of $\pi_i$.

Thus, if

$$(1 - P r_i) \ast \pi_i \geq r \ast \pi_i,$$

that is

$$(1 - \theta^{1/L_i}) \geq r,$$
5. Social capital, ideas and relation specific investments

The firm $i$ will choose secrecy. Otherwise, the firm $i$ will choose to be protected by IPR. As $r$ increases, more firms choose IPR. For a given $r$, the higher the contract incompleteness $\theta$, the higher the number of firms choosing IPR. The firms choosing IPR are the ones with higher $L_i$. Finally, the indifferent firm can be determined:

$$1 - \theta^{1/L^*_i} = r \quad L^*_i = \frac{1}{\log(1 - r)}$$

Figures 5.1, 5.2 and 5.3 show the determination of the indifferent firm. As $r$ grows, $L^*_i$ will decrease ($\theta < 1$) (more firms will choose IPR). As the base of the logarithm $\theta$ grows, $L^*_i$ will decrease (more firms will choose IPR).

5.2.3. Concluding remarks

This work investigates a cause for the existence and the dimension of firms: the idea behind them and the secret information they try to manage without making it public. Nowadays, this aspect of a firm is increasing in importance; we listen always more often speaking about knowledge economies. Thus, the incomplete contracts cause of firms’ existence identified by Coase [1937] is assuming two important dimensions: the residual rights of control that incentive the specific investments by the owner of the firm’s assets; the possibility that someone will bring the innovative idea outside of the firm generating a competition that will push prices and profits towards the perfect competition level. The intuition that firms are always looking to possible rents is not new. The link that I hint is between this behaviour, the origin and the dimension of a firm, and the IPR.
5. Social capital, ideas and relation specific investments

Figure 5.2.: IPR defense and incomplete contracts: log-

Figure 5.3.: IPR defense and incomplete contracts: indifferent firm
5. Social capital, ideas and relation specific investments

instrument.
5.3. Relation specific investments to knowledge

When we consider an asset that is not a physical good but a knowledge, the relevant difference is its property of being not rival. When we talk about the introduction of IPR (intellectual property rights) we are debating over the excludability of this not rival kind of goods. In order to evaluate the overall effect brought by the introduction of IPR on welfare it is necessary to take into account the problem of asset specificity. In the following model, where relation specific investments are taken into account, the decentralized solution brings to an excess of $r$, the reproduction cost or the chosen amount of IPR.

5.3.1. Relation specific investments to private goods

Grossman and Hart [1986] and Hart and Moore [1990] worked out an appropriate theory of the allocation of PR (property rights) in presence of private goods or, more generally, assets. The New Property Rights approach claims well defined PR in order to achieve the most efficient outcome in a society. Once the institutions that guarantee the proper functioning of PR are established we can reach the efficient ownership structure of private goods. Under this structure the property is assigned to whoever values the most these goods: this party is willing to make the most relevant and specific investments in human capital to exploit and interact with the private good under consideration. Thanks to the residual rights of control secured by the property, the owner is covered against the risk of opportunistic behaviour by other parties. By this way, she will make the most relevant investments to increase the value of the private good. As a result, the value of the goods in the society will increase at the fastest possible path.

The problem of opportunism is that not always the private incentives are in line with the public ones. When somebody makes an investment there will be a surplus in an unknown future state of the world. The division of this surplus is hence uncertain and will be done \textit{ex-post}. Because of this unpredictability of the future, a contract to divide this surplus cannot be done \textit{ex-ante}. The impossibility of a contract \textit{ex-ante} will lead to a problem of opportunism because of the asset specificity of these investments. The party that has done the investment will be able to exploit it only in presence of the relationship with the good. Thus, the possibility that the other parties will extract all the surplus from the first one makes it reluctant to invest \textit{ex-ante}. In this setting, opportunism will lead to under-investment, the society will be poorer, everybody will be worse off.

The investments we are referring to can be done in the production either of a good or of human capital. However the important thing is that in both cases there will be
specificity to the principal good object of the PR. Thus, the dangers present in the relations are:

- The party that owns the PR does not allow the others to access the private good. There will be no cooperation.
- The party that owns the PR uses the private good in unfavorable way for the others.
- The party without PR does not collaborate with the one that owns the PR on the private good.

As a consequence of the imperfections of these relations, there is a search for the second best solution well characterized by the New Property Rights approach. However, this kind of solution will lose its efficacy as the number of agents required to make specific investments on the same good increases. Indeed, the three points sketched above hold for each of such investments and so the gap between the first and the second best will be particularly wide with many specific investments. The efficient allocation of ownership rights will therefore display a limited efficacy as an incentive mechanism.

Central to all this reasoning are the peculiar characteristics of the good at the heart of the relation, i.e. a private good. That means rivalry and excludability. Rivalry in particular motivates the dangers presents in the relations involving the good. On the other hand, knowledge is not rival and this has a deep impact in all the line of reasoning for this kind of good.

5.3.2. Relation specific investments to knowledge

When we consider an asset that is not a physical good but a knowledge, the relevant difference is its property of being not rival. When we talk about the introduction of IPR (intellectual property rights), we are debating over the excludability of this not rival kind of goods.

If IPR are set up and effectively defended by institutions, efficiency claims ownership of knowledge by whoever values the most the knowledge protected by IPR. On the contrary, if IPR are not instituted the efficiency problem does not exist and whoever will invest in something specific to this knowledge will not fear neither the excludability from it nor the expropriation of the surplus generated by her investment. All the cooperation that is Pareto-efficient will rise, thanks also to the side payments parties would realize every time it is efficient from a social perspective. When the specific investments required are many or the level of specificity is high the gap between the first and the second
best solution is also high. Hence, in this last occurrence, the non rivalry property of knowledge calls for an absence of IPR.

An example to confirm this intuition of the forces that are acting is the setting up of joint ventures to make R&D. When it occurs, different firms perceive the necessity to construct a space free of IPR impositions, a space where the problems of specificity of assets and opportunism cannot damage the path of innovation. Typically, the knowledge produced by the joint venture is afterwards developed by every firm in a different final rival good.

Another important point to understand is how the absence of IPR solves the problem of specificity of assets. Asset specificity is a problem because the second best option for the owner of the specific asset is markedly worse than the first best. This gives bargaining power to the owner of the good the asset is specific to. As a result, the investment in assets that are specific will be deterred for other parties not being the owner of the good. When there is absence of IPR, the second best option is pushed towards the first best for the owner of the asset that is specific. Knowledge is not defended but at the same time it will flow not at an infinite pace. The level of diffusion of knowledge will determine the formation of alternatives of usage of specific assets, causing at the same time the decay of the spread between the first and the second best option for asset specific owners. Outside the relevant transaction, the ex-post value of specific assets or investments will converge to their ex-ante next best alternatives. The result is that the ex-post bargaining power falls and also falls the reduction of the incentive to invest caused by the feared expropriation.

Another way to see the problem is considering the reproduction of the central good. Asset specificity is important whenever the good is unique. When the good is reproduced, asset specificity is not more stringent because the party that possesses the specific asset can interact with whoever owns one copy of that good. Thus, if the reproduction is feasible at a low cost (as in the case of knowledge diffusion), letting reproduction to happen leads to a solution of the asset specificity problem.

5.3.3. The model

With this model, a quantitative validation of the thesis expressed above will be given. To do that, three different settings will be compared: private good, public good with no cost of reproduction and public good with IPR.
5. Social capital, ideas and relation specific investments

Relation specific investments to a PRIVATE GOOD - RIVALRY

There is a good $g$ and a continuum of humans or individuals $h_i$, with $i \in (0, \infty)$. Every individual $h_i$ can make investments $v_i$. These investments will be productive only if $h_i$ will work with $g$ (specificity of investment). Every individual, if works in conjunction with the good $g$, will exploit her capacity to produce value:

$$c_i = \frac{1}{1 + i} \sqrt{v_i}$$

$$c'_i = \frac{1}{2(1 + i)} \frac{1}{\sqrt{v_i}}$$

Figure 5.4 represents the capacity function and its derivative w.r.t. investments $v_i$ for the individuals 1, 2 and 3.

Rivalry means that only one $h_i$ can work with $g$. The paradigm stated by Grossman-Hart-Moore (GHM) has three properties:

- Investments in human capital;
- Heterogeneity;
- Contractual incompleteness.

All these 3 aspects are present in the model and the result, as we will see, will confirm the general efficiency result found by GHM.

There are two periods. In the first one, $v_i$ will be chosen by individuals; in the second one $\pi_i$ will be realized (we measure $\pi_i$ in $v_i$ units).

Thus,

- $\pi_i = c_i - v_i = \frac{1}{1 + i} \sqrt{v_i} - v_i$, if the individual can work in relation to $g$ in the second period and so benefit of the result;
- $\pi_i = 0 - v_i$, if the individual invest but she has not $g$ in the second period;
- $\pi_i = 0$, if the individual does not invest.
5. Social capital, ideas and relation specific investments

With contractual completeness, the maximum surplus in the society will be reached: $h_0$ will get $g$ and will apply the condition $\pi_0' = 0$ to get the maximum.

$$1 = \frac{1}{2} \frac{1}{v_0} \quad v_0^* = \frac{1}{4} \quad \pi_0 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

**Relation specific investments to a PUBLIC GOOD - NO RIVALRY - NO REPRODUCTION COST**

If $g$ is a knowledge it has the characteristic to be not rival. Now, every individual can interact with a copy of $g$ without hindering the others to do the same. Being so, the social maximum, achieved also in a private decentralized settings, is:

$$\forall h_i \quad \pi_i' = 0$$

That brings to

$$c_i' = 1 \quad \frac{1}{2(i + 1)} \frac{1}{v_i} = 1 \quad v_i^* = \frac{1}{4(i + 1)^2}$$

Thus, in continuous manner we have $v_0^* = \frac{1}{16}, v_1^* = \frac{1}{36}, v_2^* = \frac{1}{36}, ..., v_i^*$. Now we can find the profits for all the individuals

$$\pi_i^* = \frac{1}{(i + 1)^2(i + 1)} - \frac{1}{4(i + 1)^2} = \frac{1}{4(i + 1)^2}$$

Summing up, in this setting, every individual will invest in her $c_i$ a different $v_i$; therefore also the profits will be all different.

**Relation specific investments to a PUBLIC GOOD - NO RIVALRY - IPR**

Now, we assume that it is introduced a reproduction cost of $g$: $r$

The profit maximized by individuals becomes:

$$\pi_{i,r} = c_i - v_i - r = \frac{1}{i + 1} \sqrt{v_i} - v_i - r \quad \pi_{i,r}^* = \frac{1}{4(i + 1)^2} - r$$

Being max $\pi_i = \frac{1}{4}, r < \frac{1}{4}$. With a presence of $r$ reproduction costs not every individual will purchase a copy of the knowledge and will invest in her $c_i$. Only the most productive individuals will do it. The indifferent one will have a profit of 0

$$\pi_{i,r}^* = 0 \quad \frac{1}{4(i + 1)^2} = r \quad \frac{1}{2(i + 1)} = \sqrt{r} \quad i^* = \frac{1}{2\sqrt{r}} - 1$$

86
For example, if we have a $r = \frac{1}{100}$, then $i^* = 4$

Now, we compute the profit loss for the entire industry in presence of $r$

$$\Delta \Pi_r \equiv \int_{i_0}^{\infty} \frac{1}{2} \frac{1}{4(i+1)^2} di + \int_{0}^{i^*} r \text{di} = \frac{1}{4} \left[ -(i+1)^{-1} \right]_{i^*}^{\infty} + ri^* =$$

$$= \frac{1}{4(i^*+1)} + ri^* = \frac{\sqrt{r}}{2} + \frac{\sqrt{r}}{2} - r = \sqrt{r} - r > 0$$

$$\frac{d \Delta \Pi_r}{dr} = \frac{1}{2\sqrt{r}} - 1 > 0$$

that are both $> 0$ being $0 \leq r \leq \frac{1}{4}$

The total welfare in this industry is

$$I \equiv \int_{0}^{i^*} \left( \frac{1}{2} \frac{1}{4(i+1)^2} - r \right) \text{di} = \frac{1}{4} \left[ -(i+1)^{-1} \right]_{0}^{i^*} - ri^* = \frac{1}{4} - \frac{1}{4(i^*+1)} - ri^* =$$

$$= \frac{1}{4} - \frac{\sqrt{r}}{2} - \frac{\sqrt{r}}{2} + r = \sqrt{r} + \frac{1}{4}$$

In Figure 5.5, the continuum of profits and the indifferent individual are presented.

**The $r$ industry.** Now we suppose that

$$\int_{0}^{i^*} r \text{di} = ri^* = \frac{\sqrt{r}}{2} - r \equiv R$$
are taken by who invents the good \( g \). We call \( R \) this revenue. Being \( 0 \leq r \leq \frac{1}{4} \), \( R \) is also \( > 0 \).

\[
\frac{dR}{dr} = \frac{1}{4\sqrt{r}} - 1
\]

\[
\frac{dR}{dr} = 0 \quad \Rightarrow \quad \frac{1}{4\sqrt{r}} = 1 \quad r^* = \frac{1}{16}
\]

\( r^* \) will be chosen by the owner of the right of reproduction of \( g \) as result of her FOC for the profit maximization. By this, we know that in the interval

\[
\frac{1}{16} < r < \frac{1}{4} \quad \Rightarrow \quad \frac{dR}{dr} < 0
\]

At this point, we assume more \( g \) possible: \( g_j \) with \( j \in (0, \infty) \). The cost of developing \( g_j \) is exactly \( j \); the first idea to develop is easy 0, then it is incrementally more costly to develop new ideas.

Now, it is only a matter of counts:

The profit for the \( j \)-idea is

\[
\pi_{g_j} = R - j = \frac{\sqrt{r}}{2} - r - j
\]

Therefore, there will be development of ideas till \( j^* \)

\[
\pi_{g_j^*} = 0 \quad \Rightarrow \quad j^* = \frac{\sqrt{r}}{2} - r
\]

There will be a quantity of \( g_j \) developed equal to \((0, j^*)\)

The sum of all the profits in the \( r \) industry is

\[
D = \int_0^{j^*} \pi_{g_j} dj = \frac{1}{2} \left( \frac{\sqrt{r}}{2} - r \right)^2
\]

The social planner. We have seen that the decentralized choice of \( r \) will be \( r^* = \frac{1}{16} \). Let us see what a social maximizer would choose.

We have that

\[
\frac{dI}{dr} = 1 - \frac{1}{2\sqrt{r}} \leq 0 \quad \Leftrightarrow \quad 0 \leq r \leq \frac{1}{4}
\]

\[
\frac{dj^*}{dr} = \frac{1}{4\sqrt{r}} - 1 \geq 0 \quad \Leftrightarrow \quad 0 \leq r \leq \frac{1}{16}
\]

The sum of all the welfare in all the \((0, j^*)\) industries, where the \( I \) is the constant welfare

88
5. Social capital, ideas and relation specific investments

Figure 5.6.: Social planner solution

and the social value of every industry is

\[ W = \int_0^{\frac{r_*}{2}} I d_j = \int_0^{\frac{r_*}{2}} \left( r - \sqrt{r} + \frac{1}{4} \right) d_j = \left( r - \sqrt{r} + \frac{1}{4} \right) \left( \frac{\sqrt{r}}{2} - r \right) = \]

\[ = \frac{r_*^2}{2} - \frac{r}{2} \left( r - \frac{\sqrt{r}}{4} - r^2 + \frac{r^2}{4} - \frac{3}{2} \sqrt{r} - \frac{3}{4} \sqrt{r} + \frac{\sqrt{r}}{8} \right) \]

\[ \frac{dW}{dr} = -2r + \frac{9}{4} \sqrt{r} - \frac{3}{4} + \frac{1}{16 \sqrt{r}} \]

Now we compute how the sum of all the profits in the \( r \) industry behaves when \( r \) varies:

\[ \frac{dD}{dr} \]

\[ \frac{dD}{dr} = \frac{dV'}{dr} \]

\[ = r - \frac{3}{4} \sqrt{r} + \frac{1}{8} \]

The FOC for the social planner is

\[ V' = \frac{dD}{dr} + \frac{dW}{dr} = 0 \]

\[ \Rightarrow -r^2 + \frac{3}{2} \sqrt{r} - \frac{5}{8} \sqrt{r} + \frac{1}{16} = 0 \]

\[ \left( \frac{1}{2} - \sqrt{r} \right) \left( r - \sqrt{r} + \frac{1}{8} \right) = \left( \frac{1}{2} - \frac{\sqrt{5}}{4} - \sqrt{r} \right) \left( \frac{1}{2} - \sqrt{r} \right) \left( \frac{1}{2} + \frac{\sqrt{5}}{4} - \sqrt{r} \right) = 0 \]

The first solution is \( r = 0.02145 \), the second one is \( r = \frac{1}{4} \), the third one is \( > \frac{1}{4} \) and so we discharge it. Figure 5.6 plots \( V' \), showing that the solution \( r_s = 0.02145 \) is the maximum.

By remembering that \( r^* = 0.0625 \), we can see that \( r_s < r^* \). Therefore, the decentral-
5. Social capital, ideas and relation specific investments

ized solution brings to an excess of $r$, the reproduction cost or IPR.
Part III.

Workplace safety
6. Safety at the workplace: accidents and illnesses

6.1. Introduction

Workplace safety is an increasingly important topic around the world. According to new estimates by the International Labour Organization, every year, across the globe, there are 2.3 million deaths from occupational accidents and work related illnesses, in addition to 337 million nonfatal serious accidents [ILO, 2010]. In 2002 and then in 2007 the European Union defined new strategies [European Commission, 2002, 2007] to foster a continual reduction in work related accidents and illnesses. Even though progress has been made in decreasing both serious accidents/illnesses and deaths at the workplace in Europe over the past few years, work related risks have not been reduced in a uniform way leaving some categories of workers, companies and sectors overexposed to workplace risks [Venema et al., 2009]. Furthermore, the nature of workplace risks is changing due to technological innovation, changes in production organization, and to important transformations that the European labour market has been experiencing (e.g. new type of contracts, the rising number of women in the workplace, new and large waves of immigrants, etc.). For all these reasons, the topic of work safety is a very relevant problem for policy makers and merits being studied in depth.

In economic literature, the problem of work related accidents and illnesses has been analysed from various points of view. Some studies explore the relationship between wages and workplace risk, considering a job as a good characterized by amenities and disamenities (salary, time of work, safety, etc.). For instance, Hamermesh [1999] demonstrates that increasing wage inequality is accompanied by increasing job disamenities, including the risk of accidents. Viscusi [1978]'s model shows that wages paid to workers are not only dependent on job characteristics but also on perceived occupational risk by the workers. Variables other than wages have also been investigated. Worrall and Butler [1983] focus their analysis on the differences in workplace accident risk between blue-collar union workers and blue-collar non-union workers. Others [Fabiano et al., 2004]
focus on the relationship between workplace accident frequency and firm size/type, finding an inverse correlation between the frequency index, referring to all injuries, and firm size. An analysis on the effects of imperfect information on safety at the workplace shows that imperfect information leads to non-optimal levels of precaution both for firms and workers, and not necessarily to an under-provision of it [Lanoie, 1991]. A recent Italian study [Leombruni et al., 2010] investigates the causal effect of displacement on the job-related injury rate, finding that re-employed displaced workers have a higher probability of being injured in their post-displacement period than non-displaced workers. Butler et al. [2006] estimate the patterns of productivity losses that occur after an injured worker returns to work, finding large productivity losses that continue at least three years after accident, even for workers that return to stable employment. Other analyses, addressing the problem of racial inequality in the workplace, find that black workers face a higher rate of work related death than white workers [Leigh, 1983, Stout et al., 1996]. Although, Oh and Shin [2003] find no association between race and non-fatal work accidents, they find that the crucial determinants of work related accidents are human capital (both education and work experience) and occupational conditions (occupational position and work activity). Zimmermann et al. [1999] analyse the interdependence between native and foreign workers, finding no significant differences between the two groups of workers, even though the employment of foreign workers seems to have a strong positive effect on the job safety of natives. Lastly, Marvasti [2010] studies the link between language proficiency, or cultural differences, of foreign-born workers in the United States with American workers, and the prevalence of work related injuries, finding some support for the adverse effect of inadequate English language proficiency on foreign-born workers, and robust results for a cultural gap hypothesis.

A part of economic literature has focused on the relationship between nonfatal accidents in the workplace and types of worker contracts. It is worth noticing that between 1999 and 2007 the percentage of temporary employees with fixed-term contract has risen from 11.8% to 14.5% in the European Union (considering 27 Member States) [European Union, 2010]. This point is very relevant in the analysis of workplace safety because, as shown in a recent study on working conditions in 31 European countries [Parent-Thirion et al., 2007], temporary workers have fewer opportunities to receive training at the workplace. Furthermore, limited experience might also lead the worker to underestimate the risk associated with a particular work situation [Quinlan, 1999]. Therefore, the kind of contract and the amount of work experience might influence the likelihood of accidents. In addition, working conditions could be considered worse for those workers hired with temporary contracts due to their low bargaining power and it is possible that employers
tend to invest less in health and safety for temporary workers than for permanent workers, and tend not to involve temporary workers in safety protection plans [Hebdon and Stern, 1998, Amuedo-Dorantes, 2001].

Reviewing the empirical literature reveals analyses that focus on particular sectors or specific aspects of nonfatal accidents and illnesses. It is worth noting that some empirical analyses based on micro-data find different or even contradictory results about work related accidents and fixed-term (temporary) contracts. Results presented by Guadalupe [2003], using data from the 1989-1998 period in Spain, are the most relevant in supporting the case of a strong positive effect of fixed-term contracts on work related accidents. Blank et al. [1995] find that in the Swedish mining sector accidents occurring to temporary workers (contractor workers) are more frequent and more severe than those occurring to permanent workers. Dupré [2001], on the basis of 1999 data regarding European countries, discovers that the risk of accidents for temporary workers who had been employed for less than two years, is particularly high in the construction, health, and social sectors. Medical studies also investigate the issue of workplace safety and fixed-term contracts. For example, a survey [Virtanen et al., 2005] on reports about fixed-term contracts and health shows that temporary workers may have a higher risk of psychological morbidity and work related injuries as compared to permanent workers. However, Benavides et al. [2006] find that even though temporary workers seem to have a higher risk for work related injuries than permanent workers, the probability of accidents is quite similar between the groups after controlling for the length of employment. In fact, other literature seems to confirm that a worker with either a fixed-term or open-ended contract has a similar probability of suffering a work related accident when all the relevant variables are taken into account. For instance, Amuedo-Dorantes [2001], using 1997 data from Spain, finds that the higher rates of accidents/illnesses for temporary workers are due to worse working conditions than other workers. In addition, the results of the study by Hernanz and Toharia [2006], based on 1999 Spanish and Italian data, show that the differential of accident rates between temporary and permanent workers vanishes once personal and job characteristics are controlled for. Finally, García-Serrano et al. [2010], using 2004-2007 data from Spain find that workers hired through temporary help agency have lower probability to suffer serious accidents, and the duration of their absences are shorter than absences of workers hired both through open-ended contracts and through direct temporary contracts.

Differing from studies on specific sectors, this paper aims to enrich the comprehension of workplace safety as a whole by offering a broad view of the phenomenon. In particular, it studies the key determinants of workplace safety in an attempt to document the
relationships between work related accidents, illnesses, and types of contracts and then personal, firm, and job characteristics, including working conditions. This work tests specifically whether temporary and permanent workers are subject to different types of workplace safety. The only other similar study uses 1997 data from Spain [Amuedo-Dorantes, 2001], while this analysis uses data collected ten years later, when the labour market had changed, and focuses on Italy. Moreover, a wider array of econometric tools controls for the robustness of the results.

The paper is organized as follows: Section 6.2 describes the data from the Italian Labour Force Survey in 2007 used in the regressions. Section 6.3 briefly presents the econometric methodologies employed. Section 6.4 reports the results of the estimated models: two probit regressions testing the relationship between the likelihood of accidents and illnesses and types of contracts, by controlling for job, firm, and personal characteristics; an evaluation if the repetition of accidents/illnesses depends upon a different set of variables; a bivariate probit regression model to jointly analyse the probability of accidents and of illnesses at the workplace; and finally, following the indication of diagnostic tests, the last two probit models for work accidents/illnesses restricted to a subsample (workers with a maximum tenure of three years). Section 6.5 concludes.

6.2. Data

The micro-data used in the analysis come from the Labour Force Survey carried out by Istat, the Italian National Institute of Statistics. The dataset refers to the second quarter of 2007, when an “ad hoc” module devoted to safety and health at work was added to the standard information contained in the Survey. The survey is entirely comparable (concerning the organization, the information collected and the definitions used) with those carried out in other EU countries. In fact, since 1999 Eurostat requires all member states to carry out quarterly interviews using the same scheme.

Istat collects the information each quarter by interviewing a sample of nearly 77,000 households (approximately 300,000 in one year), representing 175,000 individuals who are Italian residents, even if they are temporarily abroad. For further information, see the Istat website (http://www.istat.it/en/).

The “ad hoc” module was included in the Labour Force Survey by all Members of European Union in the second quarter of 2007 to provide an assessment of the effects of the European strategy for
6. Safety at the workplace: accidents and illnesses

employees with open-ended or fixed-term contracts, excluding all other types of worker (for instance, other fixed-term temporary workers) and unemployed individuals.

The survey collects various kinds of information on the job, firm, and personal characteristics of workers. In addition, the survey provides data on working conditions (that are added to job characteristics) and on recent work related accidents and illnesses. The reference period for work related accidents and illnesses is 12 months which could lead to an undervaluation of job insecurity since according to both Landen and Hendricks [1995] and Oh and Shin [2003] the number of accidents reported decreases as the time gap between the interview date and the actual date of the accident increases. Moreover, since fatal work accidents and fatal work diseases are not considered in the data, the true extent of the problem of insecurity at the workplace is underestimated. On the other hand, the dataset does not suffer from a systematic underreporting bias since Istat collects the information directly at the household residence which ensures statistical confidentiality.

The “ad hoc” module contains detailed information related to safety at work for each worker. Specifically, it documents if the worker had experienced:

- non-fatal accidents at the workplace in the preceding twelve months;
- non-fatal illnesses or health problems caused or exacerbated by working in the preceding twelve months;
- non-fatal accidents on the way to work in the preceding twelve months.

A part of the “ad hoc” module also includes worker exposure to risk factors for health. These factors proxy working conditions considering a set of dichotomous dummy variables grouped in the tables with job characteristics. Both physical and psychological risk factors are considered. Included in the physical risk factors are: exposure to dust, fumes, smoke, chemicals; exposure to excessive noise or vibration; bad posture induced by work requirements, movement of heavy loads; and exposure to a general risk of injury. Among the risk factors that may affect the psychological balance of workers there are: excessive workload; phenomena of bullying or discrimination; and exposure to threats or physical violence.

Following the Istat definition [Istat, 2008], work related accidents can be defined as episodes of injury that occur at work which lead to a disability, total or partial, permanent or temporary, while work related illnesses can be defined as disabilities or other physical or mental problem caused or made worse by working. In this analysis, two health and safety at work.
6. Safety at the workplace: accidents and illnesses

Dichotomous variables (work accidents/work illnesses) indicate if the worker had experienced at least one episode during the previous 12 months. The determinants of the repetition of work related accidents and work related illnesses are also examined using other two dichotomous variables.\(^5\)

In addition to the type of contract, the determinants of accidents/illnesses probabilities are grouped into three categories: job, firm, and personal characteristics. Since it seems that workers with fixed-term contracts may have a higher incidence of work related accidents or illnesses, it is important to look at working time (dummy variables full-time/part-time, overtime hours, and shift work), at professional position\(^6\), and at working conditions (described above), in order to control for job characteristics. Firm characteristics are considered by means of two variables: establishment size (dummy variable ten or fewer/more than ten workers) and main activity sector of the firm.\(^7\) The last relevant set of variables is personal characteristics which include gender, age, number of household members, and marital status. These variables help to capture different effects related to how careful different groups of workers are at the workplace. Additionally, the inclusion of variables such as birthplace (Italy or abroad) and geographic area of residence captures specific socio-cultural values otherwise unobserved [Hernanz and Toharia, 2006]. A dummy variable for accidents occurring on the way to work is also included, as done in the study by Guadalupe [2003]. Introducing this variable helps capture the personal proneness to accidents which allows for a possible ability bias in the contract coefficient. Human capital is accounted for in the variables years of education, months of current job tenure, and two dummy variables for (on the job) recent training activity and for being new to the workforce (first job). Moreover, the variable occupation identifies the specific kind of job performed by the employee in his/her workplace.\(^8\) Finally, the indication of whether the worker is looking for another job is used to capture job dissatisfaction.

The total number of employees in the sample is 45,131, of which 44,620 answered questions on work related accidents and 44,489 on work related illnesses. Table 6.6 in Appendix 6.6 provides descriptive statistics for the entire sample, for the workers who

\(^{5}\)Work related accidents/illnesses repetition is a dichotomous variable indicating whether the worker had experienced more than one accident/illness (instead of only a single accident/illness) during the previous 12 months.

\(^{6}\)Professional positions are grouped into the following categories: manager or director, white collar worker, and blue collar worker or apprentice.

\(^{7}\)Sectors of activity are grouped into the following categories: agriculture, industry excluding construction, construction, retail, and other activities.

\(^{8}\)Occupations are grouped into the following categories: executive or intellectual occupation, technical position, office clerk and qualified occupation, and craftsman and operator of industrial machinery.
6. Safety at the workplace: accidents and illnesses

Table 6.1.: Incidence (%) of work accidents and illnesses by contract type

<table>
<thead>
<tr>
<th>Worker Contract Type</th>
<th>Accidents</th>
<th>Illnesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open-ended</td>
<td>2.7</td>
<td>7.6</td>
</tr>
<tr>
<td>Fixed-term</td>
<td>2.5</td>
<td>5.6***</td>
</tr>
<tr>
<td>Total</td>
<td>2.7</td>
<td>7.3</td>
</tr>
</tbody>
</table>

Significance levels (test of proportion) – difference with respect to worker with open-ended contract: * 0.10 ** 0.05 *** 0.01
Source: 2007 second quarter Istat Labour Force Survey

experienced accidents or illnesses, and for the workers who experienced no accidents and no illnesses, reporting the mean and the standard deviation of the variables used in the analysis, grouped by job, firm, and personal characteristics.

Table 6.2 shows the occurrence of work accidents and illnesses by contract type. The raw percentage of workers that had experienced at least one work related accident is similar for both types of contract (2.7% for open-ended contracts and 2.5% for fixed-term contracts), while the rate of work illnesses is greater for open-ended contracts (7.6%) than for fixed-term contracts (5.6%). The aggregate data for accidents (2.7%) is in line with the official data by INAIL, the Workers Compensation Authority, while INAIL does not compute the incidence for illnesses.9

These preliminary results are unavoidably raw because they do not take into account all of the variables that could influence the probability of having accidents/illnesses. Conversely, by exploiting the rich dataset at disposal and by different econometric techniques, ceteris paribus results are presented in the next sections. The main aim is understanding the determinants of work related accidents/illnesses by assessing which variable among worker’s job, firm, and personal characteristics is significant and should be taken into consideration.

6.3. Econometric specification

This section presents a brief discussion of the methodologies employed to derive the results presented in subsequent sections. Cameron [2005], Greene [2007] provide in

9In Italy, INAIL maintains the official data on work accidents and illnesses. Unfortunately, to date, its database does not include any information on the type of contract. To monitor the trend of work related accidents, INAIL calculates two different indicators: (1) the index of incidence and (2) the index of frequency. Even though these indices are not built perfectly homogenous with the data used in this study, they are useful in checking a general reliability. In particular, looking at 2007 the index of incidence is 3.5% and the index of frequency is 2.8% (INAIL 2010).
depth studies of these concepts.

In order to investigate the probability of having work related accidents (the same holds for illnesses), a latent variable representing the danger of the work experience is considered. In addition, this variable is assumed to be determined by

$$y^*_i = \alpha d_i + \beta' X_i + \epsilon_i$$  \hspace{1cm} (6.1)

where the scalar $d_i$ and the vector $X_i$ are the covariates ($d_i$ is the main variable of interest, the dummy fixed-term contract; $X_i$ contains the control variables). For the characteristics of the individual $i$, the scalar $\alpha$ and the vector $\beta$ are the parameters, and $\epsilon_i \sim N(0,1)$ is the error term independent of $d_i$ and $X_i$. Notwithstanding $y^*_i$ is not observable, the binary variable $y_i$, that records if the individual declares work accidents, is observable and it is defined as:

$$y_i = \begin{cases} 1 & \text{if } y^*_i > 0 \\ 0 & \text{if } y^*_i \leq 0 \end{cases}$$  \hspace{1cm} (6.2)

Since $\epsilon_i$ is normally distributed,

$$P(y_i = 1|d_i, X_i) = P(\alpha d_i + \beta' X_i + \epsilon_i > 0|d_i, X_i) = \Phi(\alpha d_i + \beta' X_i)$$  \hspace{1cm} (6.3)

where $\Phi$ denotes the standard normal distribution function. Furthermore,

$$P(y_i = 0|d_i, X_i) = 1 - \Phi(\alpha d_i + \beta' X_i)$$  \hspace{1cm} (6.4)

By (6.3) and (6.4) it is easy to obtain the log-likelihood function $L(\alpha, \beta)$. Maximizing it, the maximum likelihood estimation of $(\alpha, \beta)$ is obtained, that is the estimation of the probit model.\footnote{As detailed in Cameron [2005], the advantage of the probit model compared to the logit model is that it can be employed a latent normal random variable (that represents the danger of the work experience) in order to obtain a more natural interpretation. This choice does not influence the empirical results in a significant way.} The coefficients presented in Tables 6.2, 6.3, and 6.5 are obtained in this way.

In order to obtain a complete representation of job safety, statistical inference treating the two phenomena in a joint manner is performed. A natural extension of the probit model for two equations with correlated disturbances is the bivariate probit model. The unobserved latent variables, that represent how dangerous and how unhealthy the work
6. Safety at the workplace: accidents and illnesses

experience is, are defined as:

\[ y_{1i}^* = \alpha_1 d_{1i} + \beta_1' X_{1i} + \epsilon_{1i} \]
\[ y_{2i}^* = \alpha_2 d_{2i} + \beta_2' X_{2i} + \epsilon_{2i} \]  (6.5)

where \( \epsilon_{1i}, \epsilon_{2i} \sim N(0, 1) \) have a correlation equal to \( \rho \). The observed variables, that record if the individual declares accidents and if he/she declares illnesses related to the job, are:

\[ y_{1i} = \begin{cases} 1 & \text{if } y_{1i}^* > 0 \\ 0 & \text{if } y_{1i}^* \leq 0 \end{cases} \]
\[ y_{2i} = \begin{cases} 1 & \text{if } y_{2i}^* > 0 \\ 0 & \text{if } y_{2i}^* \leq 0 \end{cases} \]  (6.6)

All the possible outcomes are:

\[ p_{jk,i} = P(y_{1i} = j, y_{2i} = k | d_{1i}, X_{1i}, d_{2i}, X_{2i}) = \Phi_2[(2j - 1)(\alpha_1 d_{1i} + \beta_1' X_{1i}), (2k - 1)(\alpha_2 d_{2i} + \beta_2' X_{2i}), \rho] \]  (6.7)

where \( \Phi_2(z_1, z_2, \rho) \) denotes the standard bivariate normal cumulative distribution function for \( (z_1, z_2) \) with zero mean, unit variance, and correlation \( \rho \). The log-likelihood function \( L(\alpha_1, \beta_1, \alpha_2, \beta_2) \) coming from (6.7) is maximized. The result is the maximum likelihood estimation of \( (\alpha_1, \beta_1, \alpha_2, \beta_2) \), also known as the estimation of the bivariate probit model. The coefficients presented in Table 6.4 are found in this way.

6.4. Results

The main interest is to explain job safety as a broad phenomenon. However, first let us focus on a particular variable: the type of contract. Therefore, in order to capture the contract effect, in every model there is a dummy variable that indicates if the employee is hired with a fixed-term contract or not. In addition, job, firm, and personal characteristics are controlled for.

6.4.1. The probability of accidents and illnesses at the workplace

Table 6.2 shows the results concerning the probability of suffering accidents or illnesses at the workplace. In particular, Table 6.2 reports the estimated coefficients and the robust standard errors.\(^{11}\) Since joint Wald tests on the regressions are significant, the

\(^{11}\)Tables 6.2-6.5 in the paper present only one specification of diverse models, since the results are robust to these specifications (the significant variables in the tables continue to be significant and with the same sign for all the model specifications). The results of these various model specifications are available upon request.
models succeed in explaining the probability of the dependent variables. The significant likelihood ratio test on heteroskedasticity\(^\text{12}\) is rationale for using robust\(^\text{13}\) standard errors.

### Table 6.2.: Probability of accidents and illnesses at the workplace – Probit

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Accidents</th>
<th>Illnesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>Robust S E</td>
<td>Coefficient</td>
</tr>
<tr>
<td>Fixed-term contract</td>
<td>-0.026(^*) (0.047)</td>
<td>0.047(^*) (0.038)</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.794(^***) (0.325)</td>
<td>-2.469(^***) (0.249)</td>
</tr>
<tr>
<td><strong>Job characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-time contract</td>
<td>0.124(^**) (0.051)</td>
<td>0.017 (0.035)</td>
</tr>
<tr>
<td>Overtime hours</td>
<td>0.088(^*) (0.046)</td>
<td>0.129(^***) (0.034)</td>
</tr>
<tr>
<td>Shift work</td>
<td>0.160(^***) (0.033)</td>
<td>0.074(^***) (0.026)</td>
</tr>
<tr>
<td>Manager or director as reference category</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White collar worker</td>
<td>0.049 (0.075)</td>
<td>-0.047 (0.044)</td>
</tr>
<tr>
<td>Blue collar worker or apprentice</td>
<td>0.297(^***) (0.086)</td>
<td>-0.096(^*) (0.054)</td>
</tr>
<tr>
<td>Exposure to dangers such as dust, etc.</td>
<td>-0.012 (0.039)</td>
<td>0.161(^***) (0.030)</td>
</tr>
<tr>
<td>Noisy workplace</td>
<td>0.082(^**) (0.040)</td>
<td>0.247(^***) (0.031)</td>
</tr>
<tr>
<td>Bad posture induced by work</td>
<td>0.180(^***) (0.035)</td>
<td>0.525(^*) (0.027)</td>
</tr>
<tr>
<td>Feeling exposed to risk of injury</td>
<td>0.491(^***) (0.036)</td>
<td>0.115(^**) (0.029)</td>
</tr>
<tr>
<td>Excessive workload</td>
<td>0.031 (0.038)</td>
<td>0.462(^**) (0.026)</td>
</tr>
<tr>
<td>Feeling exposed to bullying or discrimination</td>
<td>0.239(^***) (0.054)</td>
<td>0.564(^***) (0.036)</td>
</tr>
<tr>
<td>Feeling exposed to threats or physical violence</td>
<td>0.167(^*) (0.090)</td>
<td>0.225(^***) (0.064)</td>
</tr>
<tr>
<td><strong>Firm characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Establishment size</td>
<td>0.036 (0.035)</td>
<td>0.089(^***) (0.027)</td>
</tr>
<tr>
<td>Agriculture as reference category</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industry excluding construction</td>
<td>0.103 (0.094)</td>
<td>-0.069 (0.068)</td>
</tr>
<tr>
<td>Construction</td>
<td>0.080 (0.100)</td>
<td>-0.028 (0.075)</td>
</tr>
<tr>
<td>Retail</td>
<td>0.078 (0.102)</td>
<td>-0.102 (0.074)</td>
</tr>
<tr>
<td>Other activities</td>
<td>0.200(^**) (0.094)</td>
<td>0.011 (0.067)</td>
</tr>
<tr>
<td><strong>Personal characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current job tenure</td>
<td>5.8e-04 (4.4e-04)</td>
<td>9.2e-04(^***) (3.2e-04)</td>
</tr>
<tr>
<td>Square of current job tenure</td>
<td>-1.7e-06 (1.1e-06)</td>
<td>-8.4e-07 (7.3e-07)</td>
</tr>
<tr>
<td>First job</td>
<td>-0.057 (0.038)</td>
<td>-0.100(^***) (0.027)</td>
</tr>
<tr>
<td>Looking for another job</td>
<td>0.180(^***) (0.055)</td>
<td>0.347(^***) (0.042)</td>
</tr>
<tr>
<td>Number of household members</td>
<td>0.011 (0.014)</td>
<td>-0.018(^*) (0.011)</td>
</tr>
<tr>
<td>Male</td>
<td>0.132(^***) (0.035)</td>
<td>-0.141(^***) (0.025)</td>
</tr>
<tr>
<td>Born in Italy</td>
<td>-0.035 (0.051)</td>
<td>-0.016 (0.043)</td>
</tr>
<tr>
<td>Years of education</td>
<td>0.015 (0.035)</td>
<td>-0.051(^*) (0.024)</td>
</tr>
<tr>
<td>Square of years of education</td>
<td>-0.001 (0.001)</td>
<td>0.002(^**) (0.001)</td>
</tr>
</tbody>
</table>

\(^\text{12}\)All the heteroskedasticity tests are done by maximum-likelihood heteroskedastic probit models. They are generalizations of the probit models in which the normal CDF no longer has a variance fixed at 1 but can vary as a multiplicative function of the independent variables [Harvey, 1976]. The likelihood-ratio test of heteroskedasticity tests the full model with heteroskedasticity against the full model without.

\(^\text{13}\)The robust or sandwich estimator of variance is used. This estimator is robust to some types of misspecification so long as the observations are independent [Greene, 2007].
First, let us focus on work related accidents. An important finding indicates that the type of contract does not seem to affect the likelihood of having accidents. Indeed, even though the estimated coefficient for the fixed-term contract dummy variable has a negative sign, it is not statistically significant. Therefore, the variables that influence safety at work seem to be other than contract term. It is noticeable that job, personal characteristics, and to a lesser extent, firm characteristics can affect the probability of experiencing work related accidents. In the category of job characteristics, being a full-

---

**Table 6.2 – Continued from previous page**

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Accidents Coefficient</th>
<th>Accidents Robust S E</th>
<th>Illnesses Coefficient</th>
<th>Illnesses Robust S E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Educational activities in last four weeks</td>
<td>-0.117 (0.072)</td>
<td>0.193*** (0.041)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accidents occurred on the way to work</td>
<td>0.315*** (0.106)</td>
<td>0.456*** (0.078)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.003 (0.010)</td>
<td>0.020** (0.008)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Square of age</td>
<td>3.9e-05 (1.2e-04)</td>
<td>-1.0e-04 (9.3e-05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North-west as reference category</td>
<td>g</td>
<td>h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North-east</td>
<td>0.068* (0.038)</td>
<td>0.086*** (0.030)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Central</td>
<td>-0.016 (0.044)</td>
<td>-0.010 (0.034)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>-0.101** (0.043)</td>
<td>0.089*** (0.031)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Islands</td>
<td>-0.095* (0.056)</td>
<td>-0.021 (0.041)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Executive or intellectual occupation as ref. cat.</td>
<td>i</td>
<td>j</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technical position</td>
<td>0.002 (0.081)</td>
<td>0.090* (0.046)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Office clerk and qualified occupation</td>
<td>0.005 (0.088)</td>
<td>0.086* (0.051)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Craftsman and operator of industrial machinery</td>
<td>0.049 (0.091)</td>
<td>0.049 (0.058)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never married as reference category</td>
<td>k</td>
<td>l</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>-0.037 (0.040)</td>
<td>0.087*** (0.032)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Separated or divorced</td>
<td>0.073 (0.066)</td>
<td>0.065 (0.049)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Widow/widower</td>
<td>-0.017 (0.129)</td>
<td>0.045 (0.084)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of obs 39590 39471
Regression Wald test $\chi^2(41) = 1164.19^{***}$ $\chi^2(41) = 3306.45^{***}$
Pseudo $R^2$ 0.112 0.172
Heteroskedasticity LR test $\chi^2(41) = 1110.55^{***}$ $\chi^2(41) = 3525.80^{***}$

Average effects: $-0.112%$ $0.478%$

Professional position dummies: $^c\chi^2(2) = 27.87^{***}$ $^d\chi^2(2) = 3.34$
Sector dummies: $^e\chi^2(4) = 12.73^{**}$ $^f\chi^2(4) = 12.78^{**}$
Geographic area dummies: $^g\chi^2(4) = 17.26^{***}$ $^h\chi^2(4) = 20.50^{***}$
Occupation dummies: $^i\chi^2(3) = 1.13$ $^j\chi^2(3) = 4.99$
Marital status dummies: $^k\chi^2(3) = 3.37$ $^l\chi^2(3) = 7.73^{*}$
Significance levels: $^* 0.10$ $^** 0.05$ $^*** 0.01$
Source: 2007 Istat Labour Force Survey

---

14 Since the estimated coefficients are not immediately interpretable as partial effects on the dependent variable, looking at the average effects (obtained by computing the variation of the estimated probability for the mean individual) allows to say that the average individual with a fixed-term contract instead of an open-ended one would have a 0.112% lower probability to incur workplace accidents.
time worker as well as doing overtime hours and shift work (as measures of work intensity) increases the probability of having accidents. Clearly, expanding working time, *ceteris paribus*, increases the chance of accidents. When looking at working conditions, the variables considered in the analysis show positive and statistically significant coefficients. For example, the probability of having accidents is, on average, 2.9% higher for those workers that feel exposed to risk of injury. The information coming from the professional position dummies confirms the intuition that blue collar workers are exposed to higher risks (+1.3%) than managers or directors. Firm characteristics seem to slightly affect the probability of injury. On the one side, the variable for establishment size is not significant even if the sign is positive (larger size brings more accidents). On the other side, the categorical variable sector of activity presents a positive joint significance but it is not possible to isolate the effect of all single sectors.

When looking at personal characteristics, gender affects the probability of having accidents. Males, on average, have a 0.6% higher risk of accidents than females. At the end of Table 6.2 it is possible to see a significant Wald test on the coefficients of a restricted cubic spline approximation [Dupont and Dupont, 2009] for current job tenure. However, in order to obtain readable results, the presented probit model includes a quadratic form which is plotted in Figure 6.1. The contribution to the probability of work accidents increases as the current job tenure rises until 167 months (almost 14 years), the point in which it reaches the maximum. After that the contribution decreases and reaches its minimum for the maximum level of tenure.\footnote{Figure 6.1 shows the prediction and the confidence intervals of the probability of accidents for the average individual in the sample for each level of current job tenure.}

Workers who are looking for another job seem to have greater likelihood (+0.9%) of incurring workplace accidents than their colleagues. Indeed, poor motivation, lack of satisfaction and desire for a different job could easily translate into less effort and less care at work. Personal characteristics also include the geographical area of residence. The probability of work related injury decreases going from north to south. This trend may be associated with the different levels of unemployment throughout the country (higher in southern Italy). In fact, as underlined in other works (Guadalupe [2003], p. 352), a high rate of unemployment can have a negative impact on accidents because it implies lower activity. As already noted, a dummy for the accidents experienced by workers on the way to work proxies their degree of proneness to accidents: the worker that experiences accidents on the way to work has a greater likelihood (+1.9%) of also having accidents at the workplace. Finally, the worker’s place of birth (Italy or abroad), the years of education, and age are not statistically significant.
When focusing on work related illnesses, a similar result about the type of contract is obtained. Change of sign notwithstanding, the estimated coefficient for the dummy variable fixed-term contract is still not significant.\textsuperscript{16} Similar to the model on work related accidents, job, firm, and personal characteristics help explain the likelihood of work illnesses. Among job characteristics, working conditions significantly affect the probability of illnesses; for instance, workers with bad posture induced by work requirements or who are subject to excessive workload endure a 6.8% and 6.1% higher risk of illnesses respectively. Also psychological working conditions prove to have a large importance in determining work illnesses: feeling exposed to bullying or discrimination increases the risk of illnesses by 8.4%. When looking at firm characteristics, employment in larger firms increases the probability of work illnesses. Among personal characteristics, work illnesses are less likely to occur to males (-1.4%). Looking at worker’s area of residence, the geographic dummies are jointly significant and being in the south leads to a higher probability (about 0.9%) of illnesses as compared to being in the north-west. Finally, the marital status dummies jointly considered are significant, even if it is not possible to differentiate between all sorts of marital status, but the data show that being married

\textsuperscript{16}The average individual with a fixed-term contract instead of an open-ended one would have a 0.478\% higher probability of incurring work related illnesses.
increases the probability of illnesses (+0.9%) in comparison to be never married.

The main findings of the analysis, differing from other works [Guadalupe, 2003], highlight that work safety is not affected by worker contract type. Moreover, in line with Amuedo-Dorantes [2001], both work related accidents and work related illnesses seem to be influenced mainly by working conditions and by several personal characteristics variables.

The analysis on work safety continues with Table 6.3 that shows statistical inference regarding the repetition of accidents/illnesses within the 12 month reference period. Specifically, it presents the estimated coefficients and robust standard errors for two probit regression models. The dependent variable is a dummy variable that indicates if the employee experienced more than one as opposed to only one accident/illness. This analysis is useful in order to understand the causes of the repetition of accidents and illnesses. In fact, in the sample, among the employees that experienced work accidents about 7.8% reported more than one episode; and among the employees that suffered work illnesses about 13.1% cases reported more than one illness. The heteroskedasticity likelihood tests suggest the use of robust standard errors.

Table 6.3: Probability of accidents/illnesses repetition at the workplace conditional on having had at least one accident/illness – Probit

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Accidents Coefficient</th>
<th>Robust S E</th>
<th>Illnesses Coefficient</th>
<th>Robust S E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed-term contract</td>
<td>-0.077*</td>
<td>(0.202)</td>
<td>0.066</td>
<td>(0.114)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.202</td>
<td>(1.343)</td>
<td>-1.403*</td>
<td>(0.788)</td>
</tr>
<tr>
<td>Job characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-time contract</td>
<td>0.126</td>
<td>(0.222)</td>
<td>0.205*</td>
<td>(0.111)</td>
</tr>
<tr>
<td>Overtime hours</td>
<td>-0.173</td>
<td>(0.183)</td>
<td>0.050</td>
<td>(0.088)</td>
</tr>
<tr>
<td>Shift work</td>
<td>-0.069</td>
<td>(0.139)</td>
<td>-0.009</td>
<td>(0.070)</td>
</tr>
<tr>
<td>Manager or director as reference category</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White collar worker</td>
<td>-0.373</td>
<td>(0.495)</td>
<td>0.249*</td>
<td>(0.131)</td>
</tr>
<tr>
<td>Blue collar worker or apprentice</td>
<td>-0.011</td>
<td>(0.526)</td>
<td>0.036</td>
<td>(0.158)</td>
</tr>
<tr>
<td>Exposure to dangers such as dust, etc.</td>
<td>0.054</td>
<td>(0.137)</td>
<td>0.176**</td>
<td>(0.074)</td>
</tr>
<tr>
<td>Noisy workplace</td>
<td>0.021</td>
<td>(0.144)</td>
<td>-0.061</td>
<td>(0.078)</td>
</tr>
<tr>
<td>Bad posture induced by work</td>
<td>0.088</td>
<td>(0.134)</td>
<td>0.111</td>
<td>(0.070)</td>
</tr>
<tr>
<td>Feeling exposed to risk of injury</td>
<td>0.15</td>
<td>(0.142)</td>
<td>0.205***</td>
<td>(0.073)</td>
</tr>
<tr>
<td>Excessive workload</td>
<td>0.215</td>
<td>(0.151)</td>
<td>0.195***</td>
<td>(0.065)</td>
</tr>
<tr>
<td>Feeling exposed to bullying or discrimination</td>
<td>0.115</td>
<td>(0.184)</td>
<td>0.208***</td>
<td>(0.076)</td>
</tr>
<tr>
<td>Feeling exposed to threats or physical violence</td>
<td>0.358</td>
<td>(0.263)</td>
<td>0.131</td>
<td>(0.122)</td>
</tr>
<tr>
<td>Firm characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Establishment size</td>
<td>-0.003</td>
<td>(0.146)</td>
<td>-0.150*</td>
<td>(0.083)</td>
</tr>
<tr>
<td>Agriculture as reference category</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industry excluding construction</td>
<td>-0.098</td>
<td>(0.407)</td>
<td>0.148</td>
<td>(0.219)</td>
</tr>
<tr>
<td>Construction</td>
<td>0.080</td>
<td>(0.409)</td>
<td>0.219</td>
<td>(0.234)</td>
</tr>
</tbody>
</table>

Continued on next page
6. Safety at the workplace: accidents and illnesses

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Accidents Coefficient</th>
<th>Robust SE</th>
<th>Illnesses Coefficient</th>
<th>Robust SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retail</td>
<td>-0.746 (0.474)</td>
<td></td>
<td>0.130 (0.246)</td>
<td></td>
</tr>
<tr>
<td>Other activities</td>
<td>-0.138 (0.400)</td>
<td></td>
<td>0.295 (0.217)</td>
<td></td>
</tr>
</tbody>
</table>

**Personal characteristics**

<table>
<thead>
<tr>
<th></th>
<th>Accidents Coefficient</th>
<th>Robust SE</th>
<th>Illnesses Coefficient</th>
<th>Robust SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current job tenure</td>
<td>0.001 (0.002)</td>
<td></td>
<td>4.6e-04 (0.001)</td>
<td></td>
</tr>
<tr>
<td>Square of current job tenure</td>
<td>-2.4e-06 (5.5e-06)</td>
<td></td>
<td>-8.1e-07 (2.3e-06)</td>
<td></td>
</tr>
<tr>
<td>First job</td>
<td>-0.280* (0.162)</td>
<td></td>
<td>-0.079 (0.079)</td>
<td></td>
</tr>
<tr>
<td>Looking for another job</td>
<td>0.117 (0.183)</td>
<td></td>
<td>0.247** (0.101)</td>
<td></td>
</tr>
<tr>
<td>Number of household members</td>
<td>-0.032 (0.055)</td>
<td></td>
<td>-0.049 (0.033)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>-0.308* (0.159)</td>
<td></td>
<td>-0.136* (0.070)</td>
<td></td>
</tr>
<tr>
<td>Born in Italy</td>
<td>-0.186 (0.182)</td>
<td></td>
<td>-0.299*** (0.111)</td>
<td></td>
</tr>
<tr>
<td>Years of education</td>
<td>-0.024 (0.161)</td>
<td></td>
<td>-0.075 (0.070)</td>
<td></td>
</tr>
<tr>
<td>Square of years of education</td>
<td>2.1e-04 (0.007)</td>
<td></td>
<td>0.003 (0.003)</td>
<td></td>
</tr>
<tr>
<td>Educational activities in last four weeks</td>
<td>-0.036 (0.300)</td>
<td></td>
<td>0.013 (0.102)</td>
<td></td>
</tr>
<tr>
<td>Accidents occurred on the way to work</td>
<td>0.313 (0.312)</td>
<td></td>
<td>0.200 (0.167)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.011 (0.042)</td>
<td></td>
<td>-0.001 (0.026)</td>
<td></td>
</tr>
<tr>
<td>Square of age</td>
<td>-2.4e-04 (0.001)</td>
<td></td>
<td>6.9e-05 (2.9e-04)</td>
<td></td>
</tr>
<tr>
<td>North-west as reference category</td>
<td>g</td>
<td></td>
<td>h</td>
<td></td>
</tr>
<tr>
<td>North-east</td>
<td>0.315** (0.158)</td>
<td></td>
<td>0.263*** (0.087)</td>
<td></td>
</tr>
<tr>
<td>Central</td>
<td>0.092 (0.181)</td>
<td></td>
<td>0.054 (0.100)</td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>0.289 (0.182)</td>
<td></td>
<td>0.171* (0.092)</td>
<td></td>
</tr>
<tr>
<td>Islands</td>
<td>0.040 (0.256)</td>
<td></td>
<td>-0.012 (0.129)</td>
<td></td>
</tr>
<tr>
<td>Executive or intellectual occupation as ref. cat.</td>
<td>i</td>
<td></td>
<td>j</td>
<td></td>
</tr>
<tr>
<td>Technical position</td>
<td>0.163 (0.537)</td>
<td></td>
<td>-0.118 (0.148)</td>
<td></td>
</tr>
<tr>
<td>Office clerk and qualified occupation</td>
<td>0.172 (0.560)</td>
<td></td>
<td>0.160 (0.159)</td>
<td></td>
</tr>
<tr>
<td>Craftsman and operator of industrial machinery</td>
<td>-0.171 (0.573)</td>
<td></td>
<td>0.139 (0.176)</td>
<td></td>
</tr>
<tr>
<td>Never married as reference category</td>
<td>k</td>
<td></td>
<td>l</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>0.105 (0.162)</td>
<td></td>
<td>0.062 (0.100)</td>
<td></td>
</tr>
<tr>
<td>Separated or divorced</td>
<td>0.230 (0.234)</td>
<td></td>
<td>0.150 (0.130)</td>
<td></td>
</tr>
<tr>
<td>Widow/widower</td>
<td>0.056 (0.586)</td>
<td></td>
<td>-0.127 (0.238)</td>
<td></td>
</tr>
</tbody>
</table>

Number of obs 1061 2895

Regression Wald test \( \chi^2 (41) = 82.97*** \) \( \chi^2 (41) = 140.10*** \)

Pseudo R\(^2\) 0.101 0.066

Heteroskedasticity LR test \( \chi^2 (41) = 63.08** \) \( \chi^2 (41) = 142.14*** \)

Overall, the findings show that the type of contract does not seem to affect the probability of repetition of accidents/illnesses at the workplace. The variable fixed-term contract is not significant in either case, even if its sign is negative for work related ac-
cidents and positive for work related illnesses; on average, having a fixed-term contract reduces the likelihood of repetition of work related accidents by 0.868% and increases the likelihood of repetition of work related illnesses by 1.349%.

When focusing on work related accidents, it emerges that working conditions do not have a significant influence on the risk of accident repetition. On the other hand, both workers at their first job experience (-2.9%), and men (-4.0%) have a lower risk of accident repetition. When looking at work related illnesses, working conditions have a significant impact on the risk of illnesses repetition. For instance, exposure to dangers such as dust, fumes, chemicals, etc. increases the likelihood of repetition of illnesses by 3.6%, and feeling exposed to bullying or discrimination augments the probability by 4.4%.

6.4.2. The joint probability of accidents and illnesses at the workplace

The two elements of work safety were examined in a unified way as a robustness check for the results. The bivariate probit regression model jointly estimate the probability of accidents and of illnesses at the workplace. In this way, covariance in the unobservables of the two equations is allowed for. Table 6.4 reports a positive and significant $\rho$ meaning that there are unobservable factors that positively affect the two probabilities. This finding suggests that there is something similar to an unobserved proneness to having accidents that also positively affects the proneness to illnesses, in other words these two phenomena are better studied together.

Table 6.4.: Joint probability of accidents and illnesses at the workplace – Bivariate probit

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Accidents</th>
<th>Illnesses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>Robust S.E.</td>
</tr>
<tr>
<td>Fixed-term contract</td>
<td>-0.028a</td>
<td>(0.047)</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.787***</td>
<td>(0.328)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-time contract</td>
<td>0.125**</td>
<td>(0.051)</td>
</tr>
<tr>
<td>Overtime hours</td>
<td>0.087*</td>
<td>(0.047)</td>
</tr>
<tr>
<td>Shift work</td>
<td>0.159***</td>
<td>(0.034)</td>
</tr>
<tr>
<td>Manager or director as reference category</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White collar worker</td>
<td>0.049</td>
<td>(0.076)</td>
</tr>
<tr>
<td>Blue collar worker or apprentice</td>
<td>0.294***</td>
<td>(0.086)</td>
</tr>
<tr>
<td>Exposure to dangers such as dust, etc.</td>
<td>-0.012</td>
<td>(0.039)</td>
</tr>
<tr>
<td>Noisy workplace</td>
<td>0.082**</td>
<td>(0.040)</td>
</tr>
<tr>
<td>Bad posture induced by work</td>
<td>0.184***</td>
<td>(0.035)</td>
</tr>
<tr>
<td>Feeling exposed to risk of injury</td>
<td>0.490***</td>
<td>(0.036)</td>
</tr>
<tr>
<td>Excessive workload</td>
<td>0.030</td>
<td>(0.038)</td>
</tr>
</tbody>
</table>

Continued on next page
6. Safety at the workplace: accidents and illnesses

Table 6.4 – Continued from previous page

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Accidents</th>
<th>Illnesses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>Robust S E</td>
</tr>
<tr>
<td>Feeling exposed to bullying or discrimination</td>
<td>0.245***</td>
<td>(0.054)</td>
</tr>
<tr>
<td>Feeling exposed to threats or physical violence</td>
<td>0.168*</td>
<td>(0.090)</td>
</tr>
</tbody>
</table>

Firm characteristics

| Establishment size                                       | 0.037         | (0.035)       | 0.089***    | (0.027)    |
| Agriculture as reference category                        |               |               |             |            |
| Industry excluding construction                           | 0.105         | (0.094)       | -0.068      | (0.068)    |
| Construction                                              | 0.074         | (0.100)       | -0.028      | (0.075)    |
| Retail                                                    | 0.080         | (0.102)       | -0.102      | (0.074)    |
| Other activities                                          | 0.200**       | (0.094)       | 0.011       | (0.067)    |

Personal characteristics

| Current job tenure                                       | 6.2e-04       | (4.4e-04)     | 9.2e-04***  | (3.2e-04)  |
| Square of current job tenure                            | -1.9e-06*     | (1.1e-06)     | -8.3e-07    | (7.3e-07)  |
| First job                                                | -0.060        | (0.038)       | -0.100***   | (0.027)    |
| Looking for another job                                  | 0.179***      | (0.055)       | 0.347***    | (0.042)    |
| Number of household members                              | 0.011         | (0.014)       | -0.018*     | (0.011)    |
| Male                                                     | 0.134***      | (0.035)       | -0.140***   | (0.025)    |
| Born in Italy                                            | -0.036        | (0.051)       | -0.016      | (0.043)    |
| Years of education                                       | 0.018         | (0.036)       | -0.051**    | (0.024)    |
| Square of years of education                             | -0.001        | (0.001)       | 0.002**     | (0.001)    |
| Educational activities in last four weeks                | -0.106        | (0.071)       | 0.193***    | (0.041)    |
| Accidents occurred on the way to work                    | 0.317***      | (0.106)       | 0.457***    | (0.078)    |
| Age                                                      | -0.004        | (0.010)       | 0.020**     | (0.008)    |
| Square of age                                            | 4.7e-05       | (1.2e-04)     | -1.0e-04    | (9.3e-05)  |
| North-west as reference category                          |               |               |             |            |
| North-east                                               | 0.070*        | (0.038)       | 0.086***    | (0.030)    |
| Central                                                  | -0.011        | (0.044)       | -0.009      | (0.034)    |
| South                                                    | -0.102**      | (0.043)       | 0.089***    | (0.031)    |
| Islands                                                  | -0.091        | (0.056)       | -0.021      | (0.041)    |
| Executive or intellectual occupation as ref. cat.        |               |               |             |            |
| Technical position                                       | -0.006        | (0.081)       | 0.089*      | (0.046)    |
| Office clerk and qualified occupation                    | 2.5e-04       | (0.088)       | 0.086*      | (0.051)    |
| Craftsman and operator of industrial machinery            | 0.045         | (0.091)       | 0.049       | (0.058)    |
| Never married as reference category                      |               |               |             |            |
| Married                                                  | -0.037        | (0.040)       | 0.088***    | (0.032)    |
| Separated or divorced                                    | 0.064         | (0.066)       | 0.066       | (0.049)    |
| Widow/widower                                            | -0.017        | (0.129)       | 0.044       | (0.084)    |

Regression Wald test: $\chi^2(1) = 13.53^{***}$

Number of obs: 39471

Average partial effects: * -0.151%, 0.555%

Professional position dummies: $^a\chi^2(2) = 26.98^{***}$ $^b\chi^2(2) = 3.29$

Sector dummies: $^c\chi^2(4) = 12.56^{**}$ $^d\chi^2(4) = 12.66^{**}$

Geographic area dummies: $^e\chi^2(4) = 17.29^{**}$ $^f\chi^2(4) = 20.42^{***}$

Occupation dummies: $^g\chi^2(3) = 1.25^{*}$ $^h\chi^2(3) = 4.94$

Marital status dummies: $^i\chi^2(3) = 2.90^{*}$ $^j\chi^2(3) = 7.76^{*}$

Significance levels: * 0.10 ** 0.05 *** 0.01

Source: 2007 Istat Labour Force Survey
Table 6.4 reports results that are broadly in line with Table 6.2. The variable of interest, fixed-term contract, continues to have a negative non-significant coefficient for accidents. At the end of Table 6.4, the average partial effects\(^{17}\) for this variable is presented: workers that have a fixed-term contract have about a 0.2% lower probability of having accidents vis-a-vis workers with an open-ended contract.

In the model for illnesses, a second order polynomial to express the contribution of the years of attained education is significant. Figure 6.2 shows that workers with approximately 11 years of education (in Italy, 13 years corresponds to the completion of the high-school) experience the lowest contribution to the probability of having illnesses at the workplace.\(^{18}\)

### 6.4.3. The probability of accidents and illnesses at the workplace within three years of tenure

The last robustness check comes from a test of misspecification. Both RESET tests on the two models of Table 6.2 show evidence of misspecification.\(^{19}\) In Italy the contractual situation changes significantly after three years of tenure in the same firm. In fact, Legislative Decree n. 368 from 2001 providing the legal framework for fixed-term contracts, states that the total duration of a fixed-term relationship cannot exceed 36 months.\(^{20}\) Therefore, a poolability test for workers with tenure within and over three years of tenure.

\(^{17}\)Differently from the average effects quoted in the previous tables, the average partial effect is obtained by calculating the partial effect for all the observations in the sample and subsequently taking the average of these partial effects [Jones, 2007]. The result is less artificial than the average effect since there are dummies among the explanatory variables.

\(^{18}\)The partial effect of years of education in the probit estimation is

\[
\frac{\partial \Phi(\alpha d_i + \beta' X_i)}{\partial x_i} = \Phi(\alpha d_i + \beta' X_i)(\beta_1 + 2\beta_2 x_i),
\]

where \(\Phi(.)\) is the normal density function, \(x_i\) is the variable years of education, \(\beta_1\) is its coefficient for the linear part, and \(\beta_2\) is its coefficient for the quadratic part. Therefore, the turning point in the response probability is \(-\frac{\beta_1}{2\beta_2}\). With regard to the partial effect of years of education on the probability of illnesses, the turning point located at 11.05 is a minimum.

\(^{19}\chi^2(1) = 16.04***, \chi^2(1) = 26.18***\)

\(^{20}\)By virtue of article 4 of Legislative Decree n. 368/2001, the duration of an employment contract may be extended only once if its initial duration was less than three years; but even in this case, the total duration of that contract may not exceed three years. Article 10 contains a list of cases to which the rules for fixed-term contracts do not apply. For example, the 36 months rule does not apply to fixed-term employment relationships with a manager or director, whose maximum term is five years. In reality, article 5 of Legislative Decree n. 368/2001 allowed for several successive fixed-term contracts, respecting a specific period of employment interruption, twenty days from the date of expiration. This rule enabled elusive behaviors of the employers, by allowing the practical use of fixed-term contracts for a period exceeding 36 months. Law 247, 24 December 2007, prevents this practice, expressly providing that, although the time limit of employment interruption between
6. Safety at the workplace: accidents and illnesses

Figure 6.2.: Years of education partial effect on the probability of illnesses at the workplace – Table 6.4

years is performed for the two models of Table 6.2. The results conclusively reject the hypothesis of equal parameters. Hence, the analysis is restricted to only one of the two subsets of the dataset. Table 6.5 presents the results of the probit models for the work related accidents/illnesses, having restricted the dataset only to workers with tenure up to three years. This table is very reliable since the RESET tests on the models run on this restricted sample do not find misspecification.

<table>
<thead>
<tr>
<th>Table 6.5.: Probability of accidents and illnesses at the workplace within three years of tenure – Probit</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Independent Variable</strong></td>
</tr>
<tr>
<td>---------------------------</td>
</tr>
<tr>
<td>Fixed-term contract</td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td>Full-time contract</td>
</tr>
</tbody>
</table>

Job characteristics

Continued on next page

two subsequent fixed-term contracts remains valid, the employment relationship between the parties cannot exceed a total of 36 months, including extensions and renewals, regardless of outages between one contract and another.

\[ \chi^2(42) = 57.45^*, \chi^2(42) = 65.53^{**} \]

\[ \chi^2(1) = 2.52, \chi^2(1) = 1.02 \]
6. Safety at the workplace: accidents and illnesses

Table 6.5 – Continued from previous page

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Accidents Coefficient</th>
<th>Robust S E</th>
<th>Illnesses Coefficient</th>
<th>Robust S E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overtime hours</td>
<td>0.091 (0.089)</td>
<td></td>
<td>0.058 (0.075)</td>
<td></td>
</tr>
<tr>
<td>Shift work</td>
<td>0.165** (0.071)</td>
<td></td>
<td>0.014 (0.059)</td>
<td></td>
</tr>
<tr>
<td>Manager or director as reference category</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White collar worker</td>
<td>0.066 (0.203)</td>
<td>-0.042 (0.116)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blue collar worker or apprentice</td>
<td>0.330 (0.212)</td>
<td>0.038 (0.132)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exposure to dangers such as dust, etc.</td>
<td>0.128* (0.074)</td>
<td>0.082 (0.065)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noisy workplace</td>
<td>0.006 (0.077)</td>
<td>0.189*** (0.068)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bad posture induced by work</td>
<td>0.143** (0.068)</td>
<td>0.580*** (0.057)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling exposed to risk of injury</td>
<td>0.540*** (0.070)</td>
<td>0.227*** (0.062)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excessive workload</td>
<td>-0.098 (0.085)</td>
<td>0.460*** (0.057)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling exposed to bullying or discrimination</td>
<td>0.482*** (0.113)</td>
<td>0.514*** (0.086)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling exposed to threats or physical violence</td>
<td>0.219 (0.214)</td>
<td>0.291* (0.158)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Establishment size</td>
<td>-0.078 (0.060)</td>
<td>0.003 (0.051)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agriculture as reference category</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industry excluding construction</td>
<td>0.191 (0.161)</td>
<td>-0.065 (0.119)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Construction</td>
<td>0.098 (0.167)</td>
<td>0.017 (0.124)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retail</td>
<td>0.205 (0.171)</td>
<td>-0.083 (0.126)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other activities</td>
<td>0.221 (0.161)</td>
<td>-0.063 (0.113)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current job tenure</td>
<td>0.004 (0.009)</td>
<td>1.5e-04 (0.007)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Square of current job tenure</td>
<td>-5.9e-05 (2.2e-04)</td>
<td>-1.1e-04 (1.8e-04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Looking for another job</td>
<td>-0.142 (0.090)</td>
<td>-0.293*** (0.078)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of household members</td>
<td>1.9e-04 (0.028)</td>
<td>-0.018 (0.021)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.266*** (0.073)</td>
<td>-0.212*** (0.055)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Born in Italy</td>
<td>-0.055 (0.084)</td>
<td>0.090 (0.075)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years of education</td>
<td>0.071 (0.070)</td>
<td>-0.099* (0.052)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Square of years of education</td>
<td>-0.003 (0.003)</td>
<td>0.004** (0.002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Educational activities in last four weeks</td>
<td>-0.148 (0.159)</td>
<td>0.255*** (0.095)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accidents occurred on the way to work</td>
<td>0.229 (0.198)</td>
<td>0.503*** (0.160)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.002 (0.017)</td>
<td>0.007 (0.014)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Square of age</td>
<td>-7.5e-05 (2.2e-04)</td>
<td>3.0e-05 (1.8e-04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North-west as reference category</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North-east</td>
<td>0.095 (0.074)</td>
<td>0.226*** (0.063)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Central</td>
<td>-0.004 (0.085)</td>
<td>0.004 (0.074)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>-0.240*** (0.089)</td>
<td>0.083 (0.068)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Islands</td>
<td>-0.200* (0.108)</td>
<td>-0.143 (0.091)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Executive or intellectual occupation as ref. cat.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technical position</td>
<td>-0.264 (0.176)</td>
<td>0.005 (0.108)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Office clerk and qualified occupation</td>
<td>-0.185 (0.177)</td>
<td>-0.034 (0.119)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Craftsman and operator of industrial machinery</td>
<td>-0.064 (0.182)</td>
<td>-0.199 (0.131)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never married as reference category</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>-0.039 (0.079)</td>
<td>0.096 (0.061)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Separated or divorced</td>
<td>-0.035 (0.144)</td>
<td>0.118 (0.104)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Widow/widower</td>
<td>-0.115 (0.323)</td>
<td>0.109 (0.211)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of obs 11208 11174

Continued on next page
6. Safety at the workplace: accidents and illnesses

When looking for the differences between Table 6.2 and Table 6.5, the variable fixed-term contract continues not to be significant in explaining the probability of incurring accidents and the probability of incurring illnesses. Regarding the probability of accidents, the variables full-time contract, overtime hours, noisy workplace, feeling exposed to threats or physical violence, sector dummies, accidents occurred on the way to work are no longer significant. Furthermore, where the probability of illnesses is concerned, the variables overtime hours, shift work, exposure to dangers such as dust, establishment size, sector dummies, current job tenure, number of household members, age, marital status dummies are also no longer significant. Finally, many variables about working conditions and the variables regarding shift work (only for accidents), professional position dummies (only for accidents), first job (only for illnesses), looking for another job, being male, years of education (only for illnesses), and geographic area dummies continue to be significant and have the same sign on the coefficients in explaining work safety.

6.5. Conclusion

Work safety is a relevant issue at stake in modern political debates. In Europe, there is the widespread conception that the contractual position of the worker is a relevant determinant for work related accidents and illnesses. In order to investigate these ideas, individual level data from the 2007 Italian Labour Force Survey and its ad hoc module on work safety are employed.

Probit regressions for the occurrence and the repetition of accidents and illnesses at the workplace are performed. Medical literature reports a strong connection between

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>( \chi^2(41) = 416.27^{***} )</th>
<th>( \chi^2(41) = 745.03^{***} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression Wald test</td>
<td>0.144</td>
<td>0.183</td>
</tr>
<tr>
<td>Heteroskedasticity LR test</td>
<td>( \chi^2(41) = 388.43^{***} )</td>
<td>( \chi^2(41) = 792.99^{***} )</td>
</tr>
</tbody>
</table>

Average effects: 
- \( a = 0.266\% \)
- \( b = 0.040\% \)
- \( c = 0.266\% \)
- \( d = 0.040\% \)
- \( e = 0.266\% \)
- \( f = 0.040\% \)
- \( g = 0.266\% \)
- \( h = 0.040\% \)
- \( i = 0.266\% \)
- \( j = 0.040\% \)
- \( k = 0.266\% \)
- \( l = 0.040\% \)

Significance levels: * 0.10 ** 0.05 *** 0.01

Source: 2007 Istat Labour Force Survey
temporary work and psychological morbidity [Ferrie et al., 2008, Virtanen et al., 2005].
The results, however, seem to be in line with the majority of works in economic literature
[Amuedo-Dorantes, 2001, Hernanz and Toharia, 2006]. This paper shows that workplace
safety is mainly affected by working conditions, and to a lesser extent, by a worker's
personal characteristics and that the type of contract a worker has does not seem to not
influence accidents and illnesses at the workplace.

By using a bivariate probit, the results are confirmed and unobservable factors that
are positively correlated with the probability of incurring workplace accidents and the
probability of incurring workplace illnesses are found. Diagnostic tools and Italian leg-
islation indicated running probit regressions restricted to the sample of workers with up
to three years of tenure. Even though some variables do not seem to be important in
this last model specification, the findings on the main variables continue to hold.

All the model specifications reveal the importance of working conditions for non-
fatal accidents and non-fatal illnesses at the workplace. Assuming that these findings
also hold for fatal accidents and illnesses (not taken in account in this paper), the
improvement of working conditions should be the priority for new policies on work
safety. This would mean: noise reduction at the workplace, correct posture requirements,
reducing excessive workload, and the elimination of behaviour connected to bullying,
harassment or discrimination.

6.6. Appendix: Descriptive statistics

Table 6.6 provides descriptive statistics for the entire sample, for the workers who ex-
perienced accidents or illnesses, and for the workers who experienced no accidents and
no illnesses, reporting the mean and the standard deviation of the variables used in the
analysis, grouped by job, firm, and personal characteristics. The results of the test of
proportion/t test for the difference of means of workers who experienced no accidents
and no illnesses with respect to workers with accidents or illnesses is finally reported.
### Table 6.6.: Descriptive statistics of variables used in the analysis

<table>
<thead>
<tr>
<th>Dependent variables</th>
<th>All workers</th>
<th></th>
<th>Workers with accidents or illnesses</th>
<th></th>
<th>Workers with no accidents and no illnesses</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
<td>N. of obs.</td>
<td>Mean</td>
<td>Std. Dev.</td>
<td>N. of obs.</td>
</tr>
<tr>
<td>Work accidents</td>
<td>0.027</td>
<td>0.162</td>
<td>4273</td>
<td>0.283</td>
<td>0.450</td>
<td>40347</td>
</tr>
<tr>
<td>Work accidents repetition</td>
<td>0.078</td>
<td>0.268</td>
<td>1209</td>
<td>0.078</td>
<td>0.268</td>
<td>40216</td>
</tr>
<tr>
<td>Work illnesses</td>
<td>0.073</td>
<td>0.261</td>
<td>4273</td>
<td>0.765</td>
<td>0.424</td>
<td>40216</td>
</tr>
<tr>
<td>Work illnesses repetition</td>
<td>0.131</td>
<td>0.337</td>
<td>3267</td>
<td>0.131</td>
<td>0.337</td>
<td>40216</td>
</tr>
<tr>
<td>Fixed-term contract</td>
<td>0.139</td>
<td>0.345</td>
<td>4278</td>
<td>0.112</td>
<td>0.315</td>
<td>40347</td>
</tr>
<tr>
<td>Full-time contract</td>
<td>0.858</td>
<td>0.349</td>
<td>4278</td>
<td>0.896</td>
<td>0.306</td>
<td>40347</td>
</tr>
<tr>
<td>Overtime hours</td>
<td>0.085</td>
<td>0.278</td>
<td>4256</td>
<td>0.133</td>
<td>0.340</td>
<td>40068</td>
</tr>
<tr>
<td>Shift work</td>
<td>0.222</td>
<td>0.416</td>
<td>4275</td>
<td>0.335</td>
<td>0.472</td>
<td>40299</td>
</tr>
<tr>
<td>Manager or director</td>
<td>0.102</td>
<td>0.302</td>
<td>4278</td>
<td>0.104</td>
<td>0.306</td>
<td>40347</td>
</tr>
<tr>
<td>White collar worker</td>
<td>0.416</td>
<td>0.493</td>
<td>4278</td>
<td>0.363</td>
<td>0.481</td>
<td>40347</td>
</tr>
<tr>
<td>Blue collar worker or apprentice</td>
<td>0.483</td>
<td>0.500</td>
<td>4278</td>
<td>0.532</td>
<td>0.499</td>
<td>40347</td>
</tr>
<tr>
<td>Exposure to dangers such as dust, etc.</td>
<td>0.163</td>
<td>0.370</td>
<td>4258</td>
<td>0.331</td>
<td>0.471</td>
<td>40037</td>
</tr>
<tr>
<td>Noisy workplace</td>
<td>0.154</td>
<td>0.361</td>
<td>4251</td>
<td>0.325</td>
<td>0.468</td>
<td>39905</td>
</tr>
<tr>
<td>Bad posture induced by work</td>
<td>0.201</td>
<td>0.401</td>
<td>4254</td>
<td>0.475</td>
<td>0.499</td>
<td>40029</td>
</tr>
<tr>
<td>Feeling exposed to risk of injury</td>
<td>0.220</td>
<td>0.414</td>
<td>4247</td>
<td>0.461</td>
<td>0.499</td>
<td>39758</td>
</tr>
<tr>
<td>Excessive workload</td>
<td>0.141</td>
<td>0.348</td>
<td>4237</td>
<td>0.351</td>
<td>0.477</td>
<td>39910</td>
</tr>
<tr>
<td>Feeling exposed to bullying or discrimination</td>
<td>0.050</td>
<td>0.218</td>
<td>4228</td>
<td>0.180</td>
<td>0.385</td>
<td>40012</td>
</tr>
<tr>
<td>Feeling exposed to threats or physical violence</td>
<td>0.015</td>
<td>0.123</td>
<td>4251</td>
<td>0.052</td>
<td>0.223</td>
<td>40600</td>
</tr>
<tr>
<td>Establishment size</td>
<td>0.724</td>
<td>0.447</td>
<td>4278</td>
<td>0.791</td>
<td>0.406</td>
<td>40347</td>
</tr>
<tr>
<td>Sector</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>0.029</td>
<td>0.168</td>
<td>4278</td>
<td>0.028</td>
<td>0.165</td>
<td>40347</td>
</tr>
<tr>
<td>Industry excluding construction</td>
<td>0.241</td>
<td>0.428</td>
<td>4278</td>
<td>0.238</td>
<td>0.426</td>
<td>40347</td>
</tr>
<tr>
<td>Construction</td>
<td>0.073</td>
<td>0.260</td>
<td>4278</td>
<td>0.083</td>
<td>0.276</td>
<td>40347</td>
</tr>
<tr>
<td>Retail</td>
<td>0.115</td>
<td>0.319</td>
<td>4278</td>
<td>0.079</td>
<td>0.271</td>
<td>40347</td>
</tr>
<tr>
<td>Other activities</td>
<td>0.542</td>
<td>0.498</td>
<td>4278</td>
<td>0.572</td>
<td>0.495</td>
<td>40347</td>
</tr>
</tbody>
</table>

**Continued on next page**
Table 6.6 – Continued from previous page

<table>
<thead>
<tr>
<th></th>
<th>All workers</th>
<th>Workers with accidents or illnesses</th>
<th>Workers with no accidents and no illnesses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
<td>N. of obs. Mean</td>
</tr>
<tr>
<td>Current job tenure</td>
<td>1.38</td>
<td>1.23</td>
<td>4278</td>
</tr>
<tr>
<td>Square of current job tenure</td>
<td>34.425</td>
<td>4890.2</td>
<td>4278</td>
</tr>
<tr>
<td>First job</td>
<td>0.303</td>
<td>0.460</td>
<td>4278</td>
</tr>
<tr>
<td>Looking for another job</td>
<td>0.053</td>
<td>0.224</td>
<td>4278</td>
</tr>
<tr>
<td>Number of household members</td>
<td>3.187</td>
<td>1.101</td>
<td>4013</td>
</tr>
<tr>
<td>Male</td>
<td>0.564</td>
<td>0.496</td>
<td>4278</td>
</tr>
<tr>
<td>Born in Italy</td>
<td>0.929</td>
<td>0.258</td>
<td>4278</td>
</tr>
<tr>
<td>Years of education</td>
<td>11.803</td>
<td>3.451</td>
<td>4278</td>
</tr>
<tr>
<td>Square of years of education</td>
<td>1.51</td>
<td>86.489</td>
<td>4278</td>
</tr>
<tr>
<td>Educational activities in last four weeks</td>
<td>0.052</td>
<td>0.222</td>
<td>4266</td>
</tr>
<tr>
<td>Accidents occurred on the way to work</td>
<td>0.011</td>
<td>0.104</td>
<td>4273</td>
</tr>
<tr>
<td>Age</td>
<td>41.277</td>
<td>11.074</td>
<td>4278</td>
</tr>
<tr>
<td>Square of age</td>
<td>1826.430</td>
<td>911.631</td>
<td>4278</td>
</tr>
<tr>
<td>Geographic area</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North-west</td>
<td>0.279</td>
<td>0.448</td>
<td>4278</td>
</tr>
<tr>
<td>North-east</td>
<td>0.231</td>
<td>0.422</td>
<td>4278</td>
</tr>
<tr>
<td>Central</td>
<td>0.161</td>
<td>0.367</td>
<td>4278</td>
</tr>
<tr>
<td>South</td>
<td>0.227</td>
<td>0.419</td>
<td>4278</td>
</tr>
<tr>
<td>Islands</td>
<td>0.102</td>
<td>0.303</td>
<td>4278</td>
</tr>
<tr>
<td>Occupation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Executive or intellectual occupation</td>
<td>0.113</td>
<td>0.316</td>
<td>4210</td>
</tr>
<tr>
<td>Technical position</td>
<td>0.240</td>
<td>0.427</td>
<td>4210</td>
</tr>
<tr>
<td>Office clerk and qualified occupation</td>
<td>0.273</td>
<td>0.446</td>
<td>4210</td>
</tr>
<tr>
<td>Craftsman and operator industrial machinery</td>
<td>0.375</td>
<td>0.484</td>
<td>4210</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never married</td>
<td>0.320</td>
<td>0.467</td>
<td>4278</td>
</tr>
<tr>
<td>Married</td>
<td>0.603</td>
<td>0.489</td>
<td>4278</td>
</tr>
<tr>
<td>Separated or divorced</td>
<td>0.060</td>
<td>0.238</td>
<td>4278</td>
</tr>
<tr>
<td>Widow/widower</td>
<td>0.016</td>
<td>0.127</td>
<td>4278</td>
</tr>
</tbody>
</table>

Significance levels (test of proportion/t test) – difference with respect to workers with accidents or illnesses: * 0.10 ** 0.05 *** 0.01
Part IV.

Government and union
7. Government fiscal efforts vs. labour union strikes: it takes two to tango

7.1. Introduction

The interaction between union and government strategies is particularly important to understanding the economic and political performance of unionized economies. The relationship between union achievements in wage bargaining and government agendas in political economy deserves in-depth analysis. In wage bargaining, unions and governments have conflicting aims: unions want to maximize real wages and governments want to minimize unemployment. However, both agents have similar forms of political commitment: unions and governments want to maximize consensus among workers and voters. So, unions call ‘economic’ and ‘political’ strikes to maximize real wages and consensus among workers, and governments enact policies to maximize employment and consensus among voters.

Several theories have been developed to explain economic strikes against employers. Hicks [1932] pointed out that strikes are costly for all agents participating in industrial relationships. Thus, if agents were rational and well-informed, there would be no reason to strike. Ashenfelter and Johnson [1969], starting from the idea that union leaders are motivated by personal advancement and union growth [Ross, 1953], focused on the lack of information among workers and found empirical evidence that strike probabilities are linked to real wages.\footnote{Shalev [1980], Kennan and Wilson [1989], and Mumford [1993] surveyed the literature on industrial relations theory.}

Although workers usually strike to gain higher wages, this is not the only reason for strikes. Workers may try to influence government policy by means of political strikes against governments. Azam and Salmon [2004] applied an imperfect information framework in a policy-game between a union and a government. In this game, the government cannot commit credibly to a given level of employment-generating public expenditure and the union may use political strikes to force the government to increase the same public expenditures. Following a similar line, Vernby [2007] developed a policy-game
Government fiscal efforts vs. labour union strikes: it takes two to tango

in which the union has the power to call strikes in order to entice the government to increase its efforts for job creation. Contrary to Azam and Salmon [2004], he focused on how the incentive for the government to pursue an accommodationist policy depends on the type of electoral system. However, these studies impose a sequential structure of the game in which the union moves first and sets wages unilaterally, without taking into account the structure of dynamic adjustment towards the equilibrium.

Recently, Franzosi [2006] has argued that a government’s partisanship and fiscal policy affect union behavior and strikes. Fiscal policies aimed at mitigation of unemployment, stimulation of the economy and stabilization of the business cycle are common across countries and over time. Data for OECD countries reveal that: (i) fiscal stimuli based on tax cuts are more likely to increase growth than those based upon spending increases; (ii) adjustments on the spending side rather than on the tax side are less likely to create recessions [Alesina and Ardagna, 2009]; (iii) expansionary fiscal contractions are effective under certain limited hypothesis [Giavazzi and Pagano, 1990]; (iv) fiscal contraction implemented through cuts in government spending has a more expansionary effect [Ardagna, 2004]; (v) credible fiscal consolidation has a contractionary effect albeit under fiscal stress measured by high debt, deficit, tax burden, and government spending relative to GDP, the effect may be expansionary [Perotti, 1999].

Schneider and Zapal [2006], in order to capture the restrictiveness or expansiveness of fiscal policy, applied the growth accounting technique proposed by Von Hagen et al. [2002] and used by Hughes-Hallett et al. [2003]. They defined net fiscal effort as the part of the change in the budget balance not due to growth of the economy, a change in monetary policy conditions, or a change in the level of public debt. In particular, the overall balance (government budget surplus) denotes the fiscal efforts required to achieve government targets to maintain fiscal sustainability. In this framework, the government can not arbitrarily use fiscal efforts to minimize unemployment and maximize its consensus between voters, so the strategic interaction between the union and the government is crucial to the effectiveness of fiscal efforts.

Several stylized facts have emerged on strikes and fiscal policies. Strike activity has a very long history in Western countries. Over time, strikes appear cyclically and occur in

---

2In the 1970s and 1980s, several authors investigated the interaction between monetary policy and wage setting in terms of policy-games. Gylfason and Lindbeck [1994] analyzed first the non-neutrality of money through a policy-game between the government and the labour union showing that monetary expansion stimulates output and employment despite the optimal reaction of the unions as long as they care about inflation.

3The efficacy of fiscal policy is, in general, a controversial issue. Corsetti et al. [2012], Davig and Leeper [2011], and Leeper et al. [2010], using the dynamic stochastic general equilibrium modeling approach, have recently addressed this issue.

---
Government fiscal efforts vs. labour union strikes: it takes two to tango

waves. For example, during the 1970s and 1980s OECD countries had been experiencing a long strike wave, while a sharp decline in strikes has occurred in the last decades. In addition, since the 1980s the motivations for strikes have been changing: economic strikes have been declining and political strikes have been rising.\(^4\) Public debt problems have been the main policy issue in Western countries since World War II. Since the 1970s, OECD countries have experienced large and persistent public deficits that, in turn, increased the public debt-to-GDP ratio. Recently, several OECD countries have adopted a ‘fiscal straitjacket’ and have engaged in fiscal adjustment programs to reduce the government debt-to-GDP ratio and/or the budget deficit-to-GDP ratio.\(^5\) For example, in the European Union, the Maastricht Treaty and the Stability and Growth Pact have imposed institutional constraints on national fiscal policies, and, more recently, the Fiscal Compact has imposed severe fiscal consolidation. However, the interaction between fiscal policy, unemployment, strikes, and wages is still a puzzle.

To deal with this puzzle, this work proposes a novel policy-game model where the government and the labour union simultaneously interact with perfect information. In particular, the policy-game model focuses on the interaction between political strikes and fiscal efforts. Then, by means of a cointegrated Vector Autoregressive (VAR) model, the policy-game model is estimated and tested using Italian quarterly data from the period 1960-2009. By interpreting the players’ reaction functions as the long-term economic relationships and the Cournot-Nash equilibrium of the game as the steady state equilibrium, the model explains how a labour union, mainly concerned with the wages of workers, and a government, mainly concerned with unemployment, may jointly determine the optimal level of their control variables: strikes and fiscal efforts.

The Italian economy represents an ideal setting to investigate the interaction between governments and unions because its history and institutional configuration in the post-World War II era, suggests that policy questions revolved around the interactions of fiscal policy, unemployment, strikes, and wages setting are highlighted. In Italy, the importance of labour unions both in industrial relationships and in economic and social policy issues is a well-known historical fact. Despite the reduction in the rate of union

\(^4\)There are various definition of the term political strike and several measures of strikes. Hamann et al. [2012], following Hyman [1989] and Walsh [1983], defined ‘general strike’ (or ‘protest strike’) as a temporary, national stoppage of work by workers from many industries, directed against the executive or legislative arms of government, to enforce a demand or give voice to a grievance. By means of this definition, they distinguished and collected data on political and economic strikes for 16 Western European countries between 1980 and 2006.

\(^5\)Mierau et al. [2007], using a panel discrete choice model for 20 OECD countries for the period 1970-2003, found that fiscal adjustments are affected by economic and political variables (as upcoming elections or broad policy reform). For advanced countries, Lavigne [2011] found that fiscal rules contribute to avoiding situations of fiscal distress.
density and the technological changes occurring in recent years, labour unions still play an important role in the performance of the labour market.\footnote{In the post-war period, the evolution of the Italian union density of employed workers displays a long cycle of about 30 years, with two peaks (about 50%), one in 1950 and one in 1976. In 1960 and in 2009, the Italian union density was about 25% and 34%, respectively. For more details see Checchi and Corneo [2000] and Franzosi [2006].} In addition, since labour market institutions have not changed during recent years, labour unions can still affect the political and social climate.\footnote{The most important pro-labour legislative act, in Italy, was the Workers’ charter of rights, enacted in 1970. The Italian constitutional law has increased work council powers and has protected union activity. Moreover, labour courts have extended the applicability of national wage agreement to all workers.} In Italy, the experience of ‘technical governments’, of ‘national unity governments’, of ‘national solidarity governments’, and of ‘historic compromise governments’ are examples of consociationalism that have dominated the political and social culture of the country, for quite some time.\footnote{The high Italian debt originated from the fiscal policy followed in the 1970s and 1980s. The Italian government implemented expansionary policies via a continuous increase in public spending (deficit spending) to support the economic growth and to satisfy the claims of the workers (welfare state), during this period, characterized by the oil crises of the 1970s, and the first slowdown in economic growth after the economic boom of the 1960s.}

Overall, the empirical results show that the predicted Cournot-Nash equilibrium is supported by the data. This evidence suggests that the government and the labour union are both powerful economic players, and each is capable of affecting the choice of the other player. Fiscal efforts and strikes are jointly determined, in the long run, as predicted by the strategic policy game. Moreover, an additional cointegrating relationship, interpreted as a statistical Phillips-type relationship, is identified.

Following the approach developed by Azam and Salmon [2004] and Vernby [2007], this study contributes to the existing literature in several respects.

(i) A novel policy-game model analyzes the simultaneous strategic interaction between the government and the labour union. Their reaction functions characterize the long-term economic relationships and the steady state equilibrium of fiscal efforts and strikes.

(ii) The policy-game model is estimated by using Italian quarterly data and the Vector Error Correction (VEC) methodology to investigate the structure of dynamic adjustment towards equilibrium, distinguishing between short run and long run adjustment.

(iii) The new approach for the strategic interaction between the labour union and the government allows a phase diagram representation of the estimated policy-game model that sheds light on the adjustment features of the players.
7. Government fiscal efforts vs. labour union strikes: it takes two to tango

The remainder of the paper is organized as follows. In Section 7.2, the policy-game model is introduced with the discussion on its main implications. In Section 7.3, the empirical investigation of the proposed policy-game is done by the VEC methodology. Specifically, in Sub-section 7.3.1 the focus is on the Cournot-Nash equilibrium, and in Sub-section 7.3.2 an interpretation by means of phase diagrams is provided. Section 7.4 concludes the paper.

7.2. Model

In this section, a policy-game model analyzes the strategic interaction between two players: the government \((g)\) and the labour union \((l)\).\(^9\) Starting from the objective function of each player, their reaction functions, the Cournot-Nash equilibrium of the game, and its properties are obtained.

The player’s objective functions are defined on the government’s effort \((e)\), the labour union’s strikes \((s)\), unemployment \((u)\), and wages \((w)\).\(^{10}\) All the variables and the parameters of the objective functions are positively defined: \(e, s, u, w, t, r, \tau, \rho, \sigma_i, \eta > 0\), with \(i = g, l\).

When the government and the labour union interact, the government uses its effort to maximize consensus among voters that depends on unemployment. Thus, the government’s objective function \((G)\), for any level of labour union’s strikes, depends on the control variable \(e\) and the variable \(u\). Specifically,

\[
G = tu - \tau s - \sigma g e - e^{1 + \eta g} \tag{7.1}
\]

The derivative of \(G\) with respect to the control variable of the player \(g\) is

\[
G_e = tu - \tau s - \sigma g e - (1 + \eta g) e^{\eta g} \tag{7.2}
\]

where \(tu - \tau s - \sigma g\) represents the constant marginal benefit of \(e\) in terms of consensus among voters, and \((1 + \eta g) e^{\eta g}\) is the increasing marginal cost of \(e\). Note that the constant marginal benefit of \(e\) decreases in \(s\) as voters punish the government for the inconvenience caused by strikes and mistrust its effort. In particular, since Eq. (7.1) captures the idea

\(^{9}\)Since the focus is on the interaction between political strikes and fiscal efforts, other players (e.g. the central bank and the employers’ association) are assumed to have a negligible effect on the interplay between \(g\) and \(l\). This orthogonality assumption is indeed supported by the empirical analysis.

\(^{10}\)The implicit assumption is that the long run relationships among \(e, s, u\) and \(w\) do not depend on other variables. In the empirical analysis, however, other variables that could affect the short run dynamics, as inflation \((\pi)\), are considered.
7. Government fiscal efforts vs. labour union strikes: it takes two to tango

of a trade-off between costs and benefits of the government’s effort, \( G_e \) has an inverted-U shape, that is the government’s control variable \( e \) has a positive and then a negative effect on \( G \):

\[
G_{e=0} = tu^{-\tau}s^{-\sigma_s} > 0 \quad G_{e\rightarrow+\infty} = -\infty < 0 \quad (7.3)
\]

Since unemployment and strikes reduce government consensus among voters, the variables \( u \) and \( s \) have always negative effect on \( G \).

\[
G_u = -t\tau u^{-\tau-1}s^{-\sigma_s}e < 0 \quad (7.4)
\]

\[
G_s = -tu^{-\tau}\sigma_g s^{-\sigma_s-1}e < 0 \quad (7.5)
\]

In the strategic interaction between the government and the labour union, the labour union uses strikes to maximize consensus among workers that are affected by wages. Therefore, for any given government’s effort, the labour union’s objective function \( (L) \) depends on the control variable \( s \) and the variable \( w \). Specifically,

\[
L = rw^{-\rho}\eta_l s - s^{1+\sigma_l} \quad (7.6)
\]

The derivative of \( L \) with respect to the control variable of the player \( l \) is

\[
L_s = rw^{-\rho}\eta_l s - (1 + \sigma_l)s^{\sigma_l} \quad (7.7)
\]

where \( rw^{-\rho}\eta_l s \) represents the constant marginal benefit of \( s \) in terms of consensus among workers, and \( (1 + \sigma_l)s^{\sigma_l} \) is the increasing marginal cost of \( s \). Note also that the constant marginal benefit of \( s \) increases in \( e \) as workers reward the union inducing higher government’s effort and believe more in strikes efficacy. By analogy with the government trade-off between costs and benefits, \( L_s \) has an inverted-U shape, that is the labour union’s control variable \( s \) has a positive and then a negative effect on \( L \):

\[
L_{s=0} = rw^{-\rho}\eta_l s > 0 \quad L_{s\rightarrow+\infty} = -\infty < 0 \quad (7.8)
\]

Finally, since higher wages reduce labour conflicts, the variable \( w \) has negative effect on \( L \) lowering the labour union consensus among workers. The variable \( e \) has positive effect on \( L \) because effort increases labour union consensus among workers.

\[
L_w = -rw^{-\rho-1}\eta_l s < 0 \quad (7.9)
\]

\[
L_e = rw^{-\rho}\eta_l s^{-1}e > 0 \quad (7.10)
\]
7. Government fiscal efforts vs. labour union strikes: it takes two to tango

7.2.1. Reaction functions

In the symmetric policy-game both players, \( g \) and \( l \), simultaneously choose their respective control variable, \( e \) and \( s \), to maximize their objective functions. Given the strategic interaction between two players, the optimal strategies arise from their reaction functions. As is well known, the government’s reaction function is obtained by maximizing its objective function. From the First Order Condition of this problem (\( G_e = 0 \) in Eq. (7.2)), the government’s reaction function expressed in logs is obtained:

\[
\ln e^C = \frac{1}{\eta_g} \ln \left( \frac{t}{1 + \eta_g} \right) - \frac{\tau}{\eta_g} \ln u - \frac{\sigma_g}{\eta_g} \ln s \tag{7.11}
\]

Since \(-\frac{\sigma_g}{\eta_g} < 0\), the government’s reaction function has a negative slope. Therefore, strikes are strategic substitute by government perspective, meaning that a more ‘aggressive’ labour union lowers the marginal benefit of the government’s effort. In order to simplify notation, Eq. (7.11) is rewritten in the form

\[
\ln e^C = \beta^0_g + \beta^u_g \ln u + \beta^s_g \ln s \tag{7.12}
\]

where \( \beta^0_g \leq 0, \beta^u_g, \beta^s_g < 0 \), given that all the parameters of the objective functions are positive by assumption.

The labour union’s reaction function is calculated by maximizing its objective function. Imposing \( L_s = 0 \) in Eq. (7.7) and taking logs, the labour union’s reaction function is obtained.

\[
\ln s^C = \frac{1}{\sigma_l} \ln \left( \frac{r}{1 + \sigma_l} \right) - \frac{\rho}{\sigma_l} \ln w + \frac{\eta_l}{\sigma_l} \ln e \tag{7.13}
\]

Since \( \frac{\eta_l}{\sigma_l} > 0 \), the labour union’s reaction function has a positive slope. By labour union perspective effort is strategic complement, meaning that a more aggressive government raises the marginal benefit of strike. As above, Eq. (7.13) is rewritten in the form

\[
\ln s^C = \beta^0_l + \beta^w_l \ln w + \beta^e_l \ln e \tag{7.14}
\]

where \( \beta^0_l \leq 0, \beta^w_l < 0, \beta^e_l > 0 \) by construction.
7. Government fiscal efforts vs. labour union strikes: it takes two to tango

7.2.2. Cournot-Nash equilibrium

Assuming an interior solution, the Cournot equilibrium associated with the game between \(g\) and \(l\) is obtained from Eq. (7.12) and (7.14).

\[
\ln e^* = \frac{\beta^0_g + \beta^s_g \beta^0_l}{1 - \beta^s_g \beta^c_l} + \frac{\beta^u_g}{1 - \beta^s_g \beta^c_l} \ln u + \frac{\beta^w_g \beta^s_l}{1 - \beta^s_g \beta^c_l} \ln w
\] (7.15)

\[
\ln s^* = \frac{\beta^0_l + \beta^c_l \beta^0_g}{1 - \beta^s_g \beta^c_l} + \frac{\beta^u_l \beta^c_g}{1 - \beta^s_g \beta^c_l} \ln u + \frac{\beta^w_l}{1 - \beta^s_g \beta^c_l} \ln w
\] (7.16)

Note that this Cournot equilibrium is stable only if

\[
\left| \frac{\partial \ln e^C}{\partial \ln s} \right| < \left| \frac{1}{\frac{\partial \ln s^C}{\partial \ln e}} \right|
\] (7.17)

and this inequality holds, looking to Eq. (7.12) and (7.14), only if

\[
|\beta^s_g| < \left| \frac{1}{\beta^c_l} \right|
\] (7.18)

Since \(\beta^s_g\) and \(\beta^c_l\) are, respectively, the elasticity of the government’s reaction function with respect to strike \((E_{eC,s})\) and the elasticity of the labour union’s reaction function with respect to effort \((E_{sC,e})\), Eq. (7.18) is equivalent to the condition

\[
|E_{eC,s}| \cdot |E_{sC,e}| < 1
\] (7.19)

Therefore, in order to reach a stable equilibrium, the players need reaction functions sufficiently inelastic.

Furthermore, the elasticities that follow show that, at the equilibrium, the effect of a raise in unemployment on the level of control variables (effort and strike) is negative, but the effect of a raise in wages on the level of effort and strike is respectively positive and negative. More formally,

\[
E_{e^*,u} = u \frac{\partial \ln e^*}{\partial u} = \frac{\beta^u_g}{1 - \beta^s_g \beta^c_l} < 0
\] (7.20)

\[
E_{s^*,u} = u \frac{\partial \ln s^*}{\partial u} = \frac{\beta^u_l \beta^c_g}{1 - \beta^s_g \beta^c_l} < 0
\] (7.21)

\[
E_{e^*,w} = w \frac{\partial \ln e^*}{\partial w} = \frac{\beta^w_g \beta^s_l}{1 - \beta^s_g \beta^c_l} > 0
\] (7.22)

124
7. Government fiscal efforts vs. labour union strikes: it takes two to tango

\[ E_{s^*, w} = w \frac{\partial \ln s^*}{\partial w} = \frac{\beta^w}{1 - \beta^g \beta^e} < 0 \quad (7.23) \]

The aforementioned equilibrium properties give rise to a set of empirical predictions and testable restrictions that are addressed in detail in the following empirical analysis focused on the case of Italy.

**7.3. Empirical analysis**

In this section, the policy-game is analyzed empirically by the VEC methodology, investigating, in particular, whether the reaction functions in Eq.s (7.12) and (7.14) are supported by Italian data. Then, an interpretation of the estimated model in terms of phase diagrams is provided.

The Italian economy, for its history and institutional background, is well-suited to investigating the equilibrium reaction functions expressed in Eq.s (7.12) and (7.14); see Appendix 7.5 for a thorough discussion. Since the time series that proxy the variables of the theoretical model are nonstationary, estimates are done by a cointegrated VAR model and Johansen’s methodology [Johansen, 1995, Juselius, 2006]. This methodology allows to check whether the data support the existence of at least two cointegrating relationships, consistent with the Cournot-Nash equilibrium predicted by the theoretical model.

The variables of the theoretical model are proxied by quarterly observations relative to the period 1960.Q1-2009.Q4 (detailed descriptions and data sources may be found in Appendix 7.6): \( \ln e \) is proxied by \( \text{effort}_t = \ln \left(1 + \frac{S_t}{GDP_t}\right) \), where \( S_t = -\Delta B_t \) is minus the public deficit, \( B_t \) is the nominal stock of debt, and \( GDP_t \) is the nominal Gross Domestic Product; \( \ln s \) is proxied by \( \text{strike}_t = \ln \left(1 + \frac{H_t}{POP_t}\right) \), where \( H_t \) is the amount of hours not worked due to labour disputes arising from the labour contract and \( POP_t \) is the active population aged 15-64; \( \ln u \) is proxied by \( \text{un}_t = \ln \left(1 + \frac{U_t}{LF_t}\right) \), where \( U_t \) is the official unemployment rate \( (LF_t \) is the labour force and \( U_t \) are the unemployed workers); \( \ln w \) is proxied by \( \text{wage}_t = \ln \left(1 + W_{06}^t\right) \), where \( W_{06}^t \) is the hourly nominal wage index of the manufacturing sector equal to 100 in 2006.Q1.\(^{11}\)

From the definitions above it turns out that \( \text{effort}_t \) can be interpreted as a measure of the budget surplus the government obtains, over the level of GDP, while \( \text{strike}_t \) is a

\(^{11}\)According to the efficiency-wages theory, the monopoly union model and the Phillips curve framework [Layard and Nickell, 1990, Manning, 1991, 1993], the outcome of collective bargaining is the real wage. However, it is the nominal wage that is the main object of the negotiation between the union and the government given the price level (or the inflation rate). In the Italian incomes policy, indeed, the salary has been considered as an ‘independent variable’ for many years [Sraffa, 1960].
measure of the amount of hours not worked per capita. In addition to these four variables, a measure of the inflation rate, \( \pi_t \), is included in the system. Including the inflation rate is crucial because the relationship between wages and unemployment is not explicitly considered in the model, but by using a nominal wage index as a proxy for wages, inflation is allowed to play a key role in explaining the empirical link between wages and unemployment. As a matter of the fact, inflation has always exerted an important role in the Italian policy practice and economic history. Although the literature on the wage curve or wage-setting in Italy is quite limited, the existing empirical evidence on the relationship between inflation, wages, unemployment, and labour productivity is now clear [Chiarini and Piselli, 2003]. The time series of main interest, \( \text{effort}_t \) and \( \text{strike}_t \), are plotted in Figure 7.1.

The five variables of the model are collected in the vector \( X_t=(\text{effort}_t, \text{strike}_t, \text{un}_t, \text{wage}_t, \pi_t)^\prime \) and modeled as a VAR system. After a data-oriented specification search, it was chosen a VAR model with four lags, a constant and three deterministic seasonal dummies. The deterministic seasonal dummies were included to account for residual seasonal patterns in the system. In principle, many episodes and institutional changes might have impacted on the dynamics of the variables included in \( X(t) \).\textsuperscript{12} The pre-

\textsuperscript{12}In 1975.Q1 the labour union requested the general agreement on wage indexation (Contingency allowance); on the 13 March 1979 Italy entered in the European Monetary System and signed an agreement for the maintenance of a fixed exchange parity with respect to the ECU; in 1981.Q3 the separation of the Treasury from the banking system occurred; on the 14 February 1984 the Craxi government decreased the amount of the Contingency allowance by 4 percentage points, and the Contingency allowance was finally abolished on 31 July 1992 by the Amato government; on the 23 July 1993 the Ciampi government adopted the consultation model based on participation in political decisions of trade unions, employers’ organizations, and the government; on the 1 November 1993 the Treaty on European Union came into force in Italy, bringing to the creation of the European Central Bank and the European System of Central Banks on the 1 January 1999.
liminary analyses showed that the VAR-based results on the long run Cournot-Nash equilibrium provided below are however substantially invariant to the inclusion of intervention dummies in the vector $D_t$ associated with several historical episodes.\(^{13}\) Hence, the occurrence of structural breaks is not explicitly accounted for.

The reference VAR model is given by

$$X_t = \sum_{i=1}^{4} A_i X_{t-i} + \mu + \Phi D_t + \epsilon_t \quad t = 1, \ldots, T \quad (7.24)$$

where $A_i$ is a $5 \times 5$ matrix of parameters, $i = 1, 2, 3, 4$, $\mu$ is a $5 \times 1$ vector of constants, $D_t$ is a $3 \times 1$ vector containing three centered deterministic seasonal dummies with associated coefficients in the $5 \times 3$ matrix $\Phi$, and $\epsilon_t$ is a 5-dimensional white noise process with covariance matrix $\Sigma_\epsilon$.

Table 7.1.: Diagnostic tests on the estimated VAR system in Eq. (7.24)

<table>
<thead>
<tr>
<th>Equations:</th>
<th>effort&lt;sub&gt;t&lt;/sub&gt;</th>
<th>strike&lt;sub&gt;t&lt;/sub&gt;</th>
<th>un&lt;sub&gt;t&lt;/sub&gt;</th>
<th>wage&lt;sub&gt;t&lt;/sub&gt;</th>
<th>(\pi_t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LM-AR 1-5 test:</td>
<td>0.33 [0.90]</td>
<td>0.52 [0.76]</td>
<td>2.22 [0.55]</td>
<td>1.46 [0.21]</td>
<td>1.49 [0.19]</td>
</tr>
<tr>
<td>LM-Normality test:</td>
<td>22.30 [0.00]</td>
<td>50.36 [0.00]</td>
<td>12.12 [0.00]</td>
<td>10.72 [0.00]</td>
<td>41.91 [0.00]</td>
</tr>
<tr>
<td>System: LM-AR 1-5 vector test:</td>
<td>1.07 [0.30]</td>
<td>1.07 [0.30]</td>
<td>1.07 [0.30]</td>
<td>1.07 [0.30]</td>
<td>1.07 [0.30]</td>
</tr>
<tr>
<td>System: LM-Normality vector test:</td>
<td>135.00 [0.00]</td>
<td>135.00 [0.00]</td>
<td>135.00 [0.00]</td>
<td>135.00 [0.00]</td>
<td>135.00 [0.00]</td>
</tr>
</tbody>
</table>

Largest eigenvalues of estimated companion matrix

<table>
<thead>
<tr>
<th>Real</th>
<th>Imag.</th>
<th>Modulus</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9987</td>
<td>0.0000</td>
<td>0.9987</td>
</tr>
<tr>
<td>0.9628</td>
<td>0.0335</td>
<td>0.9634</td>
</tr>
<tr>
<td>0.9628</td>
<td>-0.0335</td>
<td>0.9634</td>
</tr>
<tr>
<td>0.0134</td>
<td>0.8532</td>
<td>0.8533</td>
</tr>
<tr>
<td>0.0134</td>
<td>-0.8532</td>
<td>0.8533</td>
</tr>
</tbody>
</table>

The estimation sample is 1960.Q1-2009.Q4, including initial lags.

UPPER PANEL: LM-AR 1-5 is a Lagrange multiplier test for the null of absence of autocorrelation in the VAR disturbances against the alternative of autocorrelations up to lag order 5. LM-Normality is a Lagrange multiplier test for the null of normally distributed disturbances against the alternative of a non-Gaussian distribution. P-values are reported in squared brackets.

LOWER PANEL: largest estimated roots (eigenvalues) of VAR companion matrix.

\(^{13}\) Many robustness checks were done taking into account the government’s ideology (left vs. right party), the electoral system (majoritarian vs. proportional system), the election date, and the union membership. Since these political variables do not result significant, the relevant political-economic environment has been identified [Castañeda, 1995]. Results are available upon request.
System (7.24) was estimated on the period 1960.Q1-2009.Q4 \((T = 196, \text{excluding the four initial lags})\). The upper panel of Table 7.1 reports some residuals diagnostic tests on the estimated model. The diagnostic checks suggest that the specified model captures the dynamics of the five variables fairly well; indeed, albeit the VAR disturbances are not Gaussian, the residuals obtained from the estimated system are consistent with the occurrence of serially uncorrelated disturbances. Therefore, the choice of the variables included in the policy-game model seems to be supported by data.

To obtain indications about the persistence of the modeled time series and the data adequacy of the model, the lower panel of Table 7.1 reports the largest estimated roots (eigenvalues) of the VAR companion matrix. The estimated eigenvalues suggest that the system is likely to be driven by a minimum of one up to a maximum of three common stochastic trends (unit roots) and such evidence is consistent with the implications of the structural reaction functions. Indeed, under the null that the strategic interaction model summarized by Eq.s (7.12) and (7.14) holds true, not more than three unit roots would be expected to be found in the VAR system for \(X_t = (\text{effort}_t, \text{strike}_t, u_t, \text{wage}_t, \pi_t)'\).

Sub-section 7.3.1 investigates the long run equilibrium properties of the VAR system (7.24) by using cointegration techniques. The objective is to assess whether the structural reaction functions in Eq.s (7.12) and (7.14) are supported by the data in terms of cointegration analysis. Then, an interpretation of the estimated model in terms of phase diagrams is provided in Sub-section 7.3.2.

### 7.3.1. Equilibrium

When the VAR system (7.24) is driven by unit roots and the variables are cointegrated, it is useful to consider its VEC counterpart [Johansen, 1995]:

\[
\Delta X_t = \alpha(\beta', \mu_c) \left( \begin{array}{c} X_{t-1} \\ 1 \end{array} \right) + \sum_{j=1}^{3} \Gamma_j \Delta X_{t-j} + \mu_u + \Phi D_t + \varepsilon_t \quad t = 1, ..., T \tag{7.25}
\]

where \(\Delta X_t = X_t - X_{t-1}\) is the difference operator, \(\Gamma_j, j = 1, 2, 3\), are \(5 \times 5\) matrices which depend on the original VAR parameters, \(\alpha\) and \(\beta\) are matrices of dimension \(5 \times r\) of full-column rank \(r < 5\). In Eq. (7.25), the constant \(\mu\) has been split into two parts: the component \(\mu_c = \alpha_\perp (\alpha_\perp' \alpha_\perp)^{-1} \alpha_\perp' \mu\) representing the portion of \(\mu\) entering the cointegrating relationships, where \(\alpha_\perp\) is a basis of the orthogonal complement of \(\alpha\), and the component \(\mu_u = \alpha (\alpha' \alpha)^{-1} \alpha' \mu\) representing the portion of \(\mu\) entering the adjustment equations of the VEC system. Under a set of suitable identifying restrictions on \(\beta\) of the type \(\beta = \beta I\), the \(r\)-dimensional vector \(\beta' I X_{t-1}\) captures the steady state.
relationships embedded in the system, while the coefficients in the matrix $\alpha$ capture the short run (next-period) adjustment of the variables in $\Delta X_t$ to equilibrium. The strategic interaction game predicts that there should be at least two cointegrating relationships between the variables in $X_t$ consistent with the structural equilibrium reaction functions derived in Sub-section 7.2.1.

The departure of the error terms from the Gaussian distribution documented in Table 7.1 does not preclude the use of Johansen’s maximum likelihood (ML) method for the determination of the cointegration rank $r$ and the estimation of the parameters $\beta=\beta_I$ and $\alpha$, $\Gamma_1$, $\Gamma_2$, $\Gamma_3$, $\mu$, $\Phi$ and $\Sigma$. Indeed, the ML estimation of Eq. (7.25) based on the Gaussian likelihood can be considered robust to deviations from normality under the null of cointegration, see Gonzalo [1994].

Table 7.2.: LR Trace cointegration rank test

<table>
<thead>
<tr>
<th>$H_0 : r \leq j$</th>
<th>Trace</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$j = 0$</td>
<td>112.500</td>
<td>0.000</td>
</tr>
<tr>
<td>$j = 1$</td>
<td>64.884</td>
<td>0.000</td>
</tr>
<tr>
<td>$j = 2$</td>
<td>38.695</td>
<td>0.003</td>
</tr>
<tr>
<td>$j = 3$</td>
<td>14.558</td>
<td>0.068</td>
</tr>
<tr>
<td>$j = 4$</td>
<td>0.071</td>
<td>0.790</td>
</tr>
</tbody>
</table>

The estimation sample is 1960.Q1-2009.Q4, including initial lags. The cointegration rank test is conducted in the context of the VAR system in Eq. (7.24) considering the case of unrestricted constant $\mu$.

Table 7.2 summarizes the results of the LR Trace test [Johansen, 1995] for the determination of the cointegration rank $r$. The test is carried out by considering an unrestricted specification for the constant and treating the system as integrated at most of order one (I(1)). It can be noticed that the test selects $r = 3$ cointegrating relationships at the 5% level of significance (corresponding to two stochastic trends, or unit roots), albeit marginally.\textsuperscript{14} The subsequent empirical analysis is thus conditioned on the choice $r = 3$.

With $r = 3$, the quantity $\beta'X_{t-1}$ in Eq. (7.25) is a 3-dimensional vector and the parameters in $\beta$ need to be identified, i.e. $\beta$ must fulfill the restriction $\beta = \beta_I$, where $\beta_I$ matches the conditions discussed in Johansen [1995]. In addition, $\mu_c = -1 \cdot (\beta^0_y, \beta^0_I, \beta^0_x)'$ is the vector of the restricted constants. This entails identifying an ‘extra’ long run relationship other than the equilibrium reaction functions implied by Eqs. (7.12) and (7.14).

\textsuperscript{14}In principle, there would be room for the existence of a fourth cointegrating relationship, i.e. $r = 4$, corresponding to a single stochastic trend or unit root driving the system. It would be difficult, however, to identify two ‘additional’ equilibria other the two predicted by the strategic policy model.
7. Government fiscal efforts vs. labour union strikes: it takes two to tango

The first column of the matrix \( \beta_I, \beta_{I,1} \), identifies the beta coefficients of the long run government’s reaction function, i.e. the quantity \( v_{e,t-1} = \beta'_{I,1}X_{t-1} - \beta^0_g = \text{effort}_{t-1} - \beta^0_g - \beta^n_g w_{t-1} - \beta^s_g \text{strike}_{t-1} \) captures the deviation of the observed effort, \( \text{effort}_{t-1} \), from its long run equilibrium level as implied by Eq. (7.12). The second column of the matrix \( \beta_I, \beta_{I,2} \), identifies the beta coefficients of the long run labour union’s reaction function, i.e. \( v_{s,t-1} = \beta'_{I,2}X_{t-1} - \beta^0_s = \text{strike}_{t-1} - \beta^0_s - \beta^n_s w_{t-1} - \beta^s_s \text{effort}_{t-1} \) captures the deviation of the observed strike, \( \text{strike}_{t-1} \), from its long run equilibrium level as implied by Eq. (7.14). Finally, the third column of the matrix \( \beta_I, \beta_{I,3} \), identifies the coefficients of the Phillips curve-type statistical relationship, i.e. \( v_{\pi,t-1} = \beta'_{I,3}X_{t-1} - \beta^0_\pi = \pi_{t-1} - \beta^0_\pi - \beta^n_\pi w_{t-1} - \beta^s_\pi \text{wage}_{t-1} \). Overall, the simultaneous system of equations given by

\[
(\beta'_{I}, \mu_{e}) \begin{pmatrix} X_{t-1} \\ 1 \end{pmatrix} = v_{t-1}
\]

is exactly identified, where \( v_{t-1} = (v_{e,t-1}, v_{s,t-1}, v_{\pi,t-1})' \).

The estimated identified long run relationships are given by

\[
(\hat{\beta}'_{I}, \hat{\mu}_{e}) \begin{pmatrix} X_{t-1} \\ 1 \end{pmatrix} = \begin{cases} \hat{\beta}'_{I,1}X_{t-1} - \hat{\beta}^0_g & \\ \hat{\beta}'_{I,2}X_{t-1} - \hat{\beta}^0_s & \\ \hat{\beta}'_{I,3}X_{t-1} - \hat{\beta}^0_\pi \end{cases} = \begin{cases} \text{effort}_{t-1} + 0.183 \text{strike}_{t-1} + 2.180 w_{t-1} - 0.117 & (0.045) \\ \text{strike}_{t-1} - 2.260 \text{effort}_{t-1} + 0.972 \text{wage}_{t-1} - 0.396 & (0.731) \\ \pi_{t-1} + 0.089 w_{t-1} + 0.027 \text{wage}_{t-1} - 0.016 & (0.089) \end{cases}
\]

(7.27)

and are also reported in Table 7.3 along with the adjustment coefficients \( \alpha \). The estimation of \( \alpha \) and \( \beta_I \) was carried out simultaneously by ML. The LR test reported in the bottom part of Table 7.3 is a joint test for the zero restrictions characterizing the estimated \( \alpha \) matrix and can be regarded as a statistical test for the overall specification. In particular, the estimated long run relationships support the reduced form of the policy-game model and the orthogonality assumption underlying the interaction between the government and the labour union.

The first two long run relationships are the estimated government’s and labour union’s reaction functions. The \( \hat{\beta}_I \) coefficients represent the slopes of the reaction functions: for a strike (amount of hours not worked) raise of 1 hour per quarter per capita the best responsive effort (government budget surplus) decreases by 0.183% in terms of GDP; for an effort raise of 1% in terms of GDP the best responsive strike increases by 2.260 hours.
7. Government fiscal efforts vs. labour union strikes: it takes two to tango

Table 7.3.: Estimated cointegration relationships and adjustment coefficients

<table>
<thead>
<tr>
<th>Estimated $\beta_I$ for fixed cointegration rank $r = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\hat{\beta}_I)^\prime \hat{\mu}_c$ (0.731)</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>0.183</td>
</tr>
<tr>
<td>2.180</td>
</tr>
<tr>
<td>(0.529)</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>(0.118)</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>0.180</td>
</tr>
<tr>
<td>(0.529)</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>-0.117</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Estimated adjustment coefficients $\alpha$ for $r = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \text{effort}_t$ &amp; $\Delta \text{strike}_t$ &amp; $\Delta \text{un}_t$ &amp; $\Delta \text{wage}_t$ &amp; $\Delta \pi_t$</td>
</tr>
<tr>
<td>$v_{e,t-1}$ &amp; $v_{s,t-1}$ &amp; $v_{\pi,t-1}$ &amp; $v_{\pi,t-1}$ &amp; $v_{\pi,t-1}$</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>-0.822</td>
</tr>
<tr>
<td>-0.016</td>
</tr>
<tr>
<td>(0.004)</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>(0.020)</td>
</tr>
<tr>
<td>-0.030</td>
</tr>
<tr>
<td>(0.067)</td>
</tr>
</tbody>
</table>

LR test of restrictions $\chi^2(7) = 4.791$ [0.686]

The estimation sample is 1960.Q1-2009.Q4, including initial lags.
Upper panel: estimated cointegrating relationships.
Lower panel: estimated significant adjustment coefficients. The LR test is a likelihood ratio test for the zero restrictions in the $\alpha$ matrix. Standard errors in parentheses, p-value in squared brackets.

per quarter per capita.

The third long run relationship is the estimated Phillips curve-type in which the inflation rate, the unemployment rate and the wage index are linked together. Note that, in this relationship in which nominal wages are adjusted for inflation, there is not a strong (quite significant) link between the inflation and the unemployment rate. Furthermore, the estimated coefficients imply a negative relationship between inflation and wages. A possible explanation for this finding is the following. Tax incentives for saving, investment, and work, human capital investment, deregulation, trade liberalization, infrastructure development all involve a rightward shift of the aggregate supply (AS) curve. If these positive shocks raise labour productivity, wages can increase when inflation decreases, causing the Phillips curve to shift.
Table 7.4.: Cournot-Nash equilibrium properties

<table>
<thead>
<tr>
<th>Theoretical</th>
<th>Estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>\beta_g^s</td>
</tr>
</tbody>
</table>


Table 7.4 summarizes the estimates of the strategic interaction model’s elasticities derived in Eq.s (7.20)-(7.23). First of all, by replacing the theoretical parameters with their super-consistent ML punctual estimates, the stability condition for the equilibrium is verified. Furthermore, the estimated equilibrium elasticities are in line with the predictions of the theoretical model. The effect of unemployment and wages is mainly negative on the equilibrium level of the control variables of the government and the labour union. The only positive (small in magnitude) estimated elasticity implies that, along with the historical raise in wages, the long run effort has mildly increased. Regarding the magnitude of the elasticities, changes in unemployment have stronger effects than changes in wages, and strikes are affected more than effort at the long run equilibrium. In order to give a possible interpretation of the effort elasticities, one can assume that (i) an increase in structural (long-term) unemployment corresponds to a leftward shift of the AS curve, and (ii) an increase in long-term wages, associated with a raise in productivity, corresponds to a rightward shift of the AS curve. As concerns (i), the higher the unemployment the lower the GDP. Given the estimated negative elasticity $\hat{E}_{e^*,u}$ and knowing that effort is roughly equal to government budget surplus over GDP, the decrease in GDP has to be less than proportional to the decrease in the surplus. As concerns (ii), the lower the wages the lower the GDP. Given the estimated positive elasticity $\hat{E}_{e^*,w}$, the decrease in GDP has to be less than proportional to the decrease in
the surplus. Both these results are in line with a Keynesian fiscal policy implemented in Italy in the last half century. The estimated strike elasticities testify the historical long run decrease in strikes and raise in unemployment and wages.

Figure 7.2 plots the cointegration residuals $\hat{v}_{e,t}$, $\hat{v}_{s,t}$ and $\hat{v}_{\pi,t}$ implied by Eq. (7.27). It can be noticed that the two error correction terms of interest, $\hat{v}_{e,t}$ and $\hat{v}_{s,t}$, are surprisingly stable over the investigated period, while the error correction term associated with the statistical Phillips-type relationships reflects the historical pattern of Italian inflation stabilization, which was undertaken since 1986.

**Adjustment dynamics**

The parameters in the matrix $\alpha$ of the VEC in Eq. (7.25) are short run adjustment coefficients and hence capture the next-period (next-quarter in this case) responses of the variables in $\Delta X_t$ to lagged disequilibria $v_{t-1}$ (i.e. to lagged deviations of the observed variables from their cointegration levels).

The estimated $\alpha$s reported in the lower panel of Table 7.3 suggest that the government does not adjust in the short run to any of the identified long run relationship, while the labour union adjusts significantly to the error correction terms associated with the two estimated reaction functions, in addition to the error correction term involving the statistical Phillips curve-type relationship. More precisely, there is no error-correction mechanism at work in the $\Delta \text{effort}_t$-equation, meaning that $\text{effort}_t$ is ‘weakly exogenous’ with respect to the identified $\beta_I$ [Hendry, 1995]. These results indicate, with some qualifications, that $\text{effort}_t$ acts as the stochastic (common) trend driving the sub-system formed by $\text{effort}_t$ and $\text{strike}_t$, while the labour union acts as the ‘buffer’ of the adjustment process, in the sense that $\text{strike}_t$ is the variable in charge of correcting, quarter-by-quarter, the dynamic path of adjustment such that deviations from the two cointegrating reaction functions do not drift apart too much. It turns out that the labour union’s strike policy is influenced by the government’s effort not only in the long run, as suggested by
7. Government fiscal efforts vs. labour union strikes: it takes two to tango

Eq. (7.14) and its empirical counterpart in Eq. (7.27), but over shorter horizons as well. Note that, strike policy is also affected by the deviation of inflation from its long run equilibrium level (third long run relationship), while the labour union negotiates nominal wages takes inflation into account. Finally, it is worth observing that the acceleration rate, \( \Delta \pi_t \), adjusts not only to \( \hat{\nu}_{t-1} \), as expected, but to both the estimated disequilibria \( \hat{\nu}_{e,t-1} \) and \( \hat{\nu}_{s,t-1} \) as well, suggesting that beyond business cycle and monetary policy developments, the strategic interaction between the government and the labour union plays a role in explaining Italian inflation dynamics over the last fifty years.

7.3.2. Phase diagram representation

This sub-section attempts to improve the understanding of the estimated strategic interaction game by providing an interpretation of its long run and adjustment features by means of phase diagrams.

Eqs (7.28) and (7.29) below correspond to the equations of the VEC system (7.25) associated with the government’s control variable (effort) and the labour union’s control variable (strike), respectively, i.e.

\[
\Delta \text{effort}_t = \alpha_g^e (\text{effort}_{t-1} - \ln e^\text{C}_{t-1}) + \alpha_l^e (\text{strike}_{t-1} - \ln s^\text{C}_{t-1}) + f_e (v_{\pi,t-1}; \Delta X_{t-1}, \Delta X_{t-2}, \Delta X_{t-3}; \Gamma_1, \Gamma_2, \Gamma_3) + \varepsilon_t^{\text{effort}} \tag{7.28}
\]

\[
\Delta \text{strike}_t = \alpha_g^s (\text{effort}_{t-1} - \ln e^\text{C}_{t-1}) + \alpha_l^s (\text{strike}_{t-1} - \ln s^\text{C}_{t-1}) + f_s (v_{\pi,t-1}; \Delta X_{t-1}, \Delta X_{t-2}, \Delta X_{t-3}; \Gamma_1, \Gamma_2, \Gamma_3) + \varepsilon_t^{\text{strike}} \tag{7.29}
\]

Both these two equations emphasize the two error correction terms of interest \( (v_{\pi,t-1} = \text{effort}_{t-1} - \ln e^\text{C}_{t-1}) \) (\( v_{s,t-1} = \text{strike}_{t-1} - \ln s^\text{C}_{t-1} \)). All remaining components not directly related to the adjustment coefficients \( \alpha_g^e, \alpha_l^e, \alpha_g^s, \) and \( \alpha_l^s \) have been confined to the terms \( f_e(.) \) and \( f_s(.) \), respectively, and the equation-specific disturbances are \( \varepsilon_t^{\text{effort}} \) and \( \varepsilon_t^{\text{strike}} \).

To isolate the strategic interaction between the government and the labour union, it is possible to partial out \( f_e(.) \), \( f_s(.) \), \( \varepsilon_t^{\text{effort}} \), and \( \varepsilon_t^{\text{strike}} \), putting all these terms equal to zero in Eqs (7.28) and (7.29). Furthermore, by the conditions \( \Delta \text{effort}_t = 0 \) and \( \Delta \text{strike}_t = 0 \), the loci in which the short run adjustments, respectively for \( \text{effort}_{t-1} \) and \( \text{strike}_{t-1} \), are null are obtained. By solving Eqs (7.28) and (7.29) for \( \text{effort}_{t-1} \) under the above conditions:

\[
\text{effort}_{t-1} = \frac{\alpha_g^e (\beta_g^0 + \beta_g^u u_{t-1} + \beta_g^s \text{strike}_{t-1}) - \alpha_l^e (\text{strike}_{t-1} + \beta_l^0 + \beta_l^w \text{wage}_{t-1})}{\alpha_g^e + \alpha_l^e \beta_l^e} \tag{7.30}
\]
7. Government fiscal efforts vs. labour union strikes: it takes two to tango

\[
\text{effort}_{t-1} = \frac{\alpha_g^2 (\beta_g^0 + \beta_y^u un_{t-1} + \beta_y^s strike_{t-1}) - \alpha_s^l (\text{strike}_{t-1} - \beta_y^0 + \beta_y^w \text{wage}_{t-1})}{\alpha_y^2 + \alpha_s^l \beta_e^l}
\]

(7.31)

which correspond to the dashed lines plotted in the phase diagram plotted in Figure 7.3.

The phase diagram in Figure 7.3 represents the cointegrating relationships in a unified way (the player’s reaction functions) and the loci in which the short run adjustments are null (the strategic interaction between players being isolated). The contribution of this phase diagram is to highlight the interdependencies among the steady state equilibrium (Cournot-Nash equilibrium), the long-term relationships (reaction functions), and the short run and long run dynamics (that depend on the parameters in \(\alpha\) and \(\beta_I\)). Notice that, since the VEC is stable by construction, wherever one starts from the graph, the dynamics of the system is such that it pushes directly towards the steady state. This point, identified by the intersection of \(\ln s_{t-1}^C\), \(\ln e_{t-1}^C\), \(\Delta\text{effort}_t = 0\), and \(\Delta\text{strike}_t = 0\), is both the long run equilibrium and the Cournot-Nash equilibrium.

The phase diagram in Figure 7.3 specializes under some additional hypotheses about the behaviour of the government and the labour union, once the estimated parameters of the model are used. In particular, the phase diagram in Figure 7.4 is obtained by setting \(\alpha_g^2, \alpha_e^l, \alpha_s^l, \beta_y^u, \beta_y^s, \beta_y^w, \) and \(\beta_e^l\) to their ML estimates presented in Table 7.3, by fixing \(un_{t-1}\) and \(\text{wage}_{t-1}\) at their sample averages. The constants \(\beta_g^0\) and \(\beta_y^0\), instead, have been determined ‘residually’, consistently with the following hypotheses:

(i) since industrial conflict is costly, in the long run agents are assumed to be rational and well-informed such that, following Hicks [1932], labour unions will not use strikes.

\[\text{strike}_t \leq 0\]

\[\text{effort}_t \leq 0\]

\[\ln s_{t-1}^C \leq 0\]

\[\ln e_{t-1}^C \leq 0\]

\[\Delta\text{effort}_t = 0\]

\[\Delta\text{strike}_t = 0\]

\[\text{effort}_{t-1} \leq 0\]

\[\text{strike}_{t-1} \leq 0\]

\[\ln s_{t-1}^C \leq 0\]

\[\ln e_{t-1}^C \leq 0\]

\[\Delta\text{effort}_t = 0\]

\[\Delta\text{strike}_t = 0\]

\[\text{effort}_{t-1} \leq 0\]

\[\text{strike}_{t-1} \leq 0\]

\[\ln s_{t-1}^C \leq 0\]

\[\ln e_{t-1}^C \leq 0\]

\[\Delta\text{effort}_t = 0\]

\[\Delta\text{strike}_t = 0\]

\[\text{effort}_{t-1} \leq 0\]

\[\text{strike}_{t-1} \leq 0\]

\[\ln s_{t-1}^C \leq 0\]

\[\ln e_{t-1}^C \leq 0\]

\[\Delta\text{effort}_t = 0\]

\[\Delta\text{strike}_t = 0\]

\[\text{effort}_{t-1} \leq 0\]

\[\text{strike}_{t-1} \leq 0\]

\[\ln s_{t-1}^C \leq 0\]

\[\ln e_{t-1}^C \leq 0\]

\[\Delta\text{effort}_t = 0\]

\[\Delta\text{strike}_t = 0\]

\[\text{effort}_{t-1} \leq 0\]

\[\text{strike}_{t-1} \leq 0\]

\[\ln s_{t-1}^C \leq 0\]

\[\ln e_{t-1}^C \leq 0\]

\[\Delta\text{effort}_t = 0\]

\[\Delta\text{strike}_t = 0\]

\[\text{effort}_{t-1} \leq 0\]

\[\text{strike}_{t-1} \leq 0\]

\[\ln s_{t-1}^C \leq 0\]

\[\ln e_{t-1}^C \leq 0\]

\[\Delta\text{effort}_t = 0\]

\[\Delta\text{strike}_t = 0\]

\[\text{effort}_{t-1} \leq 0\]

\[\text{strike}_{t-1} \leq 0\]

\[\ln s_{t-1}^C \leq 0\]

\[\ln e_{t-1}^C \leq 0\]

\[\Delta\text{effort}_t = 0\]

\[\Delta\text{strike}_t = 0\]

\[\text{effort}_{t-1} \leq 0\]

\[\text{strike}_{t-1} \leq 0\]

\[\ln s_{t-1}^C \leq 0\]

\[\ln e_{t-1}^C \leq 0\]

\[\Delta\text{effort}_t = 0\]

\[\Delta\text{strike}_t = 0\]

\[\text{effort}_{t-1} \leq 0\]

\[\text{strike}_{t-1} \leq 0\]

\[\ln s_{t-1}^C \leq 0\]

\[\ln e_{t-1}^C \leq 0\]

\[\Delta\text{effort}_t = 0\]

\[\Delta\text{strike}_t = 0\]

\[\text{effort}_{t-1} \leq 0\]

\[\text{strike}_{t-1} \leq 0\]

\[\ln s_{t-1}^C \leq 0\]

\[\ln e_{t-1}^C \leq 0\]

\[\Delta\text{effort}_t = 0\]

\[\Delta\text{strike}_t = 0\]
7. Government fiscal efforts vs. labour union strikes: it takes two to tango

Figure 7.4.: Estimated phase diagram for Italy under hypotheses (i) and (ii) in the text as instruments of conflict resolution; (ii) since the Maastricht Treaty and the Stability and Growth Pact are effective in Italy, the Italian governments respects the institutional constraints on fiscal policy in the long run, in particular a value of the deficit-to-GDP ratio equal to 3%. Based on (i) and (ii), Figure 7.4 plots the ‘ideal’ long run equilibrium for the Italian economy obtained by setting $\text{strike}_{t-1}$ at 0 and $\text{effort}_{t-1}$ at -0.03.

As a conceptual experiment, one can isolate the dynamics driven only by the deviations from the equilibrium reaction functions by putting $f_e(\cdot), f_s(\cdot), \varepsilon^effort_t, \text{and } \varepsilon^strike_t$ equal to zero in Eq.s (7.28) and (7.29). Figure 7.4 plots the adjustment processes that emerge starting from the points A, B, and C. Since $\alpha^g$ and $\alpha^l$ are not significant (see Table 7.3), the initial adjustment is driven by strike. Once attained the locus of $\Delta \text{strike}_t = 0$, there will be a slow convergence to the long run equilibrium. Obviously, the farther a point from the locus, the stronger the push towards it. Finally, since strikes cannot be negative, the adjustment from the point A is bounded by the vertical axis.

7.4. Conclusions

When unions can use strikes to maximize wages and consensus among workers, and governments use fiscal policy to maximize employment and consensus among voters, the interaction between their strategies is crucial to understanding economic and political performances of unionized economies. To stylize the simultaneous interaction between the government and the labour union, and to identify the players’ long-term economic relationships and the steady state equilibrium, this work developed a policy-game model
7. Government fiscal efforts vs. labour union strikes: it takes two to tango

with perfect information. Strike is strategic substitute from government perspective and effort is strategic complement from labour union perspective. The stability of the Cournot-Nash equilibrium implied by the game depends on the elasticity of the government’s reaction function and of the labour union’s reaction function. At equilibrium, furthermore, the effect of a raise in unemployment on the level of effort and strike is negative, while the effect of a raise in wages on the level of effort and strike is respectively positive and negative.

To estimate and test the policy-game model, Italian quarterly data from the period 1960-2009 and the VEC methodology were used. Measuring fiscal efforts with the government budget surplus and strikes with the amount of hours not worked due to labour disputes arising from the labour contract, the empirical results show that the predicted Cournot-Nash equilibrium is convincingly supported by the data. In particular, the data support the existence of at least two cointegrating relationships (the player’s reaction functions) consistent with the Cournot-Nash equilibrium predicted by the theoretical model. Interestingly, the estimated structure of dynamic adjustment towards the equilibrium reveals that in the short run effort does not adjust to any of the empirically detected economic relationships, suggesting that the government’s policy corresponds to the ‘driving force’ (common stochastic trend) characterizing the system. On the other hand, since strike adjusts to all three estimated equilibrium relationships, it is the labour union which uses the strike policy to restore the equilibrium over time.

The proposed model can be useful to understanding unionized economies like European countries in depth, where the strategic interaction between the government and the labour union is relevant to determining strikes, government balance, unemployment, wages, and inflation. However, this model can be applied to other unionized economies, with or without the same fiscal parameters of stability. Ultimately, the result of this work is that the effect of fiscal policies and strikes on unemployment, wages, and inflation crucially depends on whether the players are able to reach the long run equilibrium, if a stable equilibrium in the policy-game exists, and if the players’ reaction functions are both inelastic or at least one of the two reaction functions is very inelastic.

7.5. Appendix: Italian institutional configuration

Italian industrial relationships have been marked by important political participation of the three biggest labour union confederations (CGIL, CISL and UIL) that cover almost all employees. In 2009, total CGIL members (including retirees) were 5,656,155, while analogous figure for CISL and UIL were respectively 4,531,040 and 2,164,276. There is
7. Government fiscal efforts vs. labour union strikes: it takes two to tango

a fourth confederation (UGL) but official data are not available for its members. These big confederations are differentiated according to their political ideology: CGIL adheres to communist ideology, CISL to catholic ideals and UIL to socialist ideology. From the political point of view and until the early 90s, CGIL was contiguous to the Communist Party (PCI), CISL to the Catholic Party (DC), and UIL to the Socialist Party (PSI). DC ruled for all terms from 1948 until 1994. In this period, PCI was always in opposition and PSI was in office with DC for several terms. In 1994, the electoral system turned from proportional representation to majority system and led to the disappearance of the traditional parties (PCI, PSI and DC) and the emergence of new Right (PDL) and Left parties (PD). The new Right party was in government in 1994 and ruled from 2001 to 2006 and from 2008 to 2011. The new Left party governed from 1996 to 2001 and from 2006 to 2008. Despite their ideological divisions, when the labour unions negotiated economic and social policy issues with the governments their differences were vague and confederations often cooperated with each other. Italian labour unions (CGIL, CISL and UIL) and Leftist parties (PCI and PSI) have historical similar political positions as both pursue the interests of workers. However, while labour unions have fought for higher wages, left-wing parties have mainly sought to reduce the unemployment rate. The model of ‘consultation’ as policy of income was applied in Italy in the early ’90s to reduce the inflation rate and, indirectly, the interest rates. Thus, labour unions have been able to negotiate directly with governments using strikes as weapon of negotiation (political strikes), while left-wing parties have been able to reduce the unemployment rate affecting fiscal policy (political cooperation or ‘consociationalism’).

7.6. Appendix: Data

The empirical analysis is based on a hand-collected dataset comprising quarterly macroeconomic time series relative to the Italian economy. The data cover the period 1960.Q1-2009.Q4. Some descriptive statistics are presented in Table 7.5. The five variables in Table 7.5 are obtained as follows (subscript \( t \) denotes quarter):

- \( \ln e \equiv \text{effort}_t = \ln \left(1 + \frac{S_t}{GDP_t}\right) \), where \( S_t = -\Delta B_t \) is minus the deficit (source: Bank of Italy) and \( GDP_t \) is the seasonally adjusted Gross Domestic Product (source: OECD);
- \( \ln s \equiv \text{strike}_t = \ln \left(1 + \frac{H_t}{POP_t}\right) \), where \( H_t \) is the amount of hours not worked due to labour disputes arising from the labour contract (source: Italian National Institute of Statistics) and \( POP_t \) is the active population aged 15-64 (source: OECD);
7. Government fiscal efforts vs. labour union strikes: it takes two to tango

- \( \ln u \equiv \ln \left( 1 + \frac{U_t}{LF_t} \right) \), where \( U_t \) are the unemployed workers, \( LF_t \) is the labour force, and \( \frac{U_t}{LF_t} \) is the seasonally adjusted unemployment rate (source: OECD);

- \( \ln w \equiv \ln \left( 1 + W_{t06}^6 \right) \), where \( W_{t06}^6 \) is the hourly wage index of the manufacturing sector relative to its basis equal to 100 in the quarter 2006.Q1 (source: Italian National Institute of Statistics);

- \( \pi_t = \ln \left( 1 + \Delta P_t \right) \), where \( P_t \) is the Harmonised Index of Consumer Prices (source: Italian National Institute of Statistics).

<table>
<thead>
<tr>
<th>Table 7.5.: Descriptive statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>effort</td>
</tr>
<tr>
<td>strike</td>
</tr>
<tr>
<td>un</td>
</tr>
<tr>
<td>wage</td>
</tr>
<tr>
<td>( \pi )</td>
</tr>
</tbody>
</table>

Number of observations: 200
Bibliography


Bibliography


Bibliography


Bibliography


Bibliography


Bibliography


Bibliography


Max Weber. The Protestant Ethic and the Spirit of Capitalism: Translated from the German by Talcott Parsons. 1930.


Bibliography

Klaus Zimmermann, Thomas K Bauer, Ralph Rotte, and Andreas Million. Immigrant labor and workplace safety. 1999.