Theoretical and Microeconometric Perspectives on Insurance and International Economics

Inaugural-Dissertation

zur Erlangung des Grades

Doctor oeconomiae publicae (Dr. oec. publ.)

an der Ludwig-Maximilians-Universität München

2009

vorgelegt von

Iris Kesternich

Referentin:	Prof. Dr. Monika Schnitzer
Korreferent:	Prof. Dr. Florian Heiss
Promotionsabschlussberatung:	19. Mai 2010

To the memory of my father

Acknowledgements

First and foremost I wish to thank my supervisor, Monika Schnitzer. Her advice, guidance and support have been very important both for the completion of my thesis and in my decision to pursue a scientific career. I have learnt and benefitted in many distinct ways through our joint work. Further, I would like to thank Joachim Winter who has shaped my idea of what applied economics is. His advice and support have been very valuable. Both Monika Schnitzer and Joachim Winter are role models in their dedication to research, teaching and mentoring. I am also grateful to Florian Heiss and Jonas Schreyögg, who offered their support and time and kindly agreed to join my thesis committee. Florian Heiss has supported my research with many ideas and constructive comments.

I am also deeply grateful to my coauthors Claudia Buch, Alexander Lipponer and Heiner Schumacher. Working jointly with them has been and continues to be a great source of mutual support, inspiration and motivation.

I also would like to thank my current and former colleagues at the Seminar of Comparative Economics, the Seminar of Empirical Economics and the Munich Graduate School of Economics. It has been a pleasure working with you. I would like to mention explicitly Anna Gumpert, Karolina Kaiser, Stefan Vetter and Martin Watzinger. Special thanks goes to the team of student assistants at the Seminar of Comparative Economics for their excellent research assistance.

Part of this thesis has been written during visits to the Research Centre of the Deutsche Bundesbank. The hospitality of the institution as well as the possibility to work with the Foreign Direct Investment Micro-Database (MiDi) are gratefully acknowledged. Throughout my thesis I have benefitted from through financial support through the German Science Foundation under GRK 801 and SFB-TR 15.

Finally, my greatest gratitude goes to my family and friends for their love and support.

Munich, 27th of September 2010

Iris Kesternich

Contents

Chapter 1

Introduction

"Raw facts are, as such, a meaningless jumble."
(Joseph Schumpeter, Business Cycles, Vol. I)

Ever since John Maynard Keynes labeled econometrics as "statistical alchemy" rather than a science (Keynes (1940)), the theory versus empirics debate in economics has been ongoing. More recently, this debate has been taken up again in the field of development economics, lanced by the article "Is there too little theory in development economics?" by Mookherjee (2005). The debate started in view of the multitude of empirical papers whose relation to economic theory has remained at best unclear. During the last few years, the role of empirics in economics has been growing, partially due to the increased possibilities of storing and handling huge data sets. This is reflected for example by the increase in (partially) empirical papers published in top journals, by the type of work that Ph.D. students from top graduate schools do or by the number of Nobel Prizes that have been awarded to empirical researchers in the last decade (Eichengreen (2009)).

However, as unsatisfactory an empirical study without theoretical background may be considered by many economists, theory that does not explain or that even contradicts empirical facts is essentially questionable. Econometrics plays an important role in testing hypotheses derived from economic theory, but even more in helping theoretical researchers to find out which questions to ask (Banerjee (2005)). Empirical findings have often challenged existing theory and they have been a starting point for theoretical analyses. Even the critics of econometrics will not deny that the history of science is full of examples where empirical findings have preceded and encouraged important theoretical breakthroughs.

One important motivation for this thesis has been that economic research could greatly benefit from a close cooperation between theorists and econometricians. To say it in Koopmans' words: "Fuller utilization of the concepts and hypotheses of economic theory [...] as part of the process of observation and measurement promises to be a shorter road, perhaps even the only possible road [...]". A simple truth that is still sometimes forgotten is that "one needs a clear mathematical formulation of the relevant economic theory before any statistical analysis with economic data can be done" (Maddala (2001) [The order of words was changed by the author]).

This dissertation consists of four essays from the areas of insurance and international economics. Their common aim is to create a close link between theory and empirical analysis. Empirical facts that have called for an extension of existing theory have always been the starting point for my analyses.

The second link between the essays in this dissertation is the use of micro data. The past two decades have seen an explosion in the availability of micro data for both firms and households, mainly due to the increased capacity of electronical storage and a quantum leap in handling and analyzing large data volumes. This has come along with a shift in empirical analysis from aggregate models describing markets as a whole to the individual decision maker. This seems a more natural unit of analysis, as economic decisions are taken at the individual level and drawing conclusions from aggregate data has to rely on more (and often unrealistic) assumptions.[1]

In fact, Chapter 2 of my dissertation, the essay "Who is afraid of political risk?" was inspired by two empirical papers by Desai, Foley and Hines (Desai *et al*. (2004) and Desai *et al*. (2008)) who investigate how multinational firms choose the capital structure of their foreign affiliates in response to political risk. They claim that in countries where political risk is high, multinational enterprises hold a smaller equity share in their foreign affiliates while at the same time using relatively more debt for financing their affiliates. When we thought of how to theoretically model this relationship, this prediction seemed less compelling, especially with regard to the predicted impact of political risk on affiliate leverage. If political risk increases the risk of expropriation, then this leads to an increased risk of bankruptcy, and firms should use less debt financing in countries where political risk is high. We then started to think about whether there might be different types of political risk such as expropriation, corruption and confiscatory taxation, and how we could analyze these different types both empirically and theoretically. In our theoretical analysis we find that, as political risk increases, the ownership share tends to decrease, because political risk decreases expected profits, whereas leverage can both increase or decrease, depending on the type of political risk. Only when political risk takes the form of discriminatory or confiscatory taxation, it will be optimal to finance affiliates with more debt.

[1] Train *et al*. (1989), chapter 1 and Cameron and Trivedi (2005), chapter 1.

For our empirical analysis, we use the Microdatabase Direct Investment (MiDi) of the Deutsche Bundesbank, and we find supportive evidence for these effects. We follow two empirical strategies in order to show that the effects of political risk may differ for different scenarios of political risk. First, we employ various empirical measures, reflecting different types of political risk, and identify them with different scenarios. Following this strategy, we can confirm that the ownership share decreases with political risk no matter what type of measure we introduce, while leverage can either increase or decrease in response to different types of risk. The second strategy is to rank the scenarios according to their severity and to introduce a quadratic function of political risk which allows the marginal effects of political risk to vary with the severity of political risk. Here, we find indeed that leverage first increases and then decreases in political risk. For ownership, it is comparatively more likely than for leverage that the ownership share decreases with political risk.

Our essay is, to my best knowledge, the first attempt to develop and test theoretical assumptions of the impact of political risk on the financing of multinational structure choice. Few thoughts have been spent on what exactly is captured by political risk and what different types of indices reflect. Further, while the leverage choice of multinational enterprises has been analyzed frequently, mainly in the area of business taxation, the ownership share that a parent company will hold in an foreign affiliate has so far received little attention from econometricians.[2]

In chapter 3 of my dissertation, "Financial constraints and the margins of FDI", we analyze both theoretically and empirically how financial constraints affect a firm's foreign entry decision. A brief history of the research preceding our analysis can illustrate how findings from micro data have deepened the theoretical understanding of economic relationships and have led to the development of new theories.

In international economics, trade data have traditionally been analyzed on an aggregate level.[3] In particular, the fact that only few firms in each sector engage in exporting or invest in foreign countries and that these firms differ in many dimensions from their

[2] Research on the impact of tax incentives on affiliate leverage includes Desai *et al*. (2004), Huizinga *et al*. (2008), Buettner *et al*. (2009), and Mintz and Weichenrieder (2005). The reaction of affiliate leverage to differences in political risk has been analyzed in Desai *et al*. (2004), Aggarwal and Kyaw (2004) and Novaes and Werlang (2005). Asiedu and Esfahani (2001) and Javorcik and Wei (2009) are empirical papers on the ownership choice.

[3] An insightful description on how empirical findings on the micro level have altered our understanding of trade and on how theory has changed in response to empirical findings is provided by Redding (2008).

domestic counterparts has lead to the development of models of heterogeneous firms.[4] Productivity differences have served as an explanation for the fact that firms which are active in foreign markets, whether as exporters or as producers, are larger than purely national firms Helpman *et al*. (2004).) If there are fixed costs of market entry, only the most productive firms should be able to overcome these fixed costs.

However, these models completely ignore the financial side of investment: The basic assumption is that financial markets are perfect, and that firms do not face financial constraints. Using balance sheet data to analyze differences in financial indicators, empirical papers have started to compare exporters to non-exporters with respect to financial variables, and they have found that exporters are less financially constrained than their purely national counterparts. Here, the seminal contribution is a purely theoretical paper by Chaney (2005). Then, empirical research has started from a more aggregated level with a paper by Manova (2010) who uses a panel of bilateral exports for countries and sectors to show that those countries and sectors who are less financially constrained export more. Next, several empirical papers used micro-level data on the firm-level to show that there is a negative relationship between financial constraints and a firm's propensity to engage in exporting (Greenaway *et al*. (2007)).

All of the papers about how multinational activity is related to financial constraints have considered exporting firms. None has attempted to analyze what drives firms' decisions to found an affiliate in a foreign country. This is presumably not due to a lack of interest, but due to a lack of data: Datasets that contained a random sample of firms, national and multinational, as well as their foreign affiliates, simply did not exist. By merging the Dafne, a commercial dataset, providing information on a large panel of German firms, with the Microdatabase Direct Investment of the Deutsche Bundesbank (MiDi), which provides information on the foreign affiliates of German enterprises, we were able to create such a dataset. Our data are unique as they allow measuring financial constraints and productivities at the parent level both for domestic firms and for multinationals, as well as financial constraints at the affiliate level.

In our theoretical model, we analyze how productivity and financial constraints affect a firm's choice to become a multinational firm under conditions of limited internal funds and the need to obtain external debt finance. Our model features limited contract enforceability and liquidation costs as two sources of inefficiencies in financial contracting that are particularly relevant for foreign investments. The model provides a set of testable implications concerning the impact of financial constraints, productivity, and host-country

[4] Bernard and Jensen (1995) on empirical findings on exporters, Mayer and Ottaviano (2008) on FDI, and Helpman *et al*. (2004)

characteristics on firms' internationalization choices. In particular, we predict that financial constraints are more likely to affect the extensive than the intensive margin, unless financial constraints are severe. Furthermore, we predict that financial constraints are more strongly felt for large firms, as they are most likely to be interested in foreign expansion.

In our econometric analysis, we find that productivity and financial constraints have a significant impact on German firms' internationalization decisions. Economically, productivity and financial constraints are of similar importance, but financial constraints matter most to the subset of firms that consider investing abroad. As suggested by our theoretical model, financial constraints have indeed a negative impact on the extensive margin of FDI, but less so on the intensive margin. However, we also find that, in contrast to the parent-level constraints, the affiliate's financial constraints matter for the intensive margin. This observation points toward a hierarchy of financing the intensive margin, with affiliate financing being preferred over parent financing.

Chapter 4, the essay "On the use of information in repeated insurance markets", was inspired by a purely empirical paper by Finkelstein and Poterba (2006). They demonstrate the presence of adverse selection in annuity purchases by showing that there exist "unused observables", namely information available to insurance companies that is correlated with risk experience, but that is not used when setting insurance premia. In their data, the address of the insured person is an example of information that is almost always collected by insurance companies, but that is seldom used in pricing insurance, although there is a correlation between geographic information and other individual attributes. Further empirical evidence on unused observables is provided by Brown and Finkelstein (2007) (gender and place of residence in the U.S. long-term care insurance industry) and Ivaldi (1996) (smoking in the French automobile insurance industry). Finkelstein and Poterba (2006) conclude their article by stating that "a complete understanding of the limited use in pricing of available or collectable risk-related information on insurance buyers remains an open issue".

However, we are able to explain this phenomenon using a repeated model of an oligopolistic insurance market where insurance companies take into account the impact of pricing decisions on competitors' actions.

In our model, there are two types of individuals who face either high or low risk of damage. There is a finite number of insurance companies. They can distinguish between these risks and compete for costumers by setting insurance premia in each period. Companies interact strategically and preconceive the effect of their pricing decisions on the prices set by their competitors in subsequent periods. If companies fear a price war after adjusting their prices, they may refrain from doing so. We show that even if insurance companies can distinguish between risk types, equilibria exist in which first, insurance companies charge the same insurance premium to both risks and second, both risk types purchase positive amounts of insurance. Thus, we derive an equilibrium with unused observables.

We then show that the same equilibrium outcome is possible if insurance companies cannot distinguish between high and low risks, i.e. if there is asymmetric information. This renders possible the following explanation for the existence of unused observables: initially, there exists a collusive equilibrium in an insurance market with asymmetric information. Then, after analyzing their data, insurance companies learn how to distinguish between high and low risks. However, they maintain their pricing schedules in order to sustain collusion. Further, we show the robustness of our model with regard to two extensions: Market entry and explicit collusion.

With regard to this chapter of my dissertation, two further points seem worth to be mentioned in connection with the combination of theory and empirics and with regard to the use of micro data: First, although empirical work seems to suggest that the common market structure in most insurance markets is oligopolistic, the vast majority of theoretical models on the insurance market are one-shot models of either perfect competition or monopolistic behaviour of insurance companies.[5] Empirical evidence suggests that more work should be devoted to imperfectly competitive models of insurance markets - or, in the words of Chiappori et al. (2006), "there is a crying need for such models".

Second, the use of information in the insurance market itself has been changed dramatically by the revolution in information technologies which enables insurance companies to collect, analyze and make use of large amounts of information. An example of evolving information that has recently received a lot of attention is that of genetic testing (Hoy and Witt 2007, Hoy and Polborn 2000, Rees and Apps 2008). A promising field of research for theorists and econometricians will be to analyze how insurance companies will use this information, how they will cope with its constant change, and how this will be affected by the competitive structure of the insurance market.

Chapter 5, the essay "What if everybody had a choice?" addresses a topical public policy question by applying appropriate econometric methods to a specially designed survey. It contributes to the discussion on Medicare Part D by using a hypothetical choice experi-

[5] Concentration indices for the top 5 insurance companies in the non-life business in Europe in 2002 ranged from 27 percent in Germany to 89 percent in Finland (Buzzaccki and Valletti 2005). Concentration measures of the life insurance sector in most developed nations in the 1990s have been constantly high: even in the USA, the least concentrated market, concentration indices for the top 5 insurance companies have been above 25 percent, while they have been (far) above 50 percent in Australia, Canada, Japan and the Netherlands (Bikker and van Leuvensteijn 2008). Market concentration is also reflected in insurance premiums (Dafny *et al.* (2009)).

ment that was conducted using a random sample of the relevant population in order to analyze the demand for drug insurance with different levels of coverage.

In 2006, Medicare Part D, a highly subsidized market for prescription drugs for the elderly, was newly introduced in the US. Before the introduction of Medicare Part D, prescription drug insurance had not been part of the 43 million elderly and disabled U.S. residents on Medicare. Some beneficiaries had coverage from some other source, but about a third of the relevant population remained without prescription drug insurance which imposed a high financial burden on the elderly and led to cost-related non-adherence (McFadden *et al*. (2008) and Madden *et al*. (2008)). Medicare Part D has been the largest single expansion in social insurance in the US since 1965, and it is often used as a natural experiment to study the optimal design of social insurance programs. Lessons from this experiment will be crucial both for deciding whether to introduce universal health care in the USA and for the design of social insurance programs in other countries. Part D also gives important insights into the practicality of Consumer Directed Health Care, an approach achieving efficient allocation of health care resources by confronting consumers with the full marginal cost of the services they use. Medicare Part D can further give insights in how consumers behave in real-world decision situations that are characterized by complexity, ambiguity and important consequences.

Therefore, a lively discussion of the consequences of Medicare Part D has taken place in numerous studies in economics and health. Most of these studies have restricted their analysis to the relatively small group of “active deciders”: those consumers who had no prescription drug coverage before the introduction of Medicare Part D. However, many other groups of consumers were affected by the introduction of Medicare Part D. [6] We show that these groups differ significantly from the active deciders with respect to their health and socio-economic characteristics, but also with respect to their insurance demand.

Our essay contributes to the discussion about Medicare Part D by using a hypothetical choice experiment that was conducted using a random sample of the relevant population to analyze consumer demand. Thus, the whole potential market is included in the analysis instead of restricting the focus to a small group of consumers. In hypothetical choice experiments, individuals are asked to choose between different commodities whose attributes vary in order to infer the utility associated with these attributes and consumers' willingness-to-pay. Here, respondents are asked to choose between insurance contracts that differ in their level of coverage.

[6] For example those who had some privately bought or employer sponsored coverage before or those on Medicaid. Many consumers were not given an equally free choice regarding their insurance coverage - as for example the consumers who were eligible for both, Medicaid and Medicare, and who were automatically enrolled and randomly assigned to prescription drug plans.

So far, hypothetical choice experiments have mainly been used to create variation in product attributes. We want to draw attention to the fact that they can be used to elicit the demand of consumer groups whose choices cannot be observed in the actual market. This becomes important when making predictions about the impacts of policy changes. As we observe actual choices for one group of consumers, we can estimate a joint model using both real and hypothetical choices, thereby mitigating the shortcomings of a separate analysis of real and hypothetical choices.[7]

We find that willingness-to-pay for drug insurance is low for consumers with either low expected drug costs or low income. On the other hand, consumers demand extensive coverage if they are currently in poor health, expect high future drug costs, but also if they are risk averse. These findings conflict with consumers' real choices where neither health nor socio-economic indicators prove significant. A possible reason for this discrepancy is that the active deciders for whom we observe actual choices are too homogeneous. In order to analyze the impact of socio-economic conditions on insurance demand we need to consider sufficiently heterogeneous consumer groups, which is possible through our hypothetical choice experiment. Further, we show that willingness-to-pay of the passive participants is significantly higher than those of the active deciders. Therefore, welfare estimates of the introduction of Medicare Part D taking into account the active deciders only might be too low.

The introduction of Medicare Part D can also serve as an example on how empirical and theoretical research can have important policy implications: First, the market itself was designed taking into consideration lessons from economic theory: With regard to the demand side, the government has designed a standard insurance contract for this market that has features that are supposed to overcome adverse selection and excessive use of medications. On the supply side, the government has designed an auction mechanism that provides incentives for insurance companies to pass bulk discounts that they receive from pharmaceutical companies through to consumers. In order to evaluate market outcomes researchers have used consumer surveys and data from insurance companies and the government. Well-designed hypothetical choices experiments can help to predict impacts of policy changes and to generalize findings from one specific consumer group to other groups of the population.

For each chapter of my dissertation, I have pointed to some open questions for future research. In the following years, my scientific agenda will be to reassume these lines of research on important applied questions in close connection with theory. A very promis-

---

[7] Strength and shortcomings of real and hypothetical choice data are discussed in detail in chapter 5.3 of this thesis.

ing field for this type of research is to apply dynamic and/ or imperfectly competitive models from industrial organization to the fields of health and insurance. Few theoretical researchers have been active in this field so far, and the amount of data that becomes available to researchers in this area is increasing steadily. One present project is to extend the essay “Real versus financial constraints to multinational activity” by analyzing the decision to enter a foreign market via exports and/ or FDI in the presence of financial constraints. In another present project, we combine the Rothschild and Stiglitz (1976) model with a dynamic framework in which consumers can exert preventive effort in several periods in order to reduce their risk of damage, exhibit hyperbolic discounting and suffer from a non-monetary loss in case of (health) damage. Thus, we are able to overcome some features of standard models which contradict empirical findings.

Chapter 2

Who is afraid of political risk? Multinational firms and their choice of capital structure*

2.1 Introduction

Multinational enterprises (MNE) have to adapt their optimal investment strategy to local conditions worldwide. Most notably, they have to respond to different political environments that in different locations may give rise to varying political risks. Political risk encompasses not only 'sovereign risk', the risk that the sovereign will interfere with a firm's ability to pay its investors as promised, but also other forms of political, economic and country-specific risks that affect the profitability of an investment in a foreign country and that would not be present if the country had more stable and developed business environment and legal institutions (Hill (1998) and Buckley (2003)). This risk ranges from outright expropriation to more subtle forms like confiscatory taxation, corruption, or economic constraints such as exchange rate controls. MNE can try to insure against political risk, but they can never do so fully.[8]

In this chapter we investigate, both theoretically and empirically, the way MNE choose their capital structure in response to political risk. For this purpose, we distinguish different types of political risk. We find that it is important to identify the type that prevails in a particular country, because different types of risk affect the optimal financing decision in different ways.

We focus on two choice variables that determine the capital structure - the level of leverage and the ownership structure of the foreign affiliate. Choosing higher leverage reduces tax

* This chapter is based on joint work with Monika Schnitzer.

[8] First, the insurance market for political risk is incomplete because most types of political risk are not contractible and because the market suffers from severe asymmetric information (see for example Desai *et al.* (2008)). Second, many investors are unaware of the existence of political risk insurance and even those who are aware of its existence often do not have such insurance (www.political-risk.net).

payments but increases the risk of bankruptcy, involving some bankruptcy cost. The investor balances optimally these costs and benefits of debt financing. Furthermore, he chooses how much of the affiliate to sell to outside investors, taking into account how the ownership structure affects the agency costs and the value of the affiliate.

We distinguish three prototypes of political risk throughout the chapter.[9] In Scenario I, political risk takes the form of outright expropriation or nationalization, where the investor loses all assets and can no longer service his debts. In the past, this type of political risk used to be very important (Kobrin (1980), Andersson (1991)). Though in general it seems to be less prevalent nowadays there are very recent examples indicating its latent relevance.[10]

Scenario II captures political risk as a form of creeping expropriation that lowers the expected returns of the project. Examples could be lack of protection for intellectual property rights or unreliable contract enforcement, but could also be economic constraints like currency or exchange rate controls, or particular regulatory requirements directed at foreign multinationals. Political violence that negatively affects market conditions and hence expected revenues would be another example.

In Scenario III, we capture political risk that directly affects the profits of the investment, i.e., after servicing potential debt payments. This type of political risk arises if the host country imposes discriminatory and confiscatory taxation, asks for bribes or blocks the repatriation of funds from the host country to the home country.

Our analysis shows that these different forms of political risk affect equity holders and debt holders in different ways and can therefore result in the multinational choosing different capital structures. We find that the optimal debt level decreases with increasing political risk in both Scenarios I and II because (creeping) expropriation increases the risk of bankruptcy, which calls for smaller leverage. But the optimal debt level increases with political risk in Scenario III because the negative effects of discriminatory or confiscatory taxation can be contained with higher leverage. Furthermore, we find that, in all three scenarios, the optimal ownership share tends to decrease as the level of political risk increases because political risk reduces the investor's expected returns from the foreign affiliate, but does not reduce the managerial cost of running the firm. Interestingly,

[9] For a description of various forms of political risks see Buckley (2003), Hill (1998).

[10] Recent examples where expropriations have taken place are Zimbabwe and Venezuela. In 2007, a law was adopted in Zimbabwe that forces foreign investors to find a majority Zimbabwean shareholder within five to ten years (compare for example Deutsche Presse Agentur (2007) or Voice of America (2008)). In 2008, both Cemex, a Mexican cement company, and Sidor, Venezuela's biggest producer of steel, were expropriated by the Venezuelan government without paying appropriate compensation (compare for example The Associated Press (2008), Romero (2008), or Ackerman (2008)).

though, due to the interaction with the optimal debt choice, the ownership share reduction is less pronounced in case of confiscatory taxation, where the debt level increases, as compared to the case of (creeping) expropriation, where the debt level is reduced as a response to political risk.

In our empirical analysis, we use the Microdatabase Direct Investment (MiDi) of the Deutsche Bundesbank to investigate the impact of political risk on both the choice of ownership shares and the leverage of foreign affiliates of German multinationals. The dataset contains balance sheet information on the foreign affiliates. German parent companies are required by law to report this information when the balance sheet total of the affiliate and the ownership share are larger than a certain threshold. As a measure for political risk, we use the time-varying, country-specific index that is provided by the International Country Risk Guide (ICRG) and that is composed of 12 weighted variables covering both political and social attributes.

We estimate the impact of political risk on our two choice variables, ownership share and leverage. We start by looking at the linear effects of political risk. Our ownership regression indicates that MNEs hold a smaller share of the equity of the foreign affiliate when political risk is high, confirming our theoretical predictions. Regarding the leverage choice, we find that affiliates of MNE use a higher level of debt in countries with a higher level of political risk. This would suggest the prevalence of the less invasive Scenario III type of political risk.

We then attempt to capture the theoretical insight that the effects of political risk may differ for different scenarios of political risk. For this purpose we follow two different strategies. One strategy is to employ various empirical measures, reflecting different types of political risk, and to identify them with different scenarios. Following this strategy of using different measures we can confirm that the ownership share decreases in political risk no matter what type of measure we introduce, while leverage can either increase or decrease in response to different types of risk. The second strategy is to rank the scenarios according to their severity and to introduce a quadratic function of political risk which allows the marginal effects of political risk to vary with the severity of political risk. Following this approach, we find that indeed leverage first increases and then decreases in political risk. For ownership, it is comparatively more likely that the ownership share decreases with political risk.

Our essay is related to two strands of literature, the literature on political risk and the literature on the capital structure choice.

The first strand studies the effects of political risk on foreign direct investment. The early theoretical papers were primarily concerned with the question of how foreign direct investment can be sustained if there is a risk of nationalization. The seminal paper in this literature is Eaton and Gersovitz (1983), which shows under what circumstances reputation can sustain foreign direct investment. Other papers study how political risk affects the multinational's investment strategy. This may induce the investor to choose an inefficient technology (Eaton and Gersovitz (1984)), inefficient investment paths (Thomas and Worrall (1994), and Schnitzer (1999)) or excess capacity (Janeba (2000)). More recent papers have investigated the sale of shares to locals as a possible way of mitigating the risk of confiscatory taxation or creeping expropriation (Konrad and Lommerud (2001), Mueller and Schnitzer (2006)). However, none of these authors have allowed for different forms of political risk that impact on the investor's decisions in different ways. Empirical studies have focussed on the question of how country characteristics affect the ownership structure in foreign direct investment projects (Asiedu and Esfahani (2001)).

The second strand of literature has so far mainly focused on taxes as the driving force behind the capital structure choice. It has been shown both empirically and theoretically that tax incentives lead to national differences in the level of leverage of affiliates of MNE (see for example Desai *et al*. (2004), Huizinga *et al*. (2008), Buettner *et al*. (2009) and Mintz and Weichenrieder (2005)). However, there is much less evidence on how differing levels of political risk may affect the capital structure of affiliates that are located in different countries. For US-data, Desai, Foley and Hines (2004) find that political risk increases affiliate leverage. Aggarwal and Kyaw (2004) also use US data, but on a more aggregated level. In contrast to Desai, Foley and Hines, they find that political risk reduces affiliate leverage. Novaes and Werlang (2005) study foreign affiliates in Brazil and find that they are more highly leveraged than their Brazilian counterparts and that the difference increases with Brazil's political risk. This conflicting evidence suggests that the relationship between political risk and leverage is not straightforward and hence needs more examination.

The contribution of our essay is twofold. Firstly, it provides a theoretical framework that explicitly models the choice of the capital structure in response to political risk. It is also the first attempt to distinguish in a theoretical framework different kinds of political risk. Secondly, our essay contributes to the empirical literature on political risk by investigating the impact of political risk on both leverage and ownership choices, and by distinguishing the impact of different types and magnitudes of political risk. As our theoretical analysis suggests, the coefficient measuring the impact of political risk on leverage may indeed change signs, depending on the type of political risk. We find this possibility of different coefficients confirmed in our empirical analysis.

The remainder of the chapter is organized as follows. Section 2.2 introduces our theoretical model and analyzes the optimal financial structure in the baseline model. Section 2.3

introduces different types of political risk and investigates the optimal financial structure in the presence of political risk. In section 2.4, we derive empirical predictions. Section 2.5 introduces the data set. In Section 2.6 we present our empirical results. Section 2.7 concludes.

2.2 The base line model

Consider a multinational investor who intends to invest a fixed amount I in a foreign location. The project generates a stochastic return R, with R being uniformly distributed on the interval $[0, \bar{R}]$. The investment can be financed with either debt, D, or equity, E, or a combination of the two, such that $E + D = I$.

The investor has to take two decisions, he has to choose (i) how much debt finance D to use as opposed to equity finance and (ii) what share α of the affiliate to keep for himself, i.e., what share $(1 - \alpha)$ of the affiliate to sell to outside investors. In the corporate finance literature, these decisions have always been studied separately. Thus, to jointly investigate both decisions, we set up a model that draws on both strands of the corporate finance literature, the one on the debt versus equity decision and the one on the optimal ownership share.

Leverage choice

To model the debt versus equity choice, we consider the well established tradeoff that debt financing saves on tax payments but increases the probability of bankruptcy, giving rise to potential dead weight losses.[11] To capture this tradeoff, we assume that in case of debt financing D the investor's liability is restricted to the investment project. So if the investor takes up debt D, he owes $(1+r)D$, where r denotes the interest rate. He can do so only when the project is sufficiently successful, i.e., generates returns $R \geq (1+r)D$. The remaining profits, after interest payments have been deducted, are subject to local taxation, at interest rate t. If the returns are not sufficient to cover the repayment, the project is liquidated and the investor has to give up all the returns to the bank. The affiliate's value V for a given level of D is thus

$$V = \int_{(1+r)D}^{\bar{R}} (1-t)[R - (1+r)D]\frac{1}{\bar{R}}dR + D - I \tag{2.1}$$

[11] In the corporate finance literature, this tradeoff is often called the tradeoff between interest tax shields and the costs of financial distress. See, e.g., Brealey and Myers (2000), pp. 496 ff. Seminal papers on this issue are Kraus and Litzenberger (1973), and Scott (1976).

Banks are assumed to operate in a competitive market and to be risk neutral. This means that, for any level of debt that the investor wants the bank to finance, the interest rate r is chosen such that the bank can expect to break even. We assume that in case of bankruptcy transaction costs are incurred during the bankruptcy procedure such that the bank is able to seize only some share s of the returns that are generated, with $s < 1$. This assumption is supposed to capture the dead weight loss that is associated with debt financing due to the risk of bankruptcy. The interest rate is implicitly defined by the bank's break even condition

$$\int_{(1+r)D}^{\bar{R}} (1+r)D\frac{1}{\bar{R}}dR + \int_{0}^{(1+r)D} sR\frac{1}{\bar{R}}dR = D \tag{2.2}$$

where the first integral represents the expected interest payments and the second integral represents the expected payoff the bank realizes in case of bankruptcy.

Ownership choice

Consider now the ownership choice of the investor, i.e. what share $(1-\alpha)$ of the foreign affiliate to sell to outside investors. If financial markets are competitive and perfect, i.e. with no transaction costs and risk neutrality, then the assets will be valued by outside investors at their expected value. So the price the investor realizes for selling a share of $(1-\alpha)$ of the affiliate, $P((1-\alpha)V)$, is equal to $(1-\alpha)V$.[12] We will discuss below how things change if markets are not perfect, in particular if the exposure of outside investors to political risk differs from that of insiders.

To model the ownership decision, we draw again on the corporate finance literature. In this literature, two potential effects are discussed why (and how) the ownership choice may affect the firm's value. The convergence-of-interest hypothesis predicts that larger ownership stakes are associated with higher firm values, because they allow for a better alignment of the incentives of owner and manager or inside and outside investors.[13] The entrenchment hypothesis, on the other hand, suggests that insiders who control a substantial part of the shares may indulge in non-value maximizing behavior at the cost of outside investors.[14] Empirical studies on this issue suggest that the firm value as a function of the insider's ownership share α follows an inverted u-shaped pattern. This evidence has been interpreted such that the convergence-of-interest effect dominates for low and the entrenchment effect for large values of α.[15]

[12] See, for example, Ross (1977).

[13] Berle and Means (1932), Jensen *et al.* (1976).

[14] Demsetz (1983), Fama and Jensen (1983).

[15] See, e.g., Morck *et al.* (1988), McConnell and Servaes (1990) and Short and Keasey (1999).

Following the convergence-of-interest hypothesis we assume that the investor increases his effort as α increases, at cost $K(\alpha)$, with $K'(\alpha) > 0$ and $K''(\alpha) > 0$, as larger ownership stakes imply more effort to be spent on the firm. Following the entrenchment hypothesis, we assume that the investor enjoys private benefits $B(\alpha)$ that increase in α, i.e. $B'(\alpha) > 0$, but at a decreasing rate, i.e. $B''(\alpha) < 0$. The combination of effort and entrenchment effects are reflected in the revenues of the affiliate that can be appropriated by the investors in the following way: $\bar{R}$ is an inverted u-shaped function of α, with $\bar{R}'(\alpha) > 0$ for $\alpha = 0$, $\bar{R}''(\alpha) < 0$ and $\bar{R}'(\hat{\alpha}) = 0$ for $0 < \hat{\alpha} < 1$.

The investor's payoff function from holding equity share α, selling equity share $(1 - \alpha)$ and experiencing private benefits and costs can thus be summarized as follows:

$$U(D, \alpha) = \alpha V(D, \alpha) + P[(1-\alpha)V(D, \alpha)] + B(\alpha) - K(\alpha) = V(D, \alpha) + B(\alpha) - K(\alpha) \tag{2.3}$$

The investor chooses both the optimal debt level D and the optimal ownership share α in order to maximize his payoff function.

The optimal financial structure

We now turn to solving the base line model without political risk. For this, consider again the investor's payoff function, which, using the specification of the firm value in equation (2.1), is

$$\begin{aligned} U(D, \alpha) &= V(D, \alpha) + B(\alpha) - K(\alpha) && (2.4) \\ &= \int_{(1+r)D}^{\bar{R}(\alpha)} (1-t)[R - (1+r)D] \frac{1}{\bar{R}(\alpha)} dR + D - I + B(\alpha) - K(\alpha) && (2.5) \\ &= \frac{1-t}{\bar{R}(\alpha)} \left[\frac{1}{2} \bar{R}^2(\alpha) - (1+r)D\bar{R}(\alpha) + \frac{1}{2}(1+r)^2 D^2 \right] + D - I + B(\alpha) - K(\alpha) \end{aligned}$$

Using equation (2.2) we can rewrite the payoff function in the following way (see Appendix)

$$U(D, \alpha) = (1-t) \left[\frac{1}{2} \bar{R}(\alpha) - \frac{(1-s)(1+r)}{2-s} D \right] + \frac{1-s+t}{2-s} D - I + B(\alpha) - K(\alpha) \,. \tag{2.6}$$

The investor's maximization problem is characterized by the following two first order conditions.

$$\begin{aligned} \frac{dU}{d\alpha} &= (1-t) \left[\frac{1}{2} \frac{d\bar{R}}{d\alpha} - \frac{1-s}{2-s} D \frac{dr}{d\bar{R}} \frac{d\bar{R}}{d\alpha} \right] + B' - K' = 0 && (2.7) \\ \frac{dU}{dD} &= -(1-t) \frac{1-s}{2-s} \left[(1+r) + \frac{dr}{dD} D \right] + \frac{1-s+t}{2-s} = 0 && (2.8) \end{aligned}$$

The following result summarizes the solution to this maximization problem:

Result 1 *The investor's maximum payoff is*

$$U^*(D^*, \alpha^*) = \frac{\bar{R}(\alpha^*)}{2} \frac{1 - s + st}{1 - s + t} - I + B(\alpha^*) - K(\alpha^*) \tag{2.9}$$

where α^* is implicitly defined by

$$\left[\frac{1 - s + st}{1 - s + t}\right] \frac{1}{2} \frac{d\bar{R}}{d\alpha} + B' - K' = 0 \tag{2.10}$$

and D^* is given by

$$D^* = \frac{\bar{R}}{2} t \frac{2(1 - s) + st}{(1 - s + t)^2} \tag{2.11}$$

Proof: See Appendix

The following result summarizes the comparative statics with respect to the local taxation rate t.

Result 2 *The optimal debt level D^* increases as the local tax rate t increases.*

$$\frac{dD^*}{dt} = \frac{\bar{R}(1 - s)^2}{(1 - s + t)^3} > 0 \tag{2.12}$$

The optimal ownership share α^* decreases as t increases, if $\bar{R}' > 0$ in the relevant parameter range, i.e.

$$\frac{d\alpha^*}{dt} = \frac{\left[\frac{(1-s)^2}{(1-s+t)^2}\right] \frac{1}{2} \frac{d\bar{R}}{d\alpha}}{\left[\frac{1-s+st}{1-s+t}\right] \frac{1}{2} \frac{d^2\bar{R}}{d\alpha^2} + B'' - K''} < 0 \quad \textit{if} \quad \bar{R}' > 0. \tag{2.13}$$

Proof: See Appendix

Result 2 shows the expected pattern for the optimal debt level: as the tax rate increases, the debt level increases. It is also straightforward to show that $\frac{dD^*}{ds} > 0$, i.e. as the inefficiency of bankruptcy decreases, the optimal debt level increases. This captures the well established tradeoff of debt versus equity financing. For the ownership share, we find that α decreases if $\bar{R}' > 0$ in the relevant parameter range, which, as equation 2.10 shows, is the case if $B' < K'$, i.e. unless the entrenchment benefits are particularly large.

2.3 Political risk and the optimal financial structure

Consider now that the investment project is subject to political risk in the foreign location. To study how this affects the firm's financial structure, we distinguish different forms of political risk.

(1) Expropriation

The first type of political risk we consider is expropriation or nationalization. This is the classic form of political risk where a sovereign simply takes property without compensation (Buckley (2003), Hill (1998)). We capture this form of political risk by assuming that the investment is expropriated with some probability π_1, i.e. the investors lose control and cash flow rights from the investment. This leads to the following modified firm value function.

$$V_1 = (1-\pi_1)\int_{(1+r)D}^{\bar{R}}(1-t)[R-(1+r)D]\frac{1}{\bar{R}}dR + D - I \tag{2.14}$$

Credits are serviced only if the investment is not expropriated. So the zero profit condition for banks needs to be modified. This is captured by the following condition.

$$(1-\pi_1)\left[\int_{(1+r)D}^{\bar{R}}(1+r)D\frac{1}{\bar{R}}dR + \int_0^{(1+r)D} sR\frac{1}{\bar{R}}dR\right] = D \tag{2.15}$$

Of course, if the investment is expropriated, the investor also loses his benefits from potential entrenchment, $B(\alpha)$. His managerial effort costs of running the firm are incurred before the potential expropriation takes place and hence are not affected. Thus, the investor's payoff function is now

$$U_1(D,\alpha,\pi_1) = (1-\pi_1)\int_{(1+r)D}^{\bar{R}}(1-t)[R-(1+r)D]\frac{1}{\bar{R}}dR + D - I + (1-\pi_1)B(\alpha) - K(\alpha) \tag{2.16}$$

This payoff function implicitly assumes that outside investors, no matter where they are located, are affected by the danger of expropriation in exactly the same way as the multinational investor. Of course, there may be circumstances where they are better protected against expropriation, e.g. because they are local investors and the government compensates them for their losses. In this case their perceived value of the affiliate is not as negatively affected as the multinational investor's perceived value. We will discuss below how this affects our results.

(2) Creeping Expropriation

Another form of political risk we consider is creeping expropriation or political violence that negatively affects the expected returns of the investment project. Other examples would be currency or exchange rate restrictions, or a failure to enforce or respect the agreed property and contract rights (Buckley (2003), Hill (1998)). We capture this by assuming that the investor can capture only a share $(1-\pi_2)$ of the returns R. This leads to the following modified firm value function.

$$V_2 = \int_{\frac{(1+r)D}{1-\pi_2}}^{\bar{R}} (1-t)[(1-\pi_2)R - (1+r)D]\frac{1}{\bar{R}}dR + D - I \tag{2.17}$$

The expected returns of the investment project also affect the zero profit condition for banks that needs to be modified in the following way.

$$\int_{\frac{(1+r)D}{(1-\pi_2)}}^{\bar{R}} (1+r)D\frac{1}{\bar{R}}dR + \int_0^{\frac{(1+r)D}{(1-\pi_2)}} s(1-\pi_2)R\frac{1}{\bar{R}}dR = D \tag{2.18}$$

As above, also creeping expropriation jeopardizes the investor's chances of enjoying benefits from entrenchment but leaves managerial effort cost unaffected. Thus, the investor's payoff function is now

$$U_2(D,\alpha,\pi_2) = \int_{\frac{(1+r)D}{(1-\pi_2)}}^{\bar{R}} (1-t)[(1-\pi_2)R-(1+r)D]\frac{1}{\bar{R}}dR+D-I+(1-\pi_2)B(\alpha)-K(\alpha) \tag{2.19}$$

We now investigate how political risk affects the optimal financial structure. The following result describes how the investor chooses the optimal ownership share and the optimal debt level in the presence of (creeping) expropriation. Interestingly, we find that both types of political risk, expropriation and creeping expropriation, affect the financial structure and the investor's payoff in the same way. Thus, we can state the following results for U_i, where the subscript $i = \{1,2\}$ captures both cases, the case of expropriation (1) and the case of creeping expropriation (2).

Result 3 *The investor's maximum payoff in case of (creeping) expropriation is*

$$U_i^*(D_i^*,\alpha_i^*) = (1-\pi_i)\frac{\bar{R}(\alpha^*)}{2}\frac{1-s+st}{1-s+t} - I + (1-\pi_i)B(\alpha^*) - K(\alpha^*) \tag{2.20}$$

where α_i^* is implicitly defined by

$$(1-\pi_i)\left[\frac{1-s+st}{1-s+t}\frac{1}{2}\frac{d\bar{R}}{d\alpha} + B'\right] - K' = 0 \tag{2.21}$$

and D_i^* is given by

$$D_i^* = (1-\pi_i)\frac{\bar{R}}{2}t\frac{2(1-s)+st}{(1-s+t)^2} \tag{2.22}$$

Proof: See Appendix

We now determine the comparative statics with respect to the political risk parameter π_i, with $i = \{1,2\}$.

Result 4 *Both the optimal debt level and the optimal ownership share decrease as (creeping) expropriation increases*

$$\frac{dD_i^*}{d\pi_i} = -\frac{\bar{R}}{2}t\frac{(2(1-s)+st)}{(1-s+t)^2} < 0 \tag{2.23}$$

$$\frac{d\alpha_i^*}{d\pi_i} = \frac{\frac{(1-s+st)}{(1-s+t)}\frac{1}{2}\frac{d\bar{R}}{d\alpha} + B'}{(1-\pi_i)\left[\frac{1-s+st}{(1-s+t)}\frac{1}{2}\frac{d^2\bar{R}}{d\alpha^2} + B''\right] - K''} < 0 \tag{2.24}$$

Proof: See Appendix

For a given debt level, the risk of expropriation makes it less likely that interest payments are made. This has to be compensated by a higher interest rate, which in turn increases the risk of bankruptcy. As a consequence, the investor chooses a smaller debt level. In case of creeping expropriation, overall revenues are smaller, increasing the likelihood of bankruptcy for any given level of debt. This leads to the same reduction of the optimal debt level. Thus, in both cases, the problem is that (creeping) expropriation increases the risk of default. To avoid costly dead weight losses the investor reduces his debt exposure.

With the same debt level, the overall payoff of the investor is the same, for any given ownership share. (Creeping) expropriation reduces the monetary and non-monetary payoff from the investment but does not change the managerial effort cost of running the firm. Thus, the investor optimally lowers his ownership share. This effect would be even larger if outside investors were less exposed to political risk and hence would value the foreign affiliate more highly.[16]

Desai *et al*. (2008), who find empirically that debt is higher in high political risk countries, have argued that credits taken by local creditors may not react as much to political risk because local creditors may be more restricted in their choice of investment opportunities. The empirical evidence does, however, suggest that local interest rates react positively to political risk (Desai *et al*. (2004), Aggarwal and Kyaw (2004)).

(3) Confiscatory taxation

Our third scenario captures the type of political risk that directly affects the multinational's profits. Examples would be the blocking of the repatriation of funds from the

[16] This effect will also be larger if the allocation of ownership rights can be used as a means of influencing the likelihood of nationalization. As Konrad and Lommerud (2001) and Schnitzer (2002) have shown, it could be in the interest of the investor to share ownership with host country firms, even without compensation, if this makes the host country less prone to engage in expropriation or confiscatory taxation.

host country to the home country, or corruption and discriminatory and confiscatory taxation that treats foreign firms differently from domestic firms (Buckley (2003)). We model this as a form of profit tax, i.e., interest payments can be deducted and are not subject to taxation or bribes. This scenario is particularly relevant if credits are obtained locally and hence the local government has no interest in jeopardizing the repayment of local credits.

The affiliate's value is thus given by

$$V_3 = \int_{(1+r)D}^{\bar{R}} (1-t-\pi_3)[R-(1+r)D]\frac{1}{\bar{R}}dR + D - I \tag{2.25}$$

This type of political risk has no impact on the zero profit condition for banks, provided that the government indeed exempts the interest payments from discriminatory taxation.

Thus the break even condition for the bank is the same as in the baseline model

$$\int_{(1+r)D}^{\bar{R}} (1+r)D\frac{1}{\bar{R}}dR + \int_{0}^{(1+r)D} sR\frac{1}{\bar{R}}dR = D\ . \tag{2.26}$$

As this type of political risk applies to profits rather then revenues, the private benefits of the investor, $B(\alpha)$, are not likely to be affected, nor is the managerial effort cost, $K(\alpha)$. This implies the following payoff function:

$$\begin{aligned} U_3(D,\alpha,\pi_3) &= \int_{(1+r)D}^{\bar{R}} (1-t-\pi_3)[R-(1+r)D]\frac{1}{\bar{R}}dR \\ &\quad +D - I + B(\alpha) - K(\alpha) \end{aligned} \tag{2.27}$$

The following result describes how the investor chooses the optimal ownership share and the optimal debt level as a function of political risk.

Result 5 *The investor's maximum payoff is*

$$U_3^*(D_3^*,\alpha_3^*) = \frac{\bar{R}(\alpha_3^*)}{2}\frac{1-s+s(t+\pi_3)}{1-s+(t+\pi_3)} - I + B(\alpha_3^*) - K(\alpha_3^*) \tag{2.28}$$

where α_3^* is implicitly defined by

$$\left[\frac{1-s+s(t+\pi_3)}{(1-s+t+\pi_3)}\right]\frac{1}{2}\frac{d\bar{R}}{d\alpha} + B' - K' = 0 \tag{2.29}$$

and D_3^* is given by

$$D_3^* = \frac{\bar{R}}{2}(t+\pi_3)\frac{2(1-s)+s(t+\pi_3)}{(1-s+(t+\pi_3))^2} \tag{2.30}$$

Proof: See Appendix

We can now determine the comparative statics with respect to political risk π_3.

Result 6 *The optimal debt level increases as confiscatory taxation increases.*

$$\frac{dD_3^*}{d\pi_3} = \frac{\bar{R}(1-s)^2}{(1-s+t+\pi_3)^3} > 0 \tag{2.31}$$

The optimal ownership share α_3^* decreases if $\bar{R}' > 0$ in the relevant parameter range, i.e.

$$\frac{d\alpha_3^*}{d\pi_3} = \frac{\left[\frac{(1-s)^2}{(1-s+t+\pi_3)^2}\right]\frac{1}{2}\frac{d\bar{R}}{d\alpha}}{\left[\frac{1-s+s(t+\pi_3)}{1-s+t+\pi_3}\right]\frac{1}{2}\frac{d^2\bar{R}}{d\alpha^2} + B'' - K''} < 0 \quad \textit{if} \quad \bar{R}' > 0 \tag{2.32}$$

Proof: See Appendix

Note that although the value functions in case of expropriation (2.14) and confiscatory taxation (2.25) look very similar, the break even conditions for the banks (2.15 and 2.26) look different. In contrast to the case of (creeping) expropriation, confiscatory taxation does not affect the investor's ability to repay his debt and hence does not increase the likelihood of bankruptcy. Thus, it avoids the extra cost of capital that comes with potential bankruptcy. Instead, confiscatory taxation induces the investor to choose a higher debt level to avoid this drain on profits. As pointed out above, this strategy is particularly worthwhile if debt holders are not negatively affected and hence do not have to increase interest rates.[17]

The optimal ownership share is likely to decrease as confiscatory taxation increases, provided $\bar{R}' > 0$, which, according to the first order condition that implicitly defines α_3^*, is the case if $B'(\alpha_3^*) < K'(\alpha_3^*)$.

We can summarize the findings from our theoretical analysis as follows:

For the optimal debt level, we find that the effects depend on the type of political risk. In scenario (1) and (2) the effects were negative. Only in scenario (3) did we find a positive effect of political risk on the optimal debt level. The effect of political risk on ownership share tends to be negative in all three scenarios. Only in scenario (3) could it be positive, and only if the benefits from the entrenchment effect are particularly large.

[17] Note that if our interpretation is correct then we should observe more debt default in case of full or creeping expropriation, but not in case of confiscatory taxation. Although we have no direct evidence on the number of defaults, there is suggestive evidence from interest rates that is consistent with our interpretation. Desai *et al.* (2004) report in Table V that interest rates react positively to political risk, but negatively to country tax rates. To the extent that confiscatory taxation has a similar effect as taxation in general, this supports our implication that the likelihood of default differs in the different scenarios.

Result 7 *The optimal ownership share reacts more negatively to (creeping) expropriation than to confiscatory taxation.*

$$\frac{d\alpha_{1/2}^*}{d\pi_{1/2}} < \frac{d\alpha_3^*}{d\pi} \tag{2.33}$$

Proof: See Appendix

This is due to the interaction with the optimal debt choice. The ownership share reduction is smaller in case of confiscatory taxation, where the debt level increases, than in case of (creeping) expropriation, where the debt level is reduced as a response to political risk.

2.4 Empirical predictions

We now turn to the predictions that can be derived from our theoretical analysis. The following predictions capture the results 4 and 6.

Hypothesis 1 *The effect of political risk on the affiliate's debt level is negative in the case of (creeping) expropriation and positive in the case of confiscatory taxation.*

Hypothesis 2 *The effect of political risk on the ownership share is negative in the case of (creeping) expropriation and tends to be negative in the case of confiscatory taxation.*

The problem of directly testing these hypotheses is that it may be difficult to identify empirically which type of political risk is present. An alternative approach would be to rank the different political risk scenarios according to their severity. It seems natural to argue that the first two scenarios of (creeping) expropriation capture a more severe type of political risk because they describe situations where political risk increases the risk of bankruptcy, whereas the third scenario describes a situation where political risk is less severe, because it affects the profitability of the investment only, without jeopardizing the survival of the affiliate and hence without imposing additional dead weight losses. Following this approach, we can state the following hypothesis.

Hypothesis 3 *The more severe the political risk, the more likely the debt level will be negatively affected.*

Similarly, we can formulate our hypothesis about the ownership share.

Hypothesis 4 *The more severe the political risk, the more likely the ownership share will be negatively affected.*

Result 6 also allows us to compare the relative reactions of leverage and ownership share to political risk. Whereas the reaction of debt is always positive in case of confiscatory taxation, this is possible for the ownership share but not necessary. From this, we derive the following prediction.

Hypothesis 5 *The ownership share is more likely to react negatively to political risk than the debt level.*

Finally, we include one prediction about the impact of taxation on the level of debt and the ownership share, based on Result 2.

Hypothesis 6 *The effect of taxation is positive on the optimal debt level and it tends to be negative on the ownership share.*

2.5 Data

The empirical analysis presented in section 2.6 is based on the Microdatabase Direct Investment (MiDi) of the Deutsche Bundesbank. The database contains a panel dataset of yearly firm-level information on German parent companies and their foreign affiliates for the period 1996-2006. The parents are required by law to report information on their investments abroad and on the financial characteristics of their foreign affiliates if the balance sheet total of the affiliate and the ownership share are larger than a certain threshold that varies over time (Lipponer (2006)).

The MiDi contains 469,332 observations with yearly observations over 11 years. As we are interested in outward FDI, we exclude all observations on inward FDI. This leaves 303,870 observations that represent affiliate-year cells. Affiliates can be present more than once a year if several parent companies report on them. Dropping all affiliates that are present more than once in one year leaves 292,494 observations. Deleting all "indirect" FDI leaves 208,441 observations. Deleting all firms that report negative equity removes another 21,489 observations from our dataset.[18] Thus, the final dataset comprises 186,952 observations.

In each year, our sample consists of about 5,000 to 8,000 German parents and of about 15,000 to 24,500 foreign affiliates (compare Table 2.1). The affiliates are located in more than 140 countries.

We augment the MiDi dataset by country-level information. As a measure of political risk, we use the time-varying International Country Risk Guide (ICRG) index provided by

[18] These are cases of ownership chains where dependent companies of German parents invest in other companies.

Table 2.1: Overview of the number of parent companies and affiliates per year

Year	Number of parent companies	Number of affiliates
1996	6,965	17,038
1997	7,277	18,034
1998	7,617	19,108
1999	7,567	19,422
2000	7,963	20,874
2001	8,080	21,509
2002	5,358	13,893
2003	5,178	13,706
2004	5,095	13,945
2005	5,119	14,438
2006	5,124	14,985

the Political Risk Services (PRS) Group. The index is composed of 12 weighted variables covering both political and social attributes.[19] We recode the index in such a way that an increasing value represents higher political risk.

There are numerous indices that try to capture the variation of political risk across countries. A good overview is provided in Howell (2008). For our analysis, the ICRG index is the best choice for three reasons: First, it takes into account diverse dimensions of political risk like corruption, bureaucratic quality, but also ethnic and religious tensions and socioeconomic conditions. Second, while many indices provide only information on a selective sample of countries, the ICRG index covers more than 140 countries. Third, the ICRG index varies according to time and provides information for all years that are covered in the MiDi dataset.

Information on GDP, GDP per capita and the rate of inflation is taken from the World Economic Outlook Database of the IMF (www.imf.org). The Private Credit variable is based on Beck *et al*. (2000). It measures the ratio of private credit lent by deposit money banks to GDP. Statutory tax rates are taken from the Institute for Fiscal

[19] Government stability, socioeconomic conditions, investment profile, internal conflict, external conflict, corruption, military in politics, religion in politics, law and order, ethnic tensions, democratic accountability and bureaucracy quality.

Studies (www.ifs.org.uk), as well as from various issues of the Corporate Tax Guides of Ernst&Young, KPMG and PricewaterhouseCoopers.

Table 2.2 provides descriptive statistics of the variables used in our analysis. The definitions of the variables are standard, and they are also presented in Table 2.2. Comparing the descriptive statistics to those of American MNE in Desai *et al*. (2004), we find that regarding most variables used in our analysis, on average American and German MNEs tend to be quite similar. Both are active in about all countries worldwide. Consequently, the means of our country variables differ only insofar as the time period considered by Desai *et al*. (2004) is 1982 to 1994, while we analyze the year 1996-2006. In the period we analyze, both average inflation and political risk are lower. Affiliate-level variables are also quite similar for German and US-American affiliates, the sole exception being profits over total assets: while Desai *et al*. (2004) report a share of profit to assets of about 15 percent, this ratio is at only 4 percent for the German affiliates.

2.6 Econometric Analysis

The aim of our analysis is to investigate how the capital structure choice of multinational enterprises reacts to political risk. The two choice variables we consider are the ownership share and the level of leverage. We define the level of leverage as debt over total assets and the ownership share as the share of equity of the foreign affiliate held by the German parent.

Leverage has a mean of 0.61 and a standard deviation of 0.31 at the firm-level. The standard deviation of average leverage per country is 0.12. The mean of the ownership share is 0.87, while its standard deviation is 0.24, and the average ownership share per country has a standard deviation of 0.14. The political risk indicator has an average value of 0.19 and a standard deviation of 0.08.

Examples of countries with extremely high political risk in our sample (an average political risk score above 0.45) are Algeria (with a political risk of 0.50), Colombia (with an average political risk of 0.47), Nigeria (with an average political risk of 0.54), Pakistan (political risk at 0.49), Yugoslavia (with an average political risk of 0.47) and Zimbabwe with a political risk of 0.51. The country with the lowest average political risk in our sample period is Luxembourg with an average political risk score of 0.07.

Figure 2.1 presents how the mean of the ownership share by country varies with political risk. The graph suggests a negative relationship between the ownership share and political risk even if we do not control for any other country or affiliate specific factors. When

Table 2.2: Descriptive statistics

	Definition	Mean	Std. deviation	Min*	Max*
Dependent variables					
Leverage	Debt/ total capital	0.508	0.295	0.000	1.001
Ownership share	Share of affiliate's equity held by German parent	0.772	0.350	0.010	1.000
Independent variables (firm-level)					
Fixed/ total assets		0.243	0.270	0.000	1.000
Log(sales)		9.194	1.437	6.908	17.481
Profit/ total assets		0.042	0.175	-6.652	7.943
Independent variables (country-level)					
Inflation		0.418	1.461	-1.000	106.120
Log(GDP)		6.120	1.590	-1.204	9.488
Log(GDP per capita)		9.654	1.061	4.463	11.407
Political risk	Index between zero and one with a higher index reflecting higher political risk	0.190	0.080	0.039	0.748
Private credit	Ratio of private credit lent by deposit money banks to total GDP	0.814	0.409	0.013	2.179
Statutory tax		0.329	0.071	0.000	0.530

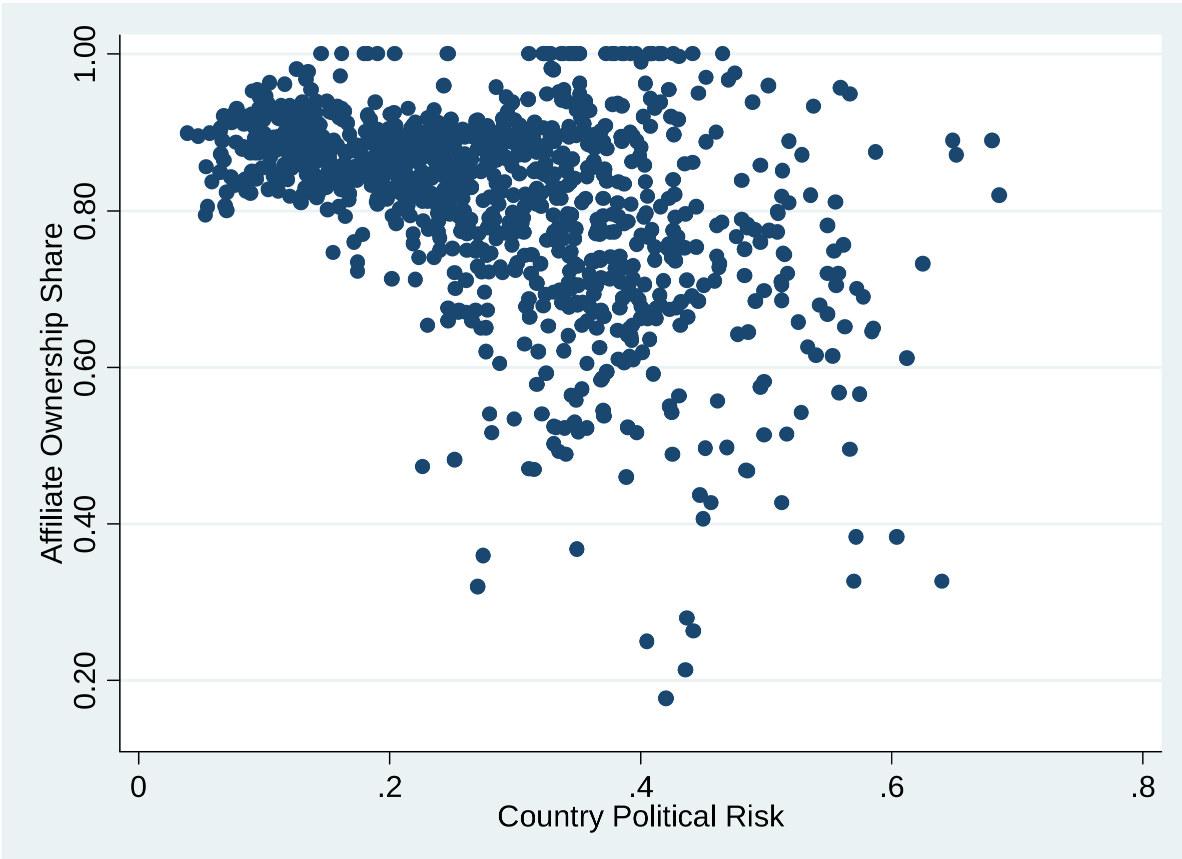
*Averaged over three affiliates

plotting leverage aggregated by country versus political risk (Figure 2.2), we cannot deduct any clear relationship between the two variables from the graph. In a regression analysis at the firm-level we are going to analyze the relationship more thoroughly.

All regressions presented in this chapter are estimated by OLS and include parent-fixed-effects in order to control for parent-specific individual heterogeneity, and we include year and affiliate industry dummies. The reason for this specification is that we can explicitly control for affiliate and country specific heterogeneity, but the only way our data allows us to control for parent-specific heterogeneity is to include parent fixed effects.[20] Thus, we

[20] See Desai *et al.* (2004), Desai *et al.* (2008) and Buettner *et al.* (2009) who use the same fixed effects and dummies in order to control for firm-specific and industry-specific considerations.

Figure 2.1: The relationship between political risk and ownership share, 1996-2006

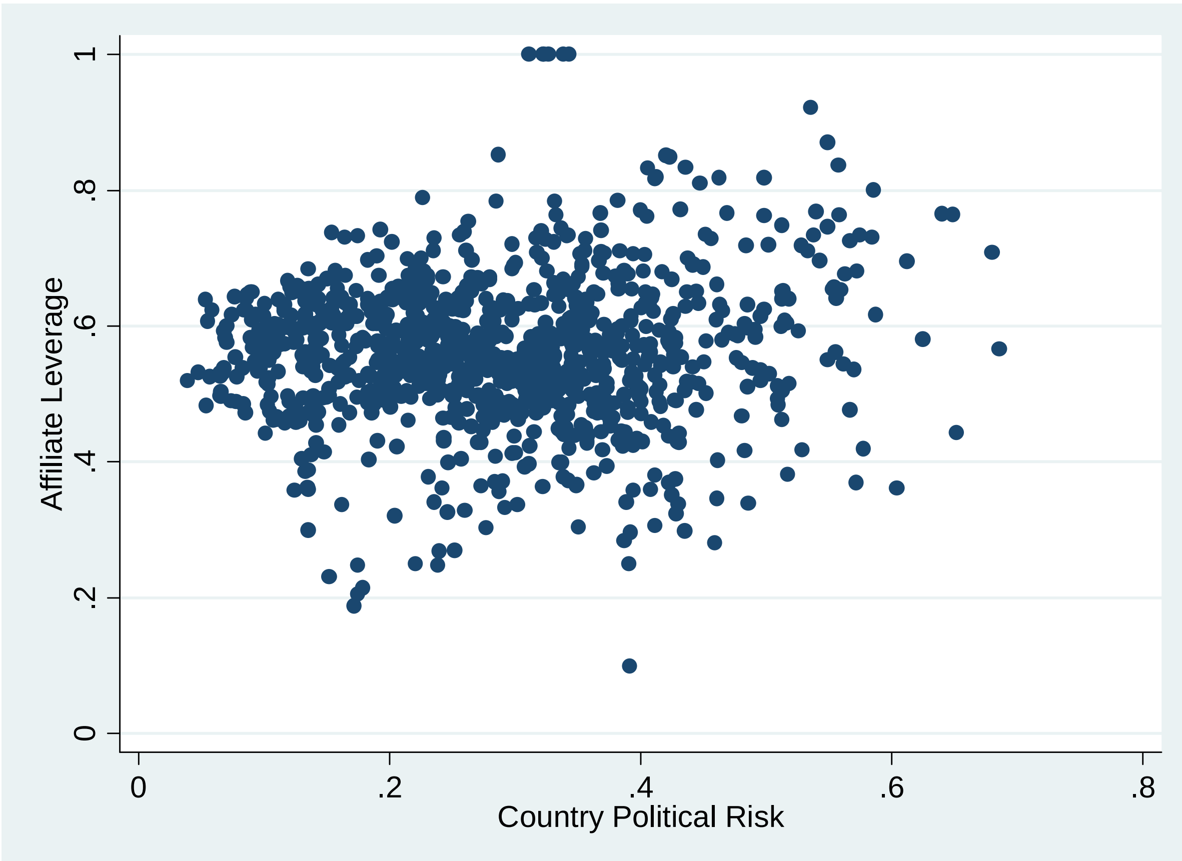


aim to make use of the observed variation between affiliates of the same parent which are located in different countries as well as of the variation over time. Indeed we find that the time variation alone is not sufficient for identification, so that we need the cross-section variation between affiliates of the same parent company located in different countries.[21] In all regressions, we use heteroscedasticity robust standard errors, and we control for clustering of the standard errors by parent company.

In Table 2.3, we present the effect of political risk on both our independent variables as a first test of our hypotheses. We report each regression with and without including our statutory tax variable, because we lose many observations when including it. The choice

[21] When we use pure time variation in the ICRG index, political risk is no longer significant in explaining the leverage choice and it positively affects the ownership choice. The latter effect is not inconsistent with our model (Result 6) but it is not what we would expect to be the dominant effect (Hypothesis 2) and what others have found in cross-section studies (see footnote 15).

Figure 2.2: The relationship between political risk and leverage level, 1996-2006



of independent variables is based on Desai *et al*. (2004) where the authors investigate the capital structure choice of US-American MNEs.

On average, an increase in political risk leads to an increase in affiliate leverage and to a decrease in the ownership share the parent holds in the foreign affiliate. This seems to be consistent with our hypothesis 5 that leverage is more likely to react positively, if at all, to political risk, than to the ownership share. In the leverage regression, the coefficient of political risk of 0.1491 can be interpreted as follows: when political risk increases by one standard deviation of 0.0797, leverage increases by 0.0119 (=0.1491*0.0797), which represents 2.3 percent of its mean value. This effect is quite close to the 2.9 percent estimated by Desai, Foley, Hines (2004) for affiliates of US-American MNEs. The estimated effect of statutory taxes on the leverage share is 0.0146 (=0.2074*0.0705) which represents 2.8 percent of its mean value. Thus, the relative impact of a change of one standard deviation

Table 2.3: The impact of political risk on affiliate leverage and ownership share

	(1)	(2)	(3)	(4)
Dependent variable	Leverage	Leverage	Ownership	Ownership
Political risk	0.117***	0.149***	-0.138***	-0.110***
	(8.11)	(9.04)	(11.41)	(8.25)
Log(sales)	0.011***	0.012***	-0.006***	-0.005***
	(14.70)	(14.35)	(9.02)	(7.44)
Profit/ total assets	-0.423***	-0.406***	-0.013***	-0.009**
	(31.84)	(29.18)	(3.32)	(2.26)
Fixed/ total assets	-0.034***	-0.021***	-0.023***	-0.029***
	(7.15)	(4.29)	(6.57)	(7.60)
Private credit	-0.029***	-0.031***	-0.030***	-0.030***
	(15.26)	(15.77)	(19.05)	(18.42)
Inflation	-0.001	-0.005***	-0.000	0.002
	(1.44)	(3.43)	(0.54)	(1.25)
Log(GDP)	0.002***	-0.001	-0.004***	-0.002**
	(3.40)	(0.71)	(7.69)	(2.39)
Log(GDP per capita)	0.005***	0.004**	0.025***	0.026***
	(3.71)	(2.43)	(20.03)	(18.49)
Statutory tax		0.207***		-0.143***
		(14.82)		(13.26)
Constant	0.312***	0.253***	0.743***	0.738***
	(12.27)	(8.89)	(34.80)	(31.45)
Observations	142,325	127,096	142,309	127,084
Number of parents	11,666	11,315	11,666	11,315
R-squared	0.11	0.12	0.04	0.04

OLS regression including parent fixed effects
Year and affiliate industry dummies included
Heteroscedasticity robust t-statistics (absolute values) in parentheses
Clustering of standard errors by parent company
*** p<0.01, ** p<0.05, * p<0.1

Data sources: Firm-level variables are taken from the Microdatabase Direct Investment of the Deutsche Bundesbank. Private credit is provided in Beck et al. (1999). Inflation, log of GDP, log of GDP per capita are taken from the IMF. Statutory tax rates are taken from the IFS, as well as from the Corporate Tax Guides of Ernst&Young, KPMG and PricewaterhouseCoopers.

in political risk on leverage versus a one standard deviation change in statutory taxes is about 81 percent. The estimated effect of political risk on the ownership share is smaller. A one standard deviation increase of political risk decreases the ownership share by 1.1 percent of its mean value.[22]

We include some other exogenous variables that are not directly related to our model in order to make our results comparable to other results presented in the literature. In particular, we use the same set of controls as Desai *et al*. (2004) and Desai *et al*. (2008), who include these variables to control for observable dimensions of country and firm heterogeneity, and we find the same signs of the coefficients in our regressions as they do. Larger affiliates, for example, use more debt and hold smaller ownership shares. These coefficients may be a reflection of financial constraints of foreign investors, something not explicitly modeled in our framework, which require them to turn to more outside finance in case of larger investment projects. Along the same lines, we find that affiliates with higher profits/ assets use less debt, as they seem to have a greater cash flow generating capacity. The fixed assets reported in the Bundesbank data include tangibles as well as intangible assets. Thus, the negative coefficient of the fixed asset variable could reflect the fact that firms with a larger share of intangible assets find it more difficult to use debt, as intangible assets are usually not accepted as collateral. We also find that firms have smaller ownership shares in more profitable firms, which seems counter-intuitive. The effect, however, is economically not very significant: For an increase in profits of one standard deviation, firms decrease their ownership share only by 0.005 standard deviations. This finding may also be due to the fact that we only observe accounting profits which in turn may reflect a firm's tax saving strategy. Including other macro variables, like inflation, GDP and GDP per capita, is important as a robustness test of our political risk variable and indeed their inclusion does not turn the political risk variable insignificant. Firms in countries with higher GDP per capita use more debt and hold larger ownership shares. If we interpret GDP per capita as a measure of the profitability of the foreign market these coefficients seem intuitive.

The fact that a firm's ownership and leverage choices might be related to each other can be captured by a correlation of the error terms of the two regressions. The two equations therefore form a system of seemingly unrelated regressions (Zellner (1962)). In general, it is more efficient to estimate this system of equations using feasible GLS and to allow for correlation of the error terms in the asymptotic variance matrix than to use OLS. In our

[22] Javorcik and Wei (2009) show that corruption on the country level, measured by three different indices, reduces the probability of whole ownership of a company. Asiedu and Esfahani (2001) show that the risk of expropriation decreases the probability of whole ownership. Both papers use cross-sectional data.

case, however, there is no efficiency gain from estimating the equations jointly, because the same regressors are included in both regressions (Cameron and Trivedi (2005)). We test for correlation of the estimated residuals and we find a positive and significant correlation in the residuals of the two regressions. This implies that a positive shock on one of the two variables is also associated with an increase in the other variable.

Desai *et al*. (2004) and Desai *et al*. (2008) assume a linear effect of political risk on the capital structure choice. By contrast, according to our model the influence of political risk differs by type or strength of political risk, as noted in our hypotheses 1 to 4. To study this effect in more detail, we follow two empirical strategies. We first continue to work with the aggregated ICRG index for political risk, but include political risk squared, to allow for the the marginal effect of political risk to change with its level. This corresponds to our interpretation above, that the different scenarios or political risk can be ranked with respect to their severity, with scenario III being the less severe and scenarios I/II the more severe. Our second strategy is to use alternative disaggregated measures for political risk, with the idea to capture more specifically one particular scenario.

Table 2.4 presents our results for the ICRG political risk index, including political risk squared. In both regressions, the estimated marginal effect of political risk on the variable in question is positive for low levels of political risk and negative for high levels of political risk. Both results are consistent with the predictions of hypotheses 3 and 4. The big difference between leverage and the ownership share is in the level of political risk where the marginal effect changes from positive to negative: for leverage, a maximum is reached at a level of political risk of about 0.3. The vast majority (about 91 percent) of affiliates of German MNE operate in countries where the political risk is weaker than this - thus, for them, the predicted effect is positive. Only for about 9 percent of all affiliates, we predict the effect of political risk on leverage to be negative. When we model leverage as a linear function of political risk, the positive effect that we predict for the majority of affiliates of German MNEs, prevails, as seen in Table 2.3. For the ownership share, the change from a positive to a negative effect of political risk is predicted to take place for a much lower level of political risk (about 0.13). This is in fact consistent with our hypothesis 5.

As outlined above, as a second strategy to capture different scenarios of political risk, we try different measures of political risk. We use two measures taken from the ICRG investment risk component (contract risk and repatriation risk) and three measures from the Heritage Index (corruption, investment risk and property rights risk). Data on contract risk and repatriation risk have only been available since 2001. *Contract risk* is defined as the risk of unilateral contract modification or cancelation and, at worst, outright expropriation of foreign owned assets. *Repatriation risk* captures to what extent profits can be transferred out of the host country. Impediments include exchange controls, excessive bureaucracy and a poor banking system. *Corruption* is defined as failure of the integrity of the system. *Investment risk* measures the degree of restrictions on foreign investment, considering a country's policies towards foreign investment, as well as its policies towards internal capital flows. *Property rights risk* measures the lack of freedom to accumulate private property as well as the risk to be expropriated.[23]

All measures have been recoded in such a way that first, higher values are associated with higher risk and, second, they lie between 0 and 1. Table 2.5 shows that, as in Table 2.3, the influence of political risk on the ownership share is negative, for all types of political risk measures considered. Interestingly, however, political risk can have either a positive or a negative influence on the level of leverage, depending on the type of risk present (Table 2.6). It is positive for corruption, investment risk and property right risk. According to our interpretation, this suggests that these three types of risk reduce the profitability of the investment without significantly increasing the risk of bankruptcy. In contrast, the coefficient is negative for contract risk, which seems to affect the chances to generate revenues and hence increases the risk of bankruptcy. Surprisingly, it is also negative for repatriation risk, even though this kind of political risk is more about the use of profits, i.e., scenario 3, for which a positive coefficient is predicted. If, however, credits are taken at home, not locally, the negative coefficient would make sense, because barriers to repatriation profits would undermine the ability to repay credits.

[23] For a more detailed description of the methodology underlying the Heritage Index see Beach and Kane (2008).

Table 2.4: The impact of political risk on affiliate leverage and ownership share (Allowing for nonlinear impact)

	(1)	(2)	(3)	(4)
Dependent variable	Leverage	Leverage	Ownership	Ownership
Political risk	0.492***	0.562***	0.167***	0.173***
	(14.75)	(15.35)	(6.11)	(5.77)
Political risk squared	-0.783***	-0.939***	-0.636***	-0.645***
	(12.23)	(12.30)	(11.51)	(9.72)
Log(sales)	0.011***	0.012***	-0.006***	-0.005***
	(14.54)	(14.14)	(9.18)	(7.60)
Profit/ total assets	-0.422***	-0.405***	-0.012***	-0.008**
	(31.85)	(29.18)	(3.13)	(2.09)
Fixed/ total assets	-0.033***	-0.021***	-0.023***	-0.028***
	(7.02)	(4.24)	(6.43)	(7.57)
Private credit	-0.026***	-0.026***	-0.027***	-0.027***
	(13.55)	(13.19)	(17.24)	(16.00)
Inflation	0.000	0.001	0.001	0.005***
	(0.56)	(0.3)	(1.56)	(4.01)
Log(GDP)	0.002***	-0.001	-0.005***	-0.002**
	(2.89)	(0.82)	(8.20)	(2.47)
Log(GDP per capita)	0.004***	0.001	0.024***	0.024***
	(2.81)	(0.68)	(19.13)	(16.86)
Statutory tax		0.208***		-0.143***
		(14.86)		(13.25)
Constant	0.283***	0.233***	0.719***	0.725***
	(11.08)	(8.19)	(33.77)	(30.91)
Observations	142,325	127,096	142,309	127,084
Number of parents	11,666	11,315	11,666	11,315
R-squared	0.12	0.12	0.04	0.04

OLS regression including parent fixed effects
Year and affiliate industry dummies included
Heteroscedasticity robust t-statistics (absolute values) in parentheses
Clustering of standard errors by parent company
*** p<0.01, ** p<0.05, * p<0.1

Data sources: Firm-level variables are taken from the Microdatabase Direct Investment of the Deutsche Bundesbank. Private credit is provided in Beck et al. (1999). Inflation, log of GDP, log of GDP per capita are taken from the IMF. Statutory tax rates are taken from the IFS, as well as from the Corporate Tax Guides of Ernst&Young, KPMG and PricewaterhouseCoopers.

Table 2.5: The impact of political risk (different measures) on ownership share

	(1) Ownership	(2) Ownership	(3) Ownership	(4) Owernship	(5) Ownerhship	(6) Ownership	(7) Ownership	(8) Ownerhship	(9) Ownership	(10) Ownership
Contract risk	-0.096***	-0.101***								
	(8.45)	(7.69)								
Corruption			-0.033***	-0.028***						
			(6.95)	(5.32)						
Investment risk					-0.044***	-0.044***				
					(8.11)	(7.40)				
Property rights risk							-0.033***	-0.040***		
							(5.34)	(5.64)		
Repatriation risk									-0.005	-0.029**
									(0.55)	(2.50)
Log(sales)	-0.005***	-0.005***	-0.006***	-0.005***	-0.006***	-0.005***	-0.006***	-0.005***	-0.005***	-0.005***
	(5.41)	(4.66)	(9.00)	(7.34)	(9.16)	(7.45)	(9.15)	(7.49)	(5.07)	(4.36)
Profit/ total assets	-0.010	-0.01	-0.013***	-0.009**	-0.012***	-0.009**	-0.012***	-0.009**	-0.011*	-0.010
	(1.6)	(1.51)	(3.27)	(2.24)	(3.11)	(2.15)	(3.19)	(2.19)	(1.69)	(1.58)
Fixed/ total assets	-0.022***	-0.025***	-0.021***	-0.027***	-0.021***	-0.027***	-0.022***	-0.028***	-0.021***	-0.025***
	(4.21)	(4.66)	(6.04)	(7.24)	(5.92)	(7.20)	(6.09)	(7.39)	(4.00)	(4.56)
Private credit	-0.027***	-0.028***	-0.031***	-0.031***	-0.027***	-0.028***	-0.028***	-0.030***	-0.029***	-0.031***
	(10.76)	(10.78)	(18.39)	(18.11)	(17.49)	(17.66)	(18.00)	(18.30)	(11.79)	(12.17)
Inflation	0.001	-0.002	-0.001*	0.000	-0.001**	-0.001	-0.001*	-0.000	-0.004**	-0.006**
	(0.28)	(0.72)	(1.70)	(0.01)	(2.09)	(0.56)	(1.94)	(0.16)	(1.96)	(2.57)
Log(GDP)	-0.005***	-0.003***	-0.005***	-0.003***	-0.005***	-0.002**	-0.006***	-0.003***	-0.005***	-0.002**
	(5.93)	(3.17)	(10.04)	(4.23)	(8.07)	(2.48)	(10.24)	(3.93)	(6.19)	(2.50)
Log(GDP per capita)	0.022***	0.022***	0.030***	0.029***	0.031***	0.030***	0.030***	0.028***	0.028***	0.025***
	(14.66)	(13.19)	(25.54)	(22.00)	(31.42)	(27.47)	(24.08)	(19.33)	(16.32)	(13.36)
Statutory tax		-0.118***		-0.139***		-0.148***		-0.141***		-0.134***
		(7.43)		(12.75)		(13.84)		(13.03)		(8.50)
Constant	0.759***	0.754***	0.683***	0.693***	0.667***	0.686***	0.678***	0.708***	0.705***	0.724***
	(25.72)	(22.97)	(33.40)	(30.62)	(34.59)	(32.17)	(32.20)	(30.14)	(23.25)	(21.31)
Observations	73,320	66,286	142,793	127,403	142,793	127,403	142,793	127,403	73,320	66,286
Number of parents	8,646	8,447	11,691	11,332	11,691	11,332	11,691	11,332	8,646	8,447
Years	2001-2006	2001-2006	1996-2006	1996-2006	1996-2006	1996-2006	1996-2006	1996-2006	2001-2006	2001-2006
R-squared	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.03	0.04

Data sources: Corruption, investment risk, property rights risk taken from the Heritage Foundation. Contract risk and repatriation risk provided by PRS. All other variables see Tables 3, 4.

Table 2.6: The impact of political risk (different measures) on affiliate leverage

	(1) Leverage	(2) Leverage	(3) Leverage	(4) Leverage	(5) Leverage	(6) Leverage	(7) Leverage	(8) Leverage	(9) Leverage	(10) Leverage
Contract risk	-0.040***	-0.070***								
	(2.94)	(4.36)								
Corruption			0.080***	0.089***						
			(14.41)	(14.47)						
Investment risk					0.004	0.028***				
					(0.66)	(3.82)				
Property rights risk							0.063***	0.081***		
							(8.87)	(9.88)		
Repatriation risk									-0.051***	-0.043***
									(4.58)	(3.09)
Log(sales)	0.013***	0.013***	0.011***	0.012***	0.011***	0.012***	0.011***	0.012***	0.013***	0.013***
	(11.73)	(10.71)	(14.52)	(14.19)	(14.60)	(14.30)	(14.76)	(14.46)	(11.74)	(10.82)
Profit/ total assets	-0.414***	-0.392***	-0.422***	-0.405***	-0.423***	-0.407***	-0.423***	-0.406***	-0.414***	-0.392***
	(21.07)	(19.04)	(31.93)	(29.22)	(31.93)	(29.25)	(31.95)	(29.25)	(21.07)	(19.04)
Fixed/ total assets	-0.010	-0.002	-0.035***	-0.022***	-0.037***	-0.024***	-0.035***	-0.022***	-0.010	-0.001
	(1.52)	(0.24)	(7.50)	(4.55)	(7.88)	(4.91)	(7.48)	(4.43)	(1.50)	(0.20)
Private credit	-0.028***	-0.030***	-0.022***	-0.023***	-0.031***	-0.033***	-0.028***	-0.030***	-0.029***	-0.031***
	(9.11)	(9.22)	(11.05)	(11.18)	(16.40)	(16.62)	(14.55)	(14.85)	(9.35)	(9.88)
Inflation	0.004*	0.006**	-0.000	-0.003*	-0.000	-0.003*	-0.000	-0.002*	0.004*	0.0042
	(1.78)	(2.39)	(0.63)	(1.88)	(0.37)	(1.77)	(0.11)	(1.69)	(1.79)	(1.62)
Log(GDP)	0.006***	0.004***	0.003***	0.001	0.004***	0.001	0.003***	0.001	0.007***	0.004***
	(6.67)	(3.23)	(4.33)	(1.4)	(5.79)	(0.65)	(5.08)	(0.98)	(7.31)	(3.83)
Log(GDP per capita)	-0.006***	-0.010***	0.009***	0.007***	-0.003**	-0.004***	0.005***	0.004***	-0.009***	-0.010***
	(3.60)	(5.54)	(6.84)	(4.75)	(2.42)	(3.64)	(3.91)	(2.83)	(4.70)	(4.89)
Statutory tax		0.201***		0.181***		0.218***		0.200***		0.190***
		(9.84)		(12.76)		(15.64)		(14.26)		(9.39)
Constant	0.330***	0.316***	0.276***	0.228***	0.403***	0.347***	0.314***	0.250***	0.354***	0.316***
	(8.51)	(6.99)	(11.37)	(8.44)	(17.40)	(13.40)	(12.61)	(8.88)	(9.03)	(6.85)
Observations	73,332	66,296	142,809	127,415	142,809	127,415	142,809	127,415	73,332	66,296
Number of parents	8,646	8,447	11,691	11,332	11,691	11,332	11,691	11,332	8,646	8,447
Years	2001-2006	2001-2006	1996-2006	1996-2006	1996-2006	1996-2006	1996-2006	1996-2006	2001-2006	2001-2006
R-squared	0.12	0.12	0.12	0.12	0.11	0.11	0.11	0.11	0.12	0.12

Data sources: Corruption, investment risk, property rights risk taken from the Heritage Foundation. Contract risk and repatriation risk provided by PRS. All other variables see Tables 3, 4.

2.7 Conclusion

In this chapter, we have investigated how MNEs adapt their capital structure choices in the presence of political risk. Our analysis suggests that, when it comes to assessing the potential effects of political risk, it is important to distinguish different types of stakeholders and how these are affected by different political measures. Only then is it possible to determine the optimal reaction of the investor to this risky environment.

Almost by definition, any form of political risk negatively affects the profitability of a MNE as a whole. The investor as one of the equity holders is residual claimant of the MNE. Thus, it is not surprising that he will want to reduce his stake in the firm when political risk becomes more severe.

But, as our analysis has shown, the effects are less straightforward for debt holders. If debt holders expect debt service to become less likely, they have to adjust their interest rates. The larger the additional dead weight losses, the more costly debt financing becomes and hence the more leverage is reduced. If, instead, debt service is less affected than the returns to equity holders, then debt can act as a shield against political risk and the balance shifts towards more debt finance relative to equity finance.

Why does it matter how political risk affects the multinational's choice of capital structure? Smaller ownership shares, for example, may negatively affect the governance structure of the MNE because they typically lead to smaller incentives for controlling the firm effectively. In addition, smaller ownership stakes could reduce the investor's incentive to transfer necessary technology. These effects, though not explicitly modeled here, are well known in the literature. Higher leverage, and hence higher risk of bankruptcy, on the other hand, lead to higher dead weight losses arising from inefficient bankruptcy procedures and, in this way, add to the social cost of political risk.

Thus, we would expect different kinds of dead weight losses to prevail, depending on how political risk affects equity holders relative to debt holders. If ownership shares are reduced and leverage increases, one may have to face deteriorating governance structures and more inefficiencies from increased risk of bankruptcy. If instead both ownership shares and leverage are reduced, it is mostly deteriorated governance structures one has to expect.

Mathematical Appendix

Proof of Result 1

Recall that the interest rate is implicitly determined by the following break even condition for the bank:

$$\int_{(1+r)D}^{\bar{R}} (1+r)D\frac{1}{\bar{R}}dR + \int_{0}^{(1+r)D} sR\frac{1}{\bar{R}}dR = D \tag{2.34}$$

Solving and rearranging yields

$$\begin{aligned} \frac{1}{\bar{R}}\left[(1+r)D\bar{R} - \frac{2-s}{2}(1+r)^2D^2\right] &= D && (2.35)\\ \frac{1}{\bar{R}}\frac{1}{2}(1+r)^2D^2 &= \frac{(1+r)D}{2-s} - \frac{D}{2-s} && (2.36)\\ (2-s)(1+r)^2D &= 2r\bar{R} && (2.37)\end{aligned}$$

From the implicit function

$$2r\bar{R} - (2-s)(1+r)^2D = 0 \tag{2.38}$$

we can derive how the interest rate reacts for a change in the debt level D, using the implicit function theorem.

$$\frac{dr}{dD} = -\frac{-(2-s)(1+r)^2}{2[\bar{R} - (2-s)(1+r)D]} \tag{2.39}$$

Using (2.38), this simplifies to

$$\frac{dr}{dD} = \frac{r(1+r)}{(1-r)D} \tag{2.40}$$

Furthermore, we can use (2.38) to derive how the interest rate reacts to a change in $\bar{R}$, again using the implicit function theorem.

$$\frac{dr}{d\bar{R}} = -\frac{-2r}{2[\bar{R} - (2-s)(1+r)D]} \tag{2.41}$$

Using (2.38), this simplifies to

$$\frac{dr}{d\bar{R}} = -\frac{r(1+r)}{(1-r)\bar{R}} \tag{2.42}$$

The investor maximizes

$$\begin{aligned} U(D,\alpha) &= V(D,\alpha) + B(\alpha) - K(\alpha) && (2.43)\\ &= \int_{(1+r)D}^{\bar{R}(\alpha)} (1-t)[R-(1+r)D]\frac{1}{\bar{R}(\alpha)} dR + D - I + B(\alpha) - K(\alpha) && (2.44)\\ &= \frac{1-t}{\bar{R}(\alpha)}\left[\frac{1}{2}\bar{R}^2(\alpha) - (1+r)D\bar{R}(\alpha) + \frac{1}{2}(1+r)^2 D^2\right] && (2.45)\\ &\quad +D - I + B(\alpha) - K(\alpha) \end{aligned}$$

Using equation (2.34) we can rewrite the payoff function in the following way

$$U(D,\alpha) = (1-t)\left[\frac{1}{2}\bar{R}(\alpha) - \frac{(1-s)(1+r)}{2-s}D\right] + \frac{1-s+t}{2-s}D - I + B(\alpha) - K(\alpha). \quad (2.46)$$

The investor's maximization problem is characterized by the following two first order conditions.

$$\begin{aligned} \frac{dU}{d\alpha} &= (1-t)\left[\frac{1}{2}\frac{d\bar{R}}{d\alpha} - \frac{1-s}{2-s}D\frac{dr}{d\bar{R}}\frac{d\bar{R}}{d\alpha}\right] + B' - K' = 0 && (2.47)\\ \frac{dU}{dD} &= -(1-t)\frac{1-s}{2-s}\left[(1+r) + \frac{dr}{dD}D\right] + \frac{1-s+t}{2-s} = 0 && (2.48) \end{aligned}$$

Using (2.40), we can rewrite the first order condition for the optimal D in the following way

$$\frac{dU}{dD} = -(1-t)\frac{1-s}{2-s}\left[1 + r + \frac{r(1+r)}{(1-r)D}D\right] + \frac{1-s+t}{2-s} = 0 \quad (2.49)$$

Rearranging yields

$$\frac{1-s+t}{2-s} = (1-t)\frac{1-s}{2-s}\frac{1+r}{1-r} \quad (2.50)$$

We can use this condition to solve for r and $1+r$:

$$r = \frac{(2-s)t}{2(1-s)+st} \qquad 1+r = \frac{2(1-s+t)}{2(1-s)+st} \quad (2.51)$$

Inserting this in (2.38), we can solve for

$$D^* = \frac{\bar{R}}{2}t\frac{2(1-s)+st}{(1-s+t)^2} \quad (2.52)$$

Using the solutions for r and D, we can finally determine the investor's payoff as a function of α

$$U = \frac{\bar{R}(\alpha)}{2}\frac{1-s+st}{1-s+t} - I + B(\alpha) - K(\alpha) \quad (2.53)$$

Q.E.D

Proof of Result 2

To see how the optimal debt level reacts to changes in t, consider the optimal debt level as characterized by (2.52):

$$\frac{dD^*}{dt} = \frac{\bar{R}}{2}\left[\frac{2(1-s)+st}{(1-s+t)^2} + t\frac{(1-s+t)^2 s - (2(1-s)+st)2(1-s+t)}{(1-s+t)^4}\right] \quad (2.54)$$

$$= \frac{\bar{R}(1-s)^2(1-t)}{(1-s+t)^3} > 0 \quad (2.55)$$

To determine how the ownership ratio α reacts to changes in t, we rewrite the first order condition (2.47) that implicitly defines α^*, using (2.42) and the solution to D^* and r. After some simplification we obtain

$$(1-t)\left[1 + \frac{t^2}{(1-s+t)(1-t)}\right]\frac{1}{2}\frac{d\bar{R}}{d\alpha} + B' - K' = 0 \quad (2.56)$$

$$\left[\frac{1-s+st}{1-s+t}\right]\frac{1}{2}\frac{d\bar{R}}{d\alpha} + B' - K' = 0 \quad (2.57)$$

From this we find, using the implicit function theorem

$$\frac{d\alpha^*}{dt} = -\frac{\left[\frac{(1-s+t)s-(1-s+st)}{(1-s+t)^2}\right]\frac{1}{2}\frac{d\bar{R}}{d\alpha}}{\left[\frac{1-s+st}{(1-s+t)}\right]\frac{1}{2}\frac{d^2\bar{R}}{d\alpha^2} + B'' - K''} \quad (2.58)$$

$$(2.59)$$

$$= \frac{\left[\frac{(1-s)^2}{(1-s+t)^2}\right]\frac{1}{2}\frac{d\bar{R}}{d\alpha}}{\underbrace{\left[\frac{1-s+st}{1-s+t}\right]}_{+}\frac{1}{2}\underbrace{\frac{d^2\bar{R}}{d\alpha^2}}_{-} + \underbrace{B''}_{-} - \underbrace{K''}_{+}} \quad (2.60)$$

Note that the sign of $\frac{d\alpha}{dt}$ depends on the sign of $\frac{d\bar{R}}{d\alpha}$. Using equation (2.57), we find that $R' > 0$ if $B' < K'$ in the relevant parameter range, and hence $\frac{d\alpha}{dt} < 0$ if $B' < K'$.

Q.E.D.

Proof of Result 3

Consider first the case of expropriation. Recall that the interest rate is implicitly determined by the following break even condition for the bank:

$$(1-\pi_1)\left[\int_{(1+r)D}^{\bar{R}} (1+r)D\frac{1}{\bar{R}}dR + \int_0^{(1+r)D} sR\frac{1}{\bar{R}}dR\right] = D \quad (2.61)$$

Solving and rearranging yields

$$\frac{1-\pi_1}{\bar{R}}\left[(1+r)D\bar{R}-\frac{2-s}{2}(1+r)^2D^2\right] = D \tag{2.62}$$

$$\frac{1-\pi_1}{\bar{R}}\frac{1}{2}(1+r)^2D^2 = \frac{(1-\pi_1)(1+r)D}{2-s}-\frac{D}{2-s} \tag{2.63}$$

$$(1-\pi_1)(2-s)(1+r)^2D = 2(1-\pi_1)(1+r)\bar{R}-2\bar{R} \tag{2.64}$$

From the implicit function

$$2[(1-\pi_1)(1+r)-1]\bar{R}-(1-\pi_1)(2-s)(1+r)^2D=0 \tag{2.65}$$

we can derive how the interest rate reacts to a change in the debt level D

$$\frac{dr}{dD}=-\frac{-(1-\pi_1)(2-s)(1+r)^2}{2[(1-\pi_1)\bar{R}-(1-\pi_1)(2-s)(1+r)D]} \tag{2.66}$$

Using (2.65), this simplifies to

$$\frac{dr}{dD}=\frac{(1+r)[(1-\pi_1)(1+r)-1]}{[2-(1-\pi_1)(1-r)]D} \tag{2.67}$$

Furthermore, we can derive how the interest rate reacts to a change in $\bar{R}$

$$\frac{dr}{d\bar{R}}=-\frac{2[(1-\pi_1)(1+r)-1]}{2(1-\pi_1)\bar{R}-2(1-\pi_1)(2-s)(1+r)D} \tag{2.68}$$

Using again (2.65), this simplifies to

$$\frac{dr}{\bar{R}}=-\frac{2[(1-\pi_1)(1+r)-1](1+r)}{2-(1-\pi_1)(1+r)\bar{R}}>0 \tag{2.69}$$

The investor maximizes

$$\begin{aligned} U_1(D,\alpha) &= (1-\pi_1)\int_{(1+r)D}^{\bar{R}(\alpha)}(1-t)[R-(1+r)D]\frac{1}{\bar{R}(\alpha)}dR && (2.70)\\ &\quad +D-I+(1-\pi_1)B(\alpha)-K(\alpha) \\ &= \frac{(1-\pi_1)(1-t)}{\bar{R}(\alpha)}\left[\frac{1}{2}\bar{R}^2(\alpha)-(1+r)D\bar{R}(\alpha)+\frac{1}{2}(1+r)^2D^2\right] && (2.71)\\ &\quad +D-I+(1-\pi_1)B(\alpha)-K(\alpha) \end{aligned}$$

Using equation (2.61) we can rewrite the payoff function in the following way

$$\begin{aligned} U_1(D,\alpha) &= (1-\pi_1)(1-t)\left[\frac{1}{2}\bar{R}(\alpha)-\frac{(1-s)(1+r)}{2-s}D\right] \\ &\quad +\frac{1-s+t}{2-s}D-I+(1-\pi_1)B(\alpha)-K(\alpha)\ . && (2.72) \end{aligned}$$

The investor's maximization problem is characterized by the following two first order conditions.

$$\frac{dU_1}{d\alpha} = (1-\pi_1)(1-t)\left[\frac{1}{2}\frac{d\bar{R}}{d\alpha} - \frac{1-s}{2-s}D\frac{dr}{d\bar{R}}\frac{d\bar{R}}{d\alpha}\right] + (1-\pi_1)B' - K' = 0 \tag{2.73}$$

$$\frac{dU_1}{dD} = -(1-\pi_1)(1-t)\frac{1-s}{2-s}\left[(1+r) + \frac{dr}{dD}D\right] + \frac{1-s+t}{2-s} = 0 \tag{2.74}$$

Using (2.67), we can rewrite the first order condition for the optimal D in the following way

$$\frac{dU_1}{dD} = -(1-\pi)(1-t)\frac{1-s}{2-s}\left[1+r+\frac{(1+r)[(1-\pi_1)(1+r)-1]}{[2-(1-\pi_1)(1+r)D}D\right] + \frac{1-s+t}{2-s} = 0 \tag{2.75}$$

Rearranging yields

$$\frac{1-s+t}{2-s} = (1-\pi_1)(1-t)\frac{1-s}{2-s}\frac{1+r}{2-(1-\pi_1)(1+r)} \tag{2.76}$$

We can use this condition to solve for r and $1+r$:

$$r_1 = \frac{(2-s)t}{(1-\pi_1)[2(1-s)+st]} \qquad 1+r_1 = \frac{2(1-s+t)}{(1-\pi_1)[2(1-s)+st]} \tag{2.77}$$

Inserting this in (2.65), we can solve for

$$D_1^* = (1-\pi_1)\frac{\bar{R}}{2}t\frac{2(1-s)+st}{(1-s+t)^2} \tag{2.78}$$

Using the solutions for r and D_1, we can finally determine the investor's payoff

$$U_1 = (1-\pi_1)\left[\frac{\bar{R}}{2}\frac{1-s+st}{1-s+t} + B(\alpha)\right] - K(\alpha) - I \tag{2.79}$$

Creeping expropriation

Consider now the case of creeping expropriation. Recall that the interest rate is implicitly determined by the following break even condition for the bank:

$$\left[\int_{\frac{(1+r)D}{(1-\pi_2)}}^{\bar{R}} (1+r)D\frac{1}{\bar{R}}dR + \int_0^{\frac{(1+r)D}{(1-\pi_2)}} s(1-\pi_2)R\frac{1}{\bar{R}}dR\right] = D \tag{2.80}$$

Solving and rearranging yields

$$\frac{1}{(1-\pi_2)\bar{R}}\left[(1+r)D(1-\pi_2)\bar{R}-\frac{2-s}{2}(1+r)^2D^2\right] = D \tag{2.81}$$

$$\frac{1}{(1-\pi_2)\bar{R}}\frac{1}{2}(1+r)^2D^2 = \frac{(1+r)D}{2-s}-\frac{D}{2-s} \tag{2.82}$$

$$(2-s)(1+r)^2D = 2(1-\pi_2)(1+r)\bar{R}-2(1-\pi_2)\bar{R}$$

$$(2-s)(1+r)^2D = 2r(1-\pi_2)\bar{R} \tag{2.83}$$

From the implicit function

$$2r(1-\pi_2)\bar{R}-(2-s)(1+r)^2D=0 \tag{2.84}$$

we can derive how the interest rate reacts for a change in the debt level D

$$\frac{dr}{dD}=-\frac{-(2-s)(1+r)^2}{2[(1-\pi_2)\bar{R}-(2-s)(1+r)D]} \tag{2.85}$$

Using (2.84), this simplifies to

$$\frac{dr}{dD}=\frac{r(1+r)}{(1-r)D} \tag{2.86}$$

Furthermore, we can derive how the interest rate reacts to a change in $\bar{R}$

$$\frac{dr}{d\bar{R}}=-\frac{2r(1-\pi_2)}{2(1-\pi_2)\bar{R}-2(2-s)(1+r)D} \tag{2.87}$$

Using again (2.84), this simplifies to

$$\frac{dr}{\bar{R}}=-\frac{r(1+r)}{(1+r)\bar{R}}>0 \tag{2.88}$$

The investor maximizes

$$U_2(D,\alpha) = \int_{\frac{(1+r)D}{(1-\pi_2)}}^{\bar{R}(\alpha)}(1-t)[(1-\pi_2)R-(1+r)D]\frac{1}{\bar{R}(\alpha)}dR + D - I + (1-\pi_2)B(\alpha)-K(\alpha) \tag{2.89}$$

$$= \frac{(1-t)}{(1-\pi_2)\bar{R}(\alpha)}\left[\frac{1}{2}(1-\pi_2)^2\bar{R}^2(\alpha)-(1+r)D(1-\pi_2)\bar{R}(\alpha) + \frac{1}{2}(1+r)^2D^2\right]+D-I+(1-\pi_2)B(\alpha)-K(\alpha) \tag{2.90}$$

Using equation (2.84) we can rewrite the payoff function in the following way

$$\begin{aligned} U_2(D,\alpha) &= (1-t)\left[\frac{1}{2}(1-\pi_2)\bar{R}(\alpha) - \frac{(1-s)(1+r)}{2-s}D\right] \\ &+\frac{1-s+t}{2-s}D - I + (1-\pi_2)B(\alpha) - K(\alpha)\ . \end{aligned} \tag{2.91}$$

The investor's maximization problem is characterized by the following two first order conditions.

$$\frac{dU}{d\alpha} = (1-t)\left[\frac{1}{2}(1-\pi_2)\frac{d\bar{R}}{d\alpha} - \frac{1-s}{2-s}D\frac{dr}{d\bar{R}}\frac{d\bar{R}}{d\alpha}\right] + (1-\pi_2)B' - K' = 0 \tag{2.92}$$

$$\frac{dU}{dD} = -(1-t)\frac{1-s}{2-s}\left[(1+r) + \frac{dr}{dD}D\right] + \frac{1-s+t}{2-s} = 0 \tag{2.93}$$

Using (2.86), we can rewrite the first order condition for the optimal D in the following way

$$\frac{dU}{dD} = -(1-\pi)(1-t)\frac{1-s}{2-s}\left[1+r+\frac{r(1+r)}{(1-r)D}D\right] + \frac{1-s+t}{2-s} = 0 \tag{2.94}$$

Rearranging yields

$$\frac{1-s+t}{2-s} = (1-t)\frac{1-s}{2-s}\frac{1+r}{(1-r)} \tag{2.95}$$

We can use this condition to solve for r_2 and $1+r_2$:

$$r_2 = \frac{(2-s)t}{2(1-s)+st} \qquad 1+r_2 = \frac{2(1-s+t)}{2(1-s)+st} \tag{2.96}$$

Inserting this in (2.84), we can solve for

$$D_2 = (1-\pi_2)\frac{\bar{R}}{2}t\frac{2(1-s)+st}{(1-s+t)^2} = D_1 \tag{2.97}$$

Using the solutions for r_2 and D_2, we can finally determine the investor's payoff

$$U_2 = (1-\pi_2)\left[\frac{\bar{R}}{2}\frac{1-s+st}{1-s+t} + B(\alpha)\right] - K(\alpha) - I \tag{2.98}$$

Q.E.D.

Proof of Result 4

In Result 3 we have seen that the optimal debt levels and the investor's payoff are the same in both cases, expropriation and creeping expropriation. We now determine the comparative statics with respect to the local taxation rate π_i, with $i = \{1, 2\}$.

To see how the optimal debt level reacts to changes in π_1, consider the optimal debt level as characterized in (2.78) and (2.97).

$$\frac{dD_i^*}{d\pi_i} = -\frac{1}{2}\bar{R}t\frac{(2(1-s)+st)}{(1-s+t)^2} < 0 \tag{2.99}$$

To determine how the ownership ratio α reacts to changes in π_i, we use the first order condition of (2.79) or (2.98) that implicitly defines α^*

$$(1-\pi_i)\left[\left[\frac{1-s+st}{1-s+t}\right]\frac{1}{2}\frac{d\bar{R}}{d\alpha}+B'\right]-K' = 0 \tag{2.100}$$

From this we find, using the implicit function theorem

$$\frac{d\alpha^*}{d\pi_i} = -\frac{\overbrace{-\frac{(1-s+st)}{(1-s+t)}\frac{1}{2}\frac{d\bar{R}}{d\alpha}-B'}^{-}}{(1-\pi_i\frac{1-s+st}{(1-s+t)}\frac{1}{2}\underbrace{\frac{d^2\bar{R}}{d\alpha^2}}_{-}+(1-\pi_i)\underbrace{B''}_{-}-\underbrace{K''}_{+}} < 0 \tag{2.101}$$

where the negative sign of the nominator is due to the fact that the first order condition (2.100) needs to be satisfied.

Q.E.D.

Proof of Result 5

Recall that in case of confiscatory taxation the interest rate is implicitly determined by the same break even condition for the bank as in the base line model:

$$\int_{(1+r)D}^{\bar{R}}(1+r)D\frac{1}{\bar{R}}dR+\int_0^{(1+r)D}sR\frac{1}{\bar{R}}dR = D \tag{2.102}$$

This implies the same implicit function and hence the same conditions for the interest rate as above.

$$2r\bar{R}-(2-s)(1+r)^2D = 0 \tag{2.103}$$

$$\frac{dr}{dD} = \frac{r(1+r)}{(1-r)D} \tag{2.104}$$

and

$$\frac{dr}{d\bar{R}} = -\frac{r(1+r)}{(1-r)\bar{R}} \tag{2.105}$$

The investor maximizes

$$
\begin{aligned}
U_3(D,\alpha) &= \int_{(1+r)D}^{\bar{R}(\alpha)} (1-t-\pi_3)[R-(1+r)D]\frac{1}{\bar{R}(\alpha)}dR \\
&\quad +D-I+B(\alpha)-K(\alpha) &(2.106)\\
&= \frac{1-t-\pi_3}{\bar{R}(\alpha)}\left[\frac{1}{2}\bar{R}^2(\alpha)-(1+r)D\bar{R}(\alpha)+\frac{1}{2}(1+r)^2D^2\right] \\
&\quad +D-I+B(\alpha)-K(\alpha) &(2.107)
\end{aligned}
$$

Using equation (2.103) we can rewrite the payoff function in the following way

$$
\begin{aligned}
U_3(D,\alpha) &= (1-t-\pi_3)\left[\frac{1}{2}\bar{R}^2(\alpha)-\frac{(1-s)(1+r)}{2-s}D\right] \\
&\quad +\frac{1-s+t+\pi_3}{2-s}D-I+B(\alpha)-K(\alpha)\ . &(2.108)
\end{aligned}
$$

The investor's maximization problem is characterized by the following two first order conditions.

$$
\frac{dU_3}{d\alpha} = (1-t-\pi_3)\left[\frac{1}{2}\frac{d\bar{R}}{d\alpha}-\frac{1-s}{2-s}D\frac{dr}{d\bar{R}}\frac{d\bar{R}}{d\alpha}\right]+B'-K'=0 \tag{2.109}
$$

$$
\frac{dU_3}{dD} = -(1-t-\pi_3)\frac{1-s}{2-s}\left[(1+r)+\frac{dr}{dD}D\right]+\frac{1-s+t+\pi_3}{2-s}=0 \tag{2.110}
$$

Using (2.104), we can rewrite the first order condition for the optimal D in the following way

$$
\frac{dU_3}{dD}=-(1-t-\pi_3)\frac{1-s}{2-s}\left[1+r+\frac{r(1+r)}{(1-r)D}D\right]+\frac{1-s+t+\pi_3}{2-s}=0 \tag{2.111}
$$

Rearranging yields

$$
\frac{1-s+t+\pi_3}{2-s}=(1-t-\pi_3)\frac{1-s}{2-s}\frac{1+r}{1-r} \tag{2.112}
$$

We can use this condition to solve for r_3 and $1+r_3$:

$$
r_3=\frac{(2-s)(t+\pi_3)}{2(1-s)+s(t+\pi_3)} \qquad 1+r_3=\frac{2(1-s+t+\pi_3)}{2(1-s)+s(t+\pi_3)} \tag{2.113}
$$

Inserting this in (2.103), we can solve for

$$
D_3^*=\frac{\bar{R}}{2}(t+\pi_3)\frac{2(1-s)+s(t+\pi_3)}{(1-s+t+\pi_3)^2} \tag{2.114}
$$

Using the solutions for r_3 and D_3, we can finally determine the investor's payoff

$$U_3 = \frac{\bar{R}}{2}\frac{1-s+s(t+\pi_3)}{1-s+t+\pi_3} - I + (1-\pi_3)B(\alpha) - K(\alpha) \tag{2.115}$$

Q.E.D.

Proof of Result 6

To see how the optimal debt level reacts to changes in π_3, consider the optimal debt level as characterized by (2.114):

$$\begin{aligned}\frac{dD_3^*}{d\pi_3} &= \frac{\bar{R}}{2}\left[\frac{2(1-s)+s(t+\pi_3)}{(1-s+t+\pi_3)^2}\right.\\ &\quad \left.+(t+\pi_3)\frac{(1-s+t+\pi_3)^2 s - 2(2(1-s)+s(t+\pi_3))(1-s+t+\pi_3)}{(1-s+t+\pi_3)^4}\right]\\ &= \frac{\bar{R}(1-s)^2}{(1-s+t+\pi_3)^3} > 0 \end{aligned}\tag{2.116}$$

To determine how the ownership ratio α reacts to changes in t, consider the first order condition of (2.115) that implicitly defines α_3^*.

$$\left[\frac{1-s+s(t+\pi_3)}{1-s+t+\pi_3}\right]\frac{1}{2}\frac{d\bar{R}}{d\alpha} + (1-\pi_3)B' - K' \quad = \quad 0 \tag{2.117}$$

From this we find, using the implicit function theorem,

$$\frac{d\alpha_3^*}{d\pi_3} = -\frac{\left[-\frac{(1-s)^2}{(1-s+t+\pi_3)^2}\right]\frac{1}{2}\frac{d\bar{R}}{d\alpha}}{\underbrace{\left[\frac{1-s+s(t+\pi_3)}{1-s+t+\pi_3}\right]}_{+}\frac{1}{2}\underbrace{\frac{d^2\bar{R}}{d\alpha^2}}_{-}+(1-\pi_3)\underbrace{B''}_{-}-\underbrace{K''}_{+}} \tag{2.118}$$

Note that the sign of $\frac{d\alpha_3}{d\pi_3}$ depends on the sign of $\frac{d\bar{R}}{d\alpha}$. Using equation (2.117), we find that $R' > 0$ if $B' < K'$ in the relevant parameter range, and hence $\frac{d\alpha_3}{d\pi_3} < 0$ if $B' < K'$.

Q.E.D.

Proof of Result 7

Consider

$$\frac{d\alpha_{1/2}^*}{d\pi_{1/2}} = \frac{\frac{(1-s+st)}{(1-s+t)}\frac{1}{2}\frac{d\bar{R}}{d\alpha} + B'}{(1-\pi_{1/2})\left[\frac{1-s+st}{(1-s+t)}\frac{1}{2}\frac{d^2\bar{R}}{d\alpha^2} + B''\right] - K''} < 0 \tag{2.119}$$

and

$$\frac{d\alpha_3^*}{d\pi_3} \quad = \quad \frac{\left[\frac{(1-s)^2}{(1-s+t+\pi_3)^2}\right]\frac{1}{2}\frac{d\bar{R}}{d\alpha}}{\left[\frac{1-s+s(t+\pi_3)}{1-s+t+\pi_3}\right]\frac{1}{2}\frac{d^2\bar{R}}{d\alpha^2}+B''-K''} < 0 \ \text{ if } \ \bar{R}'>0 \tag{2.120}$$

To see that $\frac{d\alpha_{1/2}}{d\pi_{1/2}} < \frac{d\alpha_3}{d\pi_3}$ it is sufficient to show that nominator of $\left|\frac{d\alpha_{1/2}}{d\pi_{1/2}}\right|$ is larger than the nominator of$\left|\frac{d\alpha_3}{d\pi_3}\right|$ and the denominator is smaller than the respective denominator. Simple rearranging of the respective equations prove that this is indeed the case. *Q.E.D.*

Chapter 3

Financial constraints and the margins of FDI*

3.1 Motivation

Multinational firms are larger than their domestic counterparts. For European firms, Mayer and Ottaviano (2008) show that multinational firms are also more productive, generate higher value added, pay higher wages, employ more capital per worker, and employ a larger number of skilled workers. In the theoretical literature, the characteristic size patterns of multinational firms are explained mainly by differences in productivity. According to this explanation, observed internationalization patterns reflect real constraints since only the more productive firms can afford to shoulder the fixed cost of market entry.

These stylized facts are confirmed by our data for German companies, where firms owning foreign affiliates are indeed substantially larger than purely domestic firms (Figure 3.1(a)). Yet, the two groups of firms also differ in a number of other respects. Multinational firms, for instance, have lower debt ratios and higher cash flows. This suggests difficulties in obtaining external finance as an additional impediment to foreign expansions.[24] However, most of the theoretical literature considers the impact of financial constraints to be of lesser importance, arguing that foreign direct investment (FDI) and the associated financing decisions can largely be treated separately.[25]

The purpose of this chapter is to assess the (relative) importance of real and financial barriers for the cross-border expansion of firms. In doing so, we distinguish between the decision to enter a foreign market for the first time (the extensive margin) and the decision on the volume of foreign affiliate sales (the intensive margin). We proceed in two steps.

In a first step, we theoretically analyze how productivity and financial constraints affect a firm's choice to become a multinational firm under conditions of limited internal

* This chapter is based on joint work with Claudia Buch, Alexander Lipponer, and Monika Schnitzer.

[24] In the crisis that started in 2007, for instance, an increasing number of German firms reports credit constraints as an impediment to expansion into foreign countries (DIHK 2009).

[25] See, for example, Markusen (2002).

funds and the need to obtain external debt finance. Our model features limited contract enforceability and liquidation cost as two sources of inefficiencies in financial contracting that are particularly relevant for foreign investments. The model provides a set of testable implications concerning the impact of financial constraints, productivity, and host-country characteristics on firms' internationalization choices. In particular, we predict that financial constraints are more likely to affect the extensive than the intensive margin, unless financial constraints are severe. Furthermore, we predict that financial constraints are more strongly felt for large firms, as they are more likely to be interested in foreign expansion.

In a second step, we provide empirical evidence using data for German firms. We obtain information on the foreign affiliates of German firms from a detailed firm-level database provided by the Deutsche Bundesbank, the Direct Investment Micro-Database (*MiDi*). Furthermore, we use data on the balance sheets of firms in Germany from both the *Dafne* database provided by Bureau van Dijk and the *Hoppenstedt* database. Our data are unique as they allow measuring financial constraints and productivities at the parent level for both domestic firms *and* for multinationals, as well as financial constraints at the affiliate level. This enables us to analyze the extensive and the intensive margins of FDI. Furthermore, we can evaluate the relative importance of financial constraints at both the parent and at the affiliate level, a question that has - to the best of our knowledge - not been addressed in the literature so far. In contrast to earlier work focusing on manufacturing firms, our sample also contains services firms.

Our research is motivated by recent theoretical work stressing the importance of productivity for firms' international expansions. Seminal papers focusing on firms' export decisions are Bernard *et al*. (2003) and Melitz (2003). Helpman *et al*. (2004) extend the Melitz model to account for multinational firms. The key to these models is that, ex ante, firms do not know their productivity. Upon entry, firms draw their productivity from a commonly known productivity distribution, and the level of productivity becomes common knowledge as well. Depending on the level of productivity, firms exit the market, they produce only for the domestic market, they become exporters, or they set up affiliates abroad.

The implicit assumption of these models is that firms can finance foreign operations internally and/or without incurring an external finance premium. Recent papers introduce financial constraints into the Melitz model. The focus of these models is on firms' decisions to export. Chaney (2005) predicts that financially constrained firms are less likely to be able to cover the fixed cost of exporting. Manova (2010) examines the interaction of productivity and credit constraints and their impact on the export decision as well as the volume of export.

Recent empirical work shows that financial frictions indeed affect export behavior. Using panel data on bilateral exports at the country level, Manova (2010) finds that financially more developed countries are more likely to export, and that the effect is more pronounced in financially vulnerable sectors. Firm-level studies show that financial constraints matter more for the extensive margin than for the intensive margin of exports (Berman and Héricourt (2010)), that export starters enjoy better financial conditions (Bellone *et al.* (2008)), and that financially healthy firms are more likely to export (Greenaway *et al.* (2007)).[26] Stiebale (2008), in contrast, finds no effect of financial constraints on a firm's export decision once observed and unobserved firm heterogeneity is accounted for.

This essay provides complementary evidence on the role of financial frictions for FDI. As predicted by our model, we find that productivity and financial constraints have a significant impact on German firms' internationalization decision. Economically, productivity and financial constraints are of similar importance, but financial constraints matter most to the subset of firms that consider investing abroad. Our model also suggests that the extensive margin is more likely to be affected than the intensive margin, unless financial constraints are severe. Our empirical analysis shows that parent financial constraints have indeed a negative impact on the extensive margin of FDI, but less so on the intensive margin, mirroring findings by Berman and Héricourt (2010) for exports. However, we also find that, in contrast to the parent-level constraints, the affiliate's financial constraints matter for the intensive margin. This observation points towards a hierarchy of financing the intensive margin, with affiliate financing being preferred over parent financing.

In the following section, we present our model of multinational firms. In section three, we describe our data and provide descriptive statistics. Section four provides empirical evidence, and section five concludes.

3.2 Finance and the margins of FDI: Theory

In this section, we analyze a firm's choice to become a multinational firm and the volume of sales of its foreign affiliates in the presence of financial constraints. Firms incur a fixed cost of market entry as well as a variable cost of production. They finance their foreign expansion using internally generated funds as well as an external bank credit, potentially secured by collateral. Financing decisions are made under uncertainty.

Financial constraints are firm-specific; they do not merely reflect differences across firms with regard to productivity. We do not specify the sources of "financial heterogeneity",

[26] Evidence on the reverse causality from exporting to financial conditions is mixed (Bellone *et al.* (2008), Greenaway *et al.* (2007)).

but there are several reasons why firms may have different financial constraints. Firms differ, for instance, with regard to their customer structure and, thus, the probability of being hit by an adverse demand shock. Firms also differ with regard to the quality of their management and, thus, the ability of outside lenders to extract information on the profitability of an investment project.

Financial contracting in our model suffers from potential inefficiencies due to limited enforceability of financial contracts, a problem particularly relevant when investing in a foreign country. Enforceability differs across countries and may be linked to the development of the financial market as well as the presence of home country banks abroad. With limited contract enforcement, collateral may be required to obtain credit financing. However, collecting and liquidating collateral generates transaction costs, and the amount of collateral available may be limited. The need for costly and limited collateral confines the use of external finance and thus the foreign expansion of firms.

To see how the model works, consider the decision problem of a multinational firm that can invest abroad to serve the foreign market.[27] The firm's alternative investment option is normalized to zero.[28] To set up a foreign affiliate, the firm has to incur a fixed cost of market entry F. Once the firm has decided to set up a foreign affiliate, it has to choose the level of sales. Thus, we capture both the extensive and the intensive margins of the firm's foreign expansion strategy. To fix ideas, consider the following variable production cost function, $k(x) = \frac{x^2}{2(1+\beta)}$, where x denotes the quantity produced and sold by the foreign affiliate. The productivity of the parent firm, which also spills over onto the foreign affiliate, is captured by β. The larger the fixed cost of entry and the lower a firm's productivity, the larger are the "real barriers" that a firm faces when entering foreign markets.

The firm also faces a "financial barrier" in the form of a cash-in-advance constraint because set up and production cost has to be paid before production starts and before revenues are generated. Revenues that can be generated on the foreign market are uncertain. Serving the foreign market yields positive revenues px with probability q and zero

[27] We focus on horizontal FDI. The qualitative implications of our model with regard to the impact of financial constraints would also go through for vertical FDI.

[28] It is straightforward to extend our model and to include an outside option like exports that depends positively on the firm's productivity. As we show in Buch *et al.* (2010), the firm's productivity level matters relatively more for the investment opportunity abroad than for the outside option of exporting. The qualitative results of our model are unchanged.

revenues with probability $(1-q)$, where p is the foreign price level.[29]

Benchmark case without liquidity constraints

Before we describe the impact of financial constraints on investment decisions, consider as a benchmark the first-best situation where the firm is not liquidity constrained. The firm can finance both the fixed cost of entry and the variable cost of production from internal funds L . Thus, it maximizes the following profit function:

$$\pi = qpx - k(x) - F = qpx - \frac{x^2}{2(1+\beta)} - F \tag{3.1}$$

Taking the first-order condition, solving for the optimal sales of the affiliate $x_{\text{FB}} = (1+\beta)qp$ and inserting it back into the profit function (3.1) yields the following profits under the first-best solution (FB)

$$\pi_{\text{FB}} = \frac{1}{2}q^2p^2(1+\beta) - F \tag{3.1a}$$

Thus, if liquidity is not an issue, the investment takes place if and only if $\pi_{\text{FB}} \geq 0$, i.e. if net profits of the investment are positive. Not surprisingly, profits depend positively on the firm's productivity β , i.e. less productive firms are less likely to be able to cover the fixed cost of market entry.

Foreign expansion with liquidity constraints

Consider now the situation where the firm is liquidity constrained, which we define as a situation in which its liquid assets L are not sufficient to cover the cost associated with market entry and production. Thus, the firm needs external finance. We assume that external finance is raised in the form of debt finance and, specifically, credits from banks. Firms can obtain credits from domestic or foreign banks. We do not model this choice explicitly and hence do not impose restrictions with regard to the degree of integration of financial markets. However, domestic and foreign banks may differ with regard to their ability to enforce contracts. For instance, if domestic banks maintain affiliates in the foreign country, too, they are in a better position than banks operating abroad solely to monitor the affiliates and collect collateral. This adds to the comparative advantage that they already have in terms of knowledge about the domestic parent. The focus on

[29] We abstract from exchange rate changes, i.e. revenues generated on the foreign market can be remitted 1:1 into domestic currency. Russ (2007) has a model in which endogenous adjustment of exchange rates affects firms' entry decisions.

external debt finance assumption reflects the fact that external equity finance plays a limited role for German firms (?). Also, theoretical considerations suggest a "pecking order" of external finance according to which external equity finance and portfolio capital are dominated by bank lending.

Let D denote the credit necessary to finance the fixed and variable cost of entry for a production level x, given the available liquid funds L , i.e. $D = k(x)+F-L$. Furthermore, let $(1 + r)D$ denote the repayment of principal plus interest payment that the firm is supposed to pay. Like Manova (2010) and others, we assume that credit repayment is possible only if the revenues from foreign sales are positive. In particular, we rule out the possibility that the parent firm steps in and repays the affiliate's credit if the affiliate is not able to do so. This implies also that the credit repayment $(1 + r)D$ cannot exceed the revenues px, i.e. $(1+r)D \leq px$. Banks are assumed to operate competitively and to determine the interest rate such as to just break even in expected terms.

To capture enforcement problems in financial contracts, we assume that credit repayment cannot be enforced with certainty, even if revenues are positive, but only with probability μ, with $0 \leq \mu \leq 1$. The enforcement parameter μ has two interpretations. On the one hand, it can reflect different institutional quality across countries. Legal systems may, for instance, differ with regard to the degree of creditor friendliness and the enforceability of contracts.[30] On the other hand, it could reflect a greater presence of home-country multinational banks in the host country. These banks may be able to acquire useful information on the host-country environment and be able to monitor firms more closely through their affiliates abroad. This reduces informational asymmetries and makes it more likely that credit enforcement is successful.

The firm can collateralize (part) of its credit with assets from two potential sources. First, the firm can pledge its fixed cost investment in the foreign affiliate, F, as collateral. Second, the firm can use an exogenously given collateral, $\overline{C}$, provided by the parent company, to secure the credit. Let $C \leq \overline{C} + F$ denote the collateral actually chosen to secure the credit, the exact value of which is determined endogenously below. If the credit is not repaid, the creditor can seize the collateral to cover her losses. However, she can realize only a fraction θ of the collateral when liquidating it.[31] Thus, liquidating the collateral involves a dead weight loss of $(1 - \theta)C$.

There are two situations where liquidation of a collateral (potentially) becomes an issue. Suppose the affiliate has positive revenues but the creditor fails to be able to enforce the

[30] Harrison *et al*. (2004) report that financial development lowers financial constraints.

[31] Without loss of generality, we assume that the efficiency loss is the same for both kinds of collateral goods.

repayment. Then, the bank has the option to liquidate the collateral. However, it would be inefficient to do so, due to the dead weight loss of liquidation. In this case, we assume that efficient renegotiation will make the firm pay θC, i.e. the amount that the bank can realize from liquidating the collateral, to avoid inefficient liquidation, and the bank will accept this offer.[32]

Now, consider the zero profit condition for banks which determines the interest rate for a given choice of C:

$$\mu q(1+r)D + (1-\mu q)\theta C = D \tag{3.2}$$

Banks obtain the promised credit repayment $(1+r)D$ only if credit repayment can be enforced. In all other cases, they obtain the liquidation value of the collateral, θC, either because this is what the firm pays voluntarily, after renegotiation, or this is what they receive from actually liquidating the collateral. Solving for $(1+r)D$, we find that banks charge a risk premium over and above the risk-free rate which is declining in the probability of success of the project q and in the efficiency of the liquidation procedure, θ

$$(1+r)D = \frac{D-(1-\mu q)\theta C}{\mu q} \tag{3.3}$$

Recall from above that the maximum repayment cannot exceed revenues, requiring

$$(1+r)D = \frac{D-(1-\mu q)\theta C}{\mu q} \leq px \tag{3.4}$$

Note that the smaller μ, the more important it is to pledge a collateral for this condition to be satisfied. However, due to the dead weight loss in case the collateral is actually liquidated, which happens with positive probability, the firm limits the collateral pledged to the minimum required to obtain the desired credit. Inserting $D = k(x) + F - L$ and solving for C yields the minimum collateral needed to finance the fixed cost of market entry and a given level of affiliate sales x taking into account that the collateral has to be non-negative:

$$C^*(x) = \max\left\{0, \frac{[k(x)+F-L]-\mu qpx}{(1-\mu q)\theta}\right\} \tag{3.5}$$

The larger the required credit, the larger is the minimum collateral needed. Note, however, that the collateral cannot exceed the upper bound specified above, $\overline{C}+F$. We consider,

[32] This assumes that the firm can hold the bank down to its outside option of liquidating the collateral. It would be straightforward to modify this assumption and let the two parties split the gains from not liquidating the collateral. However, given our assumption of a perfectly competitive banking market, the first assumption seems to be the most convincing one. If revenues are not positive, however, liquidation of the collateral cannot be avoided.

in turn, the cases where this upper bound of collateral constrains the firm's optimal sales choice and where it does not, starting with the case of a non-binding collateral constraint.

3.2.1 Non-binding collateral constraint

Suppose for a moment that the collateral constraint is not binding. Then, for a given level of affiliate sales x and collateral C, the firm expects the following profits:

$$\pi = qpx - \mu q(1+r)D - (1-\mu q)\theta C - (1-q)C - [k(x)+F] + D \tag{3.6}$$

The first term reflects the expected revenues, the second term the debt repayment that can be enforced with probability μ if revenues are positive, which happens with probability q. If credit repayment cannot be enforced, the firm voluntarily pays what the bank would be able to collect in the event of liquidation, θC, to avoid costly liquidation, as discussed above. If revenues are not positive, however, the collateral will be liquidated, as captured by the fourth term. The last terms capture the cost of market entry and production and the credit obtained by the firm to finance this cost, respectively.

The firm maximizes its profits by choosing the optimal sales of the affiliate x taking into account the collateral needed to finance market entry and production, $C^*(x)$.

Using $D = k(x) + F - L$ and the equations (3.3) and (3.5) for $(1+r)D$ and $C^*(x)$ we obtain:

$$\pi = qpx - k(x) - F - (1-q)(1-\theta)\max\left\{0, \frac{[k(x)+F-L-\mu qpx]}{(1-\mu q)\theta}\right\} \tag{3.6a}$$

Note that if $C^*(x) = 0$, i.e. if no collateral is needed to secure the credit, financing cost does not bias the investment decision. If collateral is needed, however, expected profits are lowered by the expected liquidation cost, $(1-q)(1-\theta)C^*(x)$.

The following proposition characterizes the solutions of the firm's maximization problem.

Proposition 1*: **Non-binding collateral constraint - Extensive and intensive margins***

The profit-maximizing sales level x^* is characterized by the following solution:

$$x^* = \begin{cases} \frac{1+\mu z}{1+z}(1+\beta)qp < x_{\mathrm{FB}},\ for\ C^*(x^*) > 0 \\ (1+\beta)qp = x_{\mathrm{FB}},\ for\ C^*(x^*) = 0 \end{cases} \quad \text{with } z = \frac{(1-q)(1-\theta)}{(1-\mu q)\theta} \tag{3.7}$$

The maximum profit the firm can attain is given by

$$\pi^* = \begin{cases} \frac{(1+\mu z)^2}{(1+z)}\frac{1}{2}(1+\beta)q^2p^2 - z(F-L) - F \leq \pi_{\text{FB}},\ for\ C^*(x^*) > 0 \\ \frac{1}{2}(1+\beta)q^2p^2 - F = \pi_{\text{FB}},\ for\ C^*(x^*) = 0 \end{cases} \tag{3.8}$$

provided that the maximum exogenous collateral is not binding, i.e.

$$\overline{C} \geq C^*(x^*) - F$$

Proof: See Appendix

Note that for $\mu = 1$, the optimum level of sales is the same as the first-best level. Also, if $\theta = 1$, then $z = 1$, and again the optimum level of sales is the same as in the first-best case. Thus, the optimum level of sales differs from the first-best choice only if both $\mu < 1$ and $\theta < 1$. The intuition for this is straightforward. Only if contract enforcement is less than perfect may a collateral be required to obtain a credit, and only if the use of a collateral is costly does it affect the marginal cost of financing the production. Thus, only if a costly collateral is required do profits fall short of first-best profits.

Of course, the firm will engage in FDI only if the maximum profits from investment are non-negative. The following proposition characterizes the comparative statics for the firm's extensive and intensive margins of investment.

Proposition 2*: **Non-binding collateral constraint - Comparative statics***
Changes in the following parameters affect the probability of non-negative profits and thus the probability of engaging in FDI:

$$\frac{d\pi^*}{d\beta} > 0, \frac{d\pi^*}{dp} > 0, \frac{d\pi^*}{d\theta} > 0, \frac{d\pi^*}{d\mu} > 0, \frac{d\pi^*}{dF} < 0, \frac{d\pi^*}{dL} > 0, \frac{d\pi^*}{d\overline{C}} = 0$$

Furthermore, the intensive margin is described by the following comparative static results for the optimal volume of foreign affiliate sales:

$$\frac{dx^*}{d\beta} > 0, \frac{dx^*}{dp} > 0, \frac{dx^*}{d\theta} > 0, \frac{dx^*}{d\mu} > 0, \frac{dx^*}{dF} = \frac{dx^*}{dL} = \frac{dx^*}{d\overline{C}} = 0$$

Proof: See Appendix

Both the optimal volume of sales and the firm's profits increase in the firm's productivity and in the lucrativeness of foreign markets. Furthermore, better contract enforcement in the host country has a positive effect on sales and profits because it lowers the requirement to use costly collateral, and improving the efficiency of liquidating collateral reduces cost. Higher fixed cost lowers expected profits not only directly but also indirectly. The larger the fixed cost, the fewer liquid funds are left for financing the investment. Less liquid funds, in turn, mean greater need for using costly collateral. Hence, there is an indirect negative effect of fixed cost over and above the direct effect. However, fixed cost and internal funds do not affect the optimal level of sales choice because the marginal cost of using collateral does not depend on how much collateral is actually needed. The maximum collateral, in turn, has no effect on profits and on the firm's choice of sales as long as it does not impose a binding constraint.

This scenario describes the situation of a mildly financially constrained investor. The need for credit financing and the requirement of providing collateral increase the marginal cost of investment and hence limit the volume of sales and profits. However, as long as the collateral requirement does not impose a binding constraint, the constraints are not as severe, as fixed cost and internal funds affect the extensive margin only, not the intensive margin.

3.2.2 Binding collateral constraint

Consider now the case where the collateral constraint is binding for the optimal sales level determined above, $x = x^*$, i.e.

$$\overline{C} + F < C^*(x^*) = \left\{0; \frac{[k(x^*) + F - L] - \mu q p x^*}{(1-\mu q)\theta}\right\} \tag{3.9}$$

In this case, x^* cannot be implemented because the credit constraint becomes binding. Instead, production settles at a smaller level $\overline{x}$ that is determined by the maximum available exogenous collateral:

$$\overline{C} + F = \frac{[k(\overline{x}) + F - L] - \mu q p \overline{x}}{(1-\mu q)\theta} \tag{3.10}$$

Solving this equation for $\overline{x}$ and inserting it into the firm's profit function yields the constrained optimal level of sales choices and profits as characterized by the following proposition.

Proposition 3: ***Binding collateral constraint - Extensive and intensive margins*** *Suppose the maximum exogenous collateral imposes a binding constraint on the firm's optimal choice of the level of sales, i.e.*

$$\overline{C} < C^*(x^*) - F \tag{3.11}$$

Then, the investor can attain a maximum profit of

$$\overline{\pi} = qp\overline{x} - [k(\overline{x}) + F] - (1-q)(1-\theta)[\overline{C} + F] \leq \pi^* \tag{3.12}$$

Where the foreign level of sales $\overline{x} < x^*$ is determined by equation (3.10)

Proof: See Appendix

Not surprisingly, profits fall short of the second-best profits that can be attained if the collateral constraint is non-binding. The following proposition characterizes the comparative static results for the extensive and intensive margins.

Proposition 4*: **Binding collateral constraint – Comparative statics***

The following comparative static results characterize the extensive margins of FDI, summarizing which parameters are more or less likely to ensure non-negative profits:

$$\frac{d\pi}{d\beta} > 0, \frac{d\pi}{dp} > 0, \frac{d\pi}{d\theta} > 0, \frac{d\pi}{d\mu} > 0, \frac{d\pi}{dF} < 0, \frac{d\pi}{dL} > 0, \frac{d\pi}{d\overline{C}} > 0$$

and

$$\frac{d^2\overline{\pi}}{d\overline{C}d\beta} > 0, \frac{d^2\overline{\pi}}{dLd\beta} > 0$$

Furthermore, the intensive margin is described by the following comparative statics for the optimal volume of foreign affiliate sales:

$$\frac{d\overline{x}}{d\beta} > 0, \frac{d\overline{x}}{dp} > 0, \frac{d\overline{x}}{d\theta} > 0, \frac{d\overline{x}}{d\mu} > 0, \frac{d\overline{x}}{dF} < 0, \frac{d\overline{x}}{dL} > 0, \frac{d\overline{x}}{d\overline{C}} > 0$$

and

$$\frac{d^2\overline{x}}{d\overline{C}d\beta} > 0, \frac{d^2\overline{x}}{dLd\beta} > 0$$

Proof: See Appendix

Like above, productivity, lucrativeness of foreign markets, contract enforcement, and the efficiency of collateral liquidation positively affect both the extensive and the intensive margin of foreign direct investment. Unlike before, however, fixed cost and internal funds now affect the level of sales as well, because higher fixed cost (or fewer internal funds) leaves fewer funds for the financing of production, which cannot be compensated by increasing credit financing if the collateral constraint becomes binding. And of course both margins are positively affected if the collateral constraint becomes less binding.

We also find that the financial status of the firm as captured by the liquid funds and the collateral available plays a more important role for more productive firms, since they are the ones more likely to invest. Thus, a high productivity is a necessary, but not a sufficient condition for foreign expansion.

This scenario captures the case of a more severely financially constrained firm that is not only exposed to higher marginal cost of credit financing, but that is also constrained in its access to collateral. The firm is constrained not only at the extensive, but also at the intensive margin of expansion. Of course, in reality, the two cases may be considered as representing the two limits of a continuous distribution, with marginal cost of using a collateral increasing in the size of the collateral. It would be straightforward to generalize our set up and to allow for a more continuous distribution of financial constraints.

3.2.3 Financial constraints at the affiliate level

So far, we have assumed the liquid funds (L) and the exogenous collateral ($\overline{C}$) to be provided by the parent firm. For the market entry decision, this is the natural assumption. Over time, however, the foreign affiliate may in turn accumulate earnings and collateral goods that may affect the financing constraints for the volume of sales. A natural extension of the model would thus be to take into account liquid funds and collateral goods provided by the affiliate itself. It seems plausible to conjecture that funds provided by the affiliate incur lower opportunity cost and/or dead weight losses than funds provided by the parent firm.[33] If this is the case, we would expect funds provided by the affiliate to be used first, and only if they are not sufficient would we expect them to be supplemented by funds provided by the parent.

3.2.4 Summing up

The model has rich implications for the determinants of firms' intensive and extensive margins of foreign activities. Higher productivity, more efficient liquidation of collateral, better contract enforcement, and more lucrative foreign markets always increase the volume of affiliate sales. Higher fixed cost decreases and higher internal funds increase activities. The impact of these variables on the intensive margin depends on whether the collateral constraint is binding. They have no effect on the intensive margin if the available collateral is sufficiently large. Likewise, the impact of the size of the collateral depends on the scenario considered. It should matter most when the collateral available

[33] This is a topic discussed extensively in the literature on internal capital markets. See for example Brusco and Panunzi (2005) or Inderst and Laux (2005). For a survey see Stein (2003).

is low. Finally, our model predicts that financial constraints matter more for larger, more productive firms, since these firms are more likely to be interested in foreign expansions. Table 3.1 gives an overview of the results of the comparative static analyses.

3.3 Data and stylized facts

3.3.1 Data sources

To investigate the importance of real and financial constraints for the foreign investment choices of firms, we use data from three sources.[34] *Dafne* and *Hoppenstedt* are commercial databases providing financial information on a large panel of firms that are active in Germany.[35] We use these datasets to obtain information on parent-level financial constraints and productivity. Information on the number of German firms' foreign affiliates, their sales, the host countries, and affiliate-level financial constraints are obtained from the firm-level database on multinational firms *MiDi* (Microdatabase Direct Investment), provided by the Deutsche Bundesbank (Lipponer (2006)).

To eliminate outliers, we start from the full *Dafne* dataset and drop firms with negative values for key variables such as sales and total assets. Also, as we need information on cash flow and sales, we eliminate observations for firms which do not report an income statement. We additionally truncate some of the data at the 1st and 99th percentile. Finally, we drop observations showing large changes in sales or in the number of employees from one year to another (increase by a factor of 10 or drop to 1/10 or less) in order to control for possible merger-induced outliers.

Table 3.4 compares the structure of the sample after the outlier correction ("corrected sample") and the sample used for the regressions in Table 3.6 ("regression sample"). The two samples are fairly similar in terms of the percentage allocation of the number of firms across sectors. We have also compared the structure of our sample to the sectoral structure of the German economy as a whole, and the rank correlation in terms of sectoral structure of sales has proven to be quite high.

[34] See Tables 3.2 and 3.3 for details.

[35] *Dafne* is the German equivalent to the European firm-level database *Amadeus*. Bayraktar *et al*. (2005) also use the German data from *Amadeus* for an analysis of firm-level domestic investment behavior.

Table 3.1: Summary of the theoretical model and empirical measurement

This table summarizes the comparative static results of the model presented in Section 3.2. See Table 3.2 for the definitions of the empirical variables.

Parameter	Measurement	Proposition 2: Non-binding collateral constraint		Proposition 4: Binding collateral constraint	
		Extensive margin	Intensive margin	Extensive margin	Intensive margin
Productivity (β)	Cost efficiency Sales / Total assets	+	+	+	+
Foreign prices (p)	GDP GDP per capita	+	+	+	+
Liquidation value (θ)		+	+	+	+
Probability of contract enforcement (μ)	Bank FDI	+	+	+	+
Probability of no contract enforcement ($1-\mu$)	(Weak) contract enforcement	–	–	–	–
Fixed costs (F)	Fixed / Total assets	–	0	–	–
Internal funds (L)	Cash flow of the parent Retained earnings of the affiliate	+	0	+	+
Collateral (C)	Debt ratio of the parent	0	0	+	+

Table 3.2: Data

Unless otherwise indicated, parent-level information comes from *Dafne* (Bureau van Dijk) and *Hoppenstedt*, affiliate level information comes from *MiDi* (Microdatabase Direct Investment, Deutsche Bundesbank). Country-level information comes from the World Bank's World Development Indicators. All values in €1,000 (unless otherwise indicated). Cash flow and cost efficiency are corrected for outliers by truncating the data at the 1st and 99th percentile. Fixed asset share, the debt ratio, and sales are corrected for outliers by truncating the data at the 99th percentile.

Variable	Definition
	Parent-level data
Cash flow	Cash flow from operations
Cost efficiency	Sales / total cost (cost of materials + labor cost)
Debt ratio	Total debt / total assets
Firms with foreign affiliate	0/1 dummy for firms with foreign affiliates from *Dafne*-*MiDi*-merge
Fixed asset share	Fixed assets / total assets
Number of foreign affiliates	Count of total number of affiliates worldwide obtained from *MiDi*
Sector definitions	We use two definition of sectors: (i) A broad definition of 28 sectoral groups is used for sample splits (see also Table 5), (ii) a narrow definition of about 64 sectors at the 2-digit-level is used to generate sector-level dummy variables
Sales	Turnover
State dummies	Dummy variables for the 16 German states
	Affiliate-level data
Debt ratio	Total debt / total assets
Sales	Aggregate turnover of parent *i* in country *j* in year *t*, i.e. data are aggregated across all affiliates in a given country for a given parent and weighted by the parent's ownership share
Retained earnings / total assets	Revenue reserves / total assets
	Country-level data
Bank FDI	Aggregate volume of FDI of German banks in country *j* in year *t*, calculated from *MiDi* in €1,000
(Weak) contract enforcement	From the World Bank's "Doing business" database (http://www.doingbusiness.org/), we use the variable "Enforcing contracts / Procedures (number)"
GDP	Host country GDP per capita in constant USD, converted into €bn, World Bank (2008)
GDP per capita	Host country GDP per capita per capita in constant USD, converted into €1,000, World Bank (2008)

Table 3.3: Descriptive statistics

This Table provides summary statistics for the regressions reported below. GDP per capita is in €1,000. Negative values in ln(GDP per capita) hence come from countries with a GDP per capita of less than €1,000. Minimum and maximum values for affiliate-level variables are not reported due to confidentiality reasons.

a) Extensive margin

Variable	Obs	Mean	Std. Dev.	Min	Max
Cash flow (log)	176,034	5.347	2.245	0.000	10.653
Cost efficiency	136,093	1.344	0.444	0.383	4.750
Debt ratio	203,325	0.561	0.291	0.000	0.999
Exporter dummy	211,205	0.072	0.259	0.000	1.000
FDI dummy	211,205	0.020	0.140	0.000	1.000
Fixed / total assets	184,882	0.267	0.269	0.000	0.970
Size (log)	211,143	7.825	2.404	0.000	18.922

b) Intensive margin

	Obs	Mean	Std. Dev.	Min	Max
Affiliate-level					
Debt ratio	17,475	0.516	0.269	…	…
Retained earnings / total assets	17,475	0.059	0.132	…	…
Sales (log)	16,582	10.095	1.286	…	…
Parent-level					
Cash flow (log)	3,980	11.171	4.207	0.000	19.441
Cost efficiency	3,682	1.307	0.330	0.393	4.690
Debt ratio	5,269	0.433	0.229	0.000	0.999
Fixed assets / total assets	4,924	0.246	0.219	0.000	0.963
Number of foreign affiliates	4,222	4.429	9.878	1.000	…
Size (log)	5,129	13.919	3.726	3.296	21.484
Country-level					
Bank FDI (log)	296	11.601	2.282	4.754	16.812
(Weak) contract enforcement (number of procedures)	243	36.078	6.373	21.000	51.000
GDP (log)	438	4.541	1.873	-0.664	9.762
GDP per capita (log)	434	1.707	1.386	-1.853	4.001

Table 3.4: Corrected versus regression sample

This table compares the sample corrected for outliers ("corrected sample") and the sample used for the regressions in Table 6 ("Regression sample"). The two samples differ because of missing observations for the explanatory variables.

	Regression sample				Corrected sample			
	Number	%	Sales (million €)	%	Number	%	Sales (million €)	%
Agriculture & Fishing	1,172	1.63	5,242	0.14	2,435	1.49	12,744	0.16
Chemicals	1,219	1.70	158,715	4.23	1,908	1.17	251,227	3.17
Construction	8,166	11.36	136,271	3.63	17,220	10.56	184,209	2.32
Education	273	0.38	3,293	0.09	798	0.49	17,049	0.21
Energy	2,308	3.21	183,807	4.90	4,271	2.62	598,190	7.54
Financial services	323	0.45	25,840	0.69	1,906	1.17	109,836	1.38
Food & Tobacco	1,568	2.18	169,024	4.50	2,865	1.76	387,082	4.88
Furniture	1,043	1.45	54,521	1.45	1,804	1.11	49,107	0.62
Glass	902	1.25	66,058	1.76	1,515	0.93	52,436	0.66
Health	2,302	3.20	75,202	2.00	4,596	2.82	158,890	2.00
Hotels & Restaurants	600	0.83	7,202	0.19	1,549	0.95	17,713	0.22
Coking	84	0.12	82,420	2.20	163	0.10	64,276	0.81
Leather	62	0.09	3,122	0.08	99	0.06	3,052	0.04
Machinery	3,502	4.87	262,599	7.00	5,934	3.64	317,839	4.01
Metals	4,063	5.65	138,351	3.69	7,619	4.67	283,242	3.57
Mining	279	0.39	13,555	0.36	572	0.35	128,760	1.62
Office equipment	2,695	3.75	235,198	6.27	4,718	2.89	253,050	3.19
Other services	2,382	3.31	72,376	1.93	6,482	3.97	213,489	2.69
Paper	1,566	2.18	62,047	1.65	3,052	1.87	138,107	1.74
Real estate & Business services	13,854	19.27	595,754	15.87	44,063	27.02	1,536,041	19.37
Rubber & Plastics	1,248	1.74	82,808	2.21	2,152	1.32	88,540	1.12
Textiles	736	1.02	22,114	0.59	1,335	0.82	55,160	0.70
Trade & repair	16,706	23.23	1,041,823	27.76	34,642	21.24	2,047,607	25.82
Transport & Communication	3,460	4.81	158,111	4.21	8,341	5.11	631,385	7.96
Vehicles	786	1.09	82,430	2.20	1,436	0.88	297,970	3.76
Wood	435	0.60	11,598	0.31	920	0.56	20,855	0.26
n.e.c	177	0.25	4,149	0.11	684	0.42	13,808	0.17
Total	71,911	100.00	3,753,631	100.00	163,079	100.00	7,931,660	100.00

3.3.2 Dependent and explanatory variables

Extensive and intensive margin

By merging the firm-level databases *Dafne* and *Hoppenstedt* with information on the foreign affiliates of German firms provided in *MiDi*, we obtain a dataset which includes two groups of firms. The first group contains purely domestic German firms, i.e. firms which do not hold affiliates abroad ("Domestic Firms") (94.5% of the firm-year observations). The second group consists of German firms with foreign affiliates ("German MNEs") (5.5%). From *MiDi*, we also obtain a count variable on the number of affiliates that a given parent operates abroad. This serves as an additional proxy for the extensive margin of foreign activities, which measures complex FDI strategies involving many affiliates. We also have information on the volume of a firm's foreign affiliates' sales as a measure of the intensive margin.

Productivity

In line with the theoretical model, we use cost efficiency as a firm-level measure of productivity. Cost efficiency is given by parent sales over total cost, i.e. labor cost plus the cost of other inputs. A higher value reflects higher cost efficiency, hence we expect a positive sign. Higher sales relative to total cost might also reflect higher mark-ups. The expected sign of the coefficient would be the same. We include the size of the parent as a measure for its productivity, and the expected sign is positive.

Fixed cost

The parent's fixed cost of investment is proxied by the ratio of fixed over total assets. We use the ratio rather than the level of this variable as we additionally account for size effects in our regressions. We expect a negative impact of the fixed asset share on the extensive margin. The impact of this variable on the intensive margin could be insignificant, according to our model, if the collateral available is sufficiently large.

Internal funds

In our model, we distinguish liquid funds from less liquid collateral as two determinants of financial constraints. Log cash flow of the parent is used to measure the internal funds available for financing a particular investment project. This variable should have a positive impact on the extensive margin of foreign activities. As in the case of fixed cost, its impact could be insignificant on the intensive margin if the collateral available is sufficiently large. In addition, we look at retained earnings of the affiliate as a measure for the liquid funds available to the affiliate to finance the intensive margin. Again, the expected sign is positive or insignificant.[36]

Collateral

The debt ratio measures leverage at the parent and at the affiliate levels *ex ante*. We can interpret the debt ratio as a measure of the firms' collateral – firms which are more highly leveraged ex ante have, *ceteris paribus*, fewer assets available that can serve as collateral for new credits. Hence, the expected sign for the parent debt ratio is negative for both the extensive and the intensive margins if the collateral constraint is binding. Similarly, the expected sign for the affiliate debt ratio is negative for the intensive margin. Firms may also report a high leverage ratio precisely because they have taken out a credit in order to finance FDI. If this were the correct interpretation, we should expect a positive sign of the coefficient.

Foreign market size

In our theoretical model, we have described the attractiveness of the foreign market in terms of the price that firms can fetch abroad for their product. In our empirical model, we distinguish two aspects of foreign market size. The first is the size of the market measured through its GDP. The second is the state of development of a foreign market measured through GDP per capita. We expect a positive sign for both variables.

Contract enforcement

The probability of contract enforcement depends on two parameters - an index measuring the difficulty of contract enforcement as well as the presence of affiliates of German banks

[36] Following Kaplan and Zingales (1997) criticism, there has been a lively debate on the usefulness of investment-cash flow sensitivities as a measure for financial constraints. The focus of the discussion have been endogeneity issues as well as issues of adequately taking into account access to external finance. See also Brown *et al.* (2009) for an overview of this discussion. We use lagged variables to address the simultaneity of firm-level variables issue. We also include the debt ratio, as discussed below.

abroad. The variable *(weak) contract enforcement* gives the number of procedures required to enforce contracts, and the expected impact is negative. This variable can be expected to influence both the entry decision as well as the volume of activities, and we include it for both margins. Affiliates of German banks should be at an advantage over other lenders with regard to monitoring foreign affiliates and enforcing contracts. We use *MiDi* to obtain information on the volume of *FDI of German banks* by country, and we expect a positive impact on the intensive margin.

3.3.3 Stylized facts

In Figure 3.1, we visualize the differences between German MNEs and Domestic Firms by plotting the Kernel densities of size (Figure 3.1(a)), cost efficiency (Figure 3.1(b)), cash flow (Figure 3.1(c)), the debt ratio (Figure 3.1(d)), and the share of fixed assets (Figure 3.1(e)).

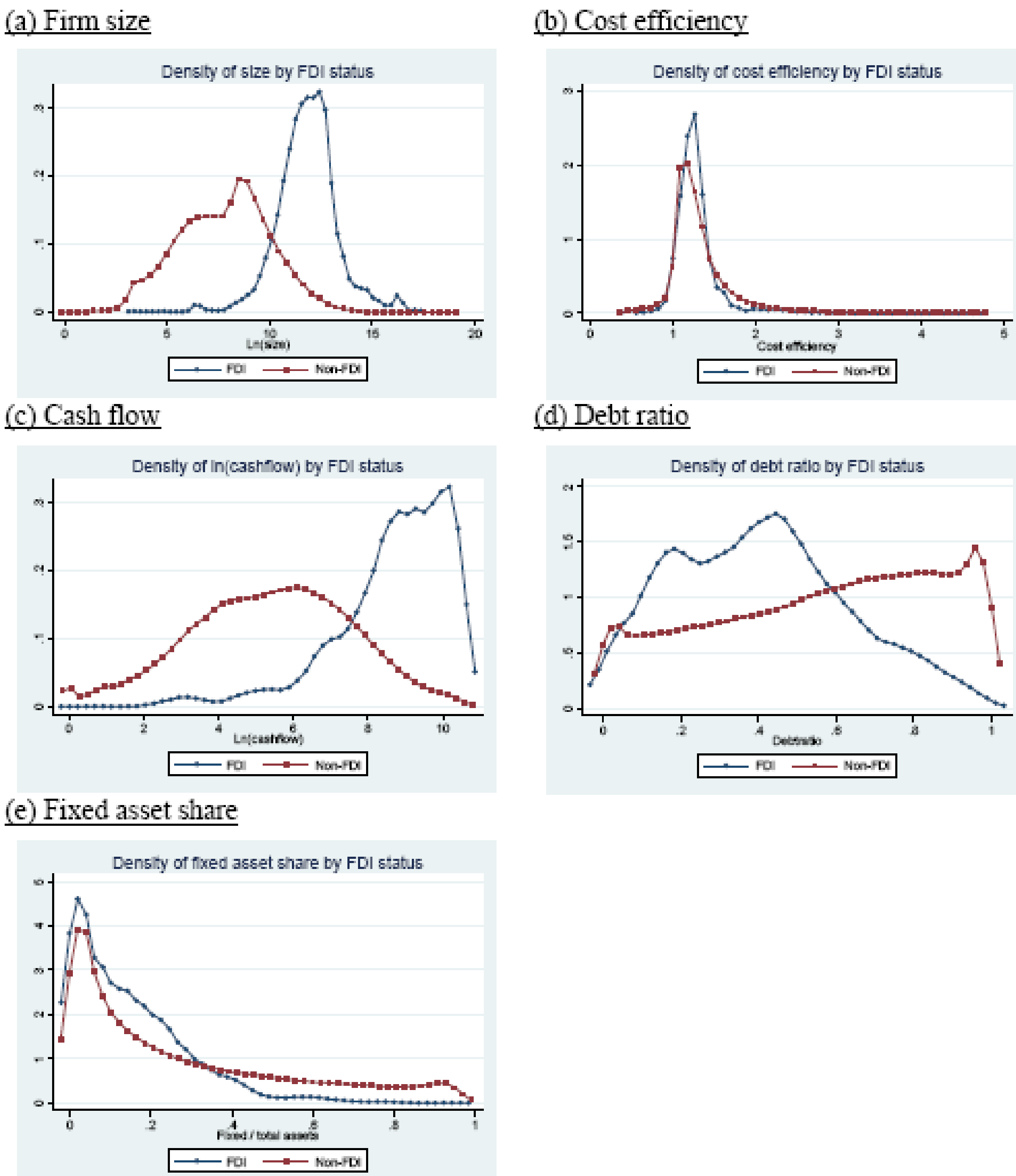
Figure 3.1(a) confirms stylized facts reported in earlier papers using firm-level data (e.g. Mayer and Ottaviano (2008)): MNEs are larger than purely domestic firms. Unreported one-sided t-tests on the equality of the means between the two sub-samples show that this difference is statistically significant. Measuring size through the volume of sales gives a very similar result. MNEs also exhibit a somewhat lower share of fixed assets (Figure 3.1(e)). Figure 3.1(b) shows that differences between the two types of firms in terms of cost efficiency are small and, in fact, not significant.

Hence, while the dividing line between multinationals and non-multinationals is not as clear-cut as might have been expected on the basis of the cost efficiency of these firms, the dividing line is clear for measures of financial status. Multinationals have significantly higher cash flow (Figure 3.1(c)) and lower debt ratios (Figure 3.1(d)). Prima facie, these graphs suggest that heterogeneity with regard to the openness and international orientation of firms could be driven by financial factors just as by real factors.

3.4 Productivity versus financial constraints: Regression results

Our main empirical model relates financial constraints and productivity to the pattern of internationalization at the firm level. We are interested in two main questions. First, do financial constraints and productivity affect the probability of investing abroad? Second, do these factors affect the volume of foreign affiliates' sales? We answer these questions in two steps. In a first step, we analyze the determinants of the firms' extensive margin of FDI using the probability of investing abroad and the number of affiliates as dependent

Figure 3.1: Firm characteristics by multinational status



variable. In a second step, we analyze the sales of affiliates across countries, i.e. the intensive margin. We also estimate the extensive and intensive margins jointly using a Heckman selection model.

Table 3.5: Probability of owning affiliates abroad

This table reports results of probit regressions using a 0/1 dummy variable of owning foreign affiliates as the dependent variable. All explanatory variables are at the parent level (P). Sample splits are at the sample median. Sector, state, and year fixed effects included. Standard errors in parentheses. Marginal effects are reported. ***, **, * = significant at the 1%, 5%, 10%-level.

	(1) Full sample	(2) Large	(3) Small	(4) Manufacturing	(5) Services	(6) Listed	(7) Unlisted
Log size t-1 (P)	0.009***	0.011***	0.006***	0.013***	0.006***	0.013***	0.008***
	(0.001)	(0.003)	(0.001)	(0.002)	(0.001)	(0.004)	(0.001)
Cost efficiency t-1 (P)	-0.010***	-0.015***	-0.003	-0.008	-0.010***	-0.006	-0.011***
	(0.003)	(0.005)	(0.002)	(0.007)	(0.002)	(0.009)	(0.003)
Debt ratio t-1 (P)	-0.005	-0.016	0.004	-0.011	0.004	0.030	-0.006
	(0.005)	(0.010)	(0.003)	(0.007)	(0.007)	(0.021)	(0.005)
Log cash flow t-1 (P)	0.005***	0.009***	0.001	0.004**	0.005***	0.008***	0.004***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.003)	(0.001)
Fixed asset share t-1 (P)	-0.037***	-0.051***	-0.017***	-0.032***	-0.032***	-0.097***	-0.031***
	(0.007)	(0.013)	(0.005)	(0.011)	(0.009)	(0.032)	(0.007)
Log GDP	0.019***	0.026***	0.014***	0.021***	0.017***	0.026***	0.018***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Log GDP per capita	0.002**	0.004**	0.001	0.001	0.004***	0.011***	0.001
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.004)	(0.001)
(Weak) contract enforcement	-0.001***	-0.001***	-0.001***	-0.001***	-0.001**	-0.001	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Exporter (0/1)	0.000	0.003	-0.003**	-0.002	0.001	-0.007	-0.000
	(0.002)	(0.005)	(0.001)	(0.003)	(0.004)	(0.007)	(0.002)
Observations	58,087	29,493	28,594	32,537	22,681	6165	51,922
Pseudo R^2	0.134	0.138	0.108	0.153	0.124	0.235	0.121
log likelihood	-9,500	-5,899	-3,532	-5,446	-3,553	-1,222	-8,197

3.4.1 Extensive margin

Our baseline regression for the extensive margin - the decision to enter a foreign market - is given by the following probit model:

$$\mathrm{Pr(FDI)_{i,k,t}} = \alpha_0 + \alpha_1 Z_{i,t-1} + \alpha_2 Z_{k,t} + \alpha_3 I + \alpha_4 S + \alpha_5 T + \epsilon_{i,t} \tag{3.13}$$

where $\mathrm{Pr(FDI)}_{i,k,t}$ indicates whether a firm i has invested abroad in year t in country k. $Z_{i,t-1}$ ($Z_{k,t}$) are vectors of firm-level (country-level) control variables.[37] We include the ratio of sales over total cost as a measure of cost efficiency. Our main proxies for financial barriers are cash flow and the debt ratio. The country-level control variables are GDP, GDP per capita, and the severity of contract enforcement. We additionally include firm size, and a full set of industry (I), German states (S), and time (T) dummies. These dummies capture systematic differences across industries and states as well as common macroeconomic effects. We also include an exporter dummy to account for the fact that exporting is typically a stepping stone into international markets (see Helpman *et al.* (2004)). This variable turns out to be positive and significant on the extensive margin regarding the number of affiliates abroad but insignificant regarding the probability of owning foreign affiliates.

Table 3.5 shows the results. Column (1) has the baseline specification for the full regression sample. In columns (2)-(7), we split the sample by size, by sector (manufacturing versus services), and by legal status (listed versus unlisted). While the sub-sample of listed firms is small (6,165 versus 51,922 firm-country-year observations), it nevertheless serves as a useful test of the impact of financial frictions. A priori, we expect financial frictions to be less important for the listed firms with access to a larger range of financial sources.

Larger and more efficient firms are more likely to be multinationals. Size has a positive and significant impact on the probability of being a multinational, and this effect is robust across specifications. Contrary to expectations, cost efficiency is negative and significant in some specifications. This effect is driven by certain sub-samples such as the large firms and the services sector firms and suggests that size is a better proxy for productivity than cost efficiency.

Our measure for fixed cost of market entry, the fixed asset share, has a strong and significantly negative impact on the probability of investing abroad for all specifications, as

[37] Firm-level regressors are lagged by one period to account for the simultaneity of the explanatory variables.

expected.[38] Berman and Héricourt (2010) as well as Manova (2010) interpret the fixed asset share as capturing the tangibility of assets, and hence as a measure of easier access to external finance secured by collateral. Following their interpretation, the expected effect is positive. The negative coefficient we find suggests that, for FDI, our interpretation is the more appropriate one. Financial constraints have a significant and robust impact on the extensive margin. Cash flow is mostly positive and significant. The debt ratio has an insignificant impact, consistent with the prediction of the model for non-binding collateral constraints.

The marginal effects reported in Table 3.5 show a similar importance of productivity and financial frictions. Generally, however, fixed cost of entry (the fixed asset share) and the country-level variables are more important than variables such as size or the debt ratio. Mean elasticities also show the strongest response to changes in log GDP (elasticity of +0.66), cost efficiency (-0.45), firm size (+0.30), the fixed asset share (+0.23), and cash flow (+0.16).

To study the interaction of productivity and financial constraints, we split the sample. We take firm size as an indicator for firm productivity. One of the financial variables - the debt ratio - is insignificant for both groups. The other - cash flow - matters for large firms, but not for small firms. The latter finding may look counterintuitive at first sight, as one would expect smaller firms to be more opaque and hence more likely to be affected by financial constraints. Our finding is, however, consistent with the prediction of our model that financial constraints should matter the more, the more productive the firm and hence the more interested it is in expanding abroad.[39] Financial constraints, in other words, do not impede the foreign expansion of small firms because these firms are not productive enough to invest abroad in the first place. It is also consistent with the finding of Berman and Héricourt (2010) who observe that productivity has no effect on a firm's export decision if the firm faces financial constraints.

The country-level variables are significant and have the expected sign. GDP is positive and significant, and GDP per capita is positive and significant for the full sample and for most of the sample splits, thus confirming the expectation that market size matters. Consistent

[38] An alternative interpretation of this finding is that firms with a large share of intangibles and thus firm-specific know-how are more likely to venture abroad. These firms would also have a lower fixed asset share.

[39] Chaney (2005) distinguishes three classes of firms, with low, intermediate and high productivity. He predicts that firms with low productivity are not affected by financial constraints, since investing abroad is not a viable option for them, even without financial constraints. More productive firms, instead, are hampered by financial constraints in their foreign expansion strategy. In his model, very productive firms are by construction not liquidity constrained and hence not affected by financial constraints.

with our model, greater difficulties with contract enforcement lower the probability that a given German firm enters a particular country.

In sum, our results show that parent-level financial constraints and productivity affect the extensive margin of foreign entry: larger, more efficient, and firms with a lower share of fixed assets are more likely to become multinationals. In addition, country-level variables capturing contract enforcement and market size play an important role for the entry decision.

3.4.2 Extensive margin: Number of affiliates

Table 3.6: Determinants of the number of affiliates

This table reports the estimated coefficients of the Poisson (Negative Binomial, Zero-Inflated Poisson ZIP) regression using the total number of affiliates of each German firm as the dependent variable. Year fixed effects included. Standard errors in parentheses. ***, **, * = significant at the 1%, 5%, 10%-level.

	(1) Poisson	(2) ZIP	(3) NegBin
Log size t-1 (P)	0.008***	0.012***	0.004***
	(0.001)	(0.002)	(0.000)
Cost efficiency t-1 (P)	-0.006***	-0.009**	-0.001
	(0.002)	(0.004)	(0.001)
Debt ratio t-1 (P)	-0.016***	-0.024***	-0.001
	(0.004)	(0.006)	(0.001)
Log cash flow t-1 (P)	0.007***	0.009***	0.002***
	(0.001)	(0.001)	(0.000)
Fixed asset share t-1 (P)	-0.048***	-0.065***	-0.019***
	(0.007)	(0.009)	(0.002)
Exporter (0/1)	0.009***	0.008***	0.005***
	(0.002)	(0.002)	(0.001)
Observations	71,911	71,911	71,911
Pseudo R^2	0.677	0.322	0.321
log likelihood	-27,438	-20,976	-18,295

An alternative way of looking at the extensive margin of firms' foreign activities is to count the number of foreign affiliates that a given parent holds. Adding an affiliate implies a new set-up cost, hence the count data models presented in Table 3.6 provide information on the determinants of complex FDI strategies. The count data models differ in their assumptions regarding the moments of the distribution and the presence of unobserved individual heterogeneity. These models, therefore, allow controlling for the large share of zeros in our data to a differing degree.[40] The basic count data model is the Poisson model which is quite restrictive in assuming that the conditional mean of the dependent variable equals the conditional variance. The Negative Binomial model allows for unobserved individual heterogeneity and for overdispersion. It is the preferable model, as the equidispersion assumption is strongly rejected for our data. Finally, zero-inflated models assign an even higher weight to the probability of observing a zero in the dependent variable.

Results from count data models support our finding that larger, less indebted parents, firms with a lower share of fixed assets, and firms with higher cash flow are more active internationally. Cost efficiency is negative or insignificant. The debt ratio has a negative impact on the extensive margin when using the number of foreign affiliates. This is consistent with the interpretation of high debt ratios as indicators of low collateral at the parent level which is available to back up new lending.[41]

3.4.3 Intensive margin: Sales of affiliates

We now focus on the sales of the foreign affiliate, while taking the decision to *become* a multinational as well as its location as given. The dependent variable $log(Sales)_{ijk,t}$ are the sales of affiliate j of parent i in country k, and the regression equation includes control variables at the parent level ($Z_{i,t-1}$), at the affiliate level ($Z_{j,t-1}$), and at the country level ($Z_{k,t}$):

$$\log(\text{Sales})_{ijk,t} = \alpha_0 + \alpha_1 Z_{i,t-1} + \alpha_2 Z_{j,t-1} + \alpha_3 Z_{k,t} + \alpha_4 S + \alpha_5 T + \epsilon_{ijk,t} \tag{3.14}$$

We estimate this equation as a parent-level fixed effects model; results are given in Table 3.7. In contrast to the results for the extensive margin, all our parent-level measures for real and financial constraints are insignificant for the intensive margin. Given that most parent characteristics are already absorbed by the fixed effects, variables that capture parents' real and financial constraints do not have an additional impact on the sales of their affiliates. The retained earnings of the affiliate enter with a positive and significant sign in all specifications. Hence, the availability of liquid funds which also reflects the profitability of the affiliate matters for the volume of activities.

Our host-country regressors again yield the expected signs. German firms have larger foreign affiliates in larger countries and in countries hosting many German banks. While the impact of market size per se is not surprising and would, in fact, be borne out by many

[40] For a detailed description of count data models, see, for example, Jones *et al.* (2007).

[41] Naturally, we omit the country-level variables from these regressions.

theoretical models, the positive impact of bank FDI is in support of our theoretical model. A greater presence of home country banks and thus familiarity of domestic lenders with the foreign market should improve the collection of information on the foreign affiliate. This increases the probability that collateral can be collected abroad, thus lowering the cost of financing and increasing the volume of lending.

In columns (2)-(5), we perform similar sample splits by size and sector. The overall findings are very similar with two exceptions. Size (negative) and cash flow (positive) are weakly significant (at the 10%-level) for the large firms. The positive sign on cash flow is consistent with the previous finding that financial constraints matter most for firms with larger foreign activities. Also, within the group of already large firms, the relatively small ones have higher foreign affiliate sales.

Whereas parent-level frictions do not matter for the volume of activities, financial frictions at the affiliate level have an impact on affiliate sales. This is a novel finding since, to the best of our knowledge, the joined impact of parent- and affiliate-level financial frictions has not been analyzed before. These results suggest a hierarchy of financing foreign expansion, where preference is given to local funds and only if they are not sufficient, parent funds are used, albeit at potentially higher opportunity cost.

3.4.4 Heckman selection model

So far, we have treated the decision whether to enter a foreign country and the decision how much to produce and sell separately. To check whether this assumption is justified, we estimate a Heckman selection model, which explicitly accounts for the selection into the FDI mode (Table 3.8). We use state dummies as exclusion restrictions, thus accounting for the fact that - historically - different regions in Germany have different degrees of international openness. Variables measured at the affiliate level and German bank FDI abroad are included in the outcome but not in the selection equation. The Mills ratio in the outcome equation – affiliate sales – is insignificant, which justifies our earlier assumption to model the extensive and the intensive margin separately.

Qualitative results by and large confirm earlier findings. It is interesting to see that some variables affect the probability of setting up an affiliate, but not the volume of its sales. Higher cash flow has a positive impact on the selection into foreign status but not on the volume of sales. This effect is, consistent with the findings reported above, driven by the large firms. Country-level variables such as GDP and GDP per capita have a strong positive impact on the extensive margin, but none on the intensive margin.

Table 3.7: Determinants of the volume of affiliate sales

This table reports results of parent fixed effects panel regressions using the log sales of affiliates of domestic multinational *i* in host country *j* as the dependent variable. (P) = parent-level variables, (A) = affiliate-level variables. In Panel (b), sample splits are at the sample median. Standard errors in parentheses. ***, **, * = significant at the 1%, 5%, 10%-level.

	(1) Full sample	(2) Large	(3) Small	(4) Manufacturing	(5) Services
Log size t-1 (P)	-0.029	-0.128*	0.055	-0.014	-0.078
	(0.048)	(0.076)	(0.061)	(0.060)	(0.093)
Cost efficiency t-1 (P)	-0.112	-0.014	-0.118	-0.479	-0.059
	(0.142)	(0.332)	(0.150)	(0.322)	(0.184)
Debt ratio t-1 (P)	-0.165	-0.072	-0.285	-0.065	-0.100
	(0.277)	(0.404)	(0.404)	(0.375)	(0.483)
Log cash flow t-1 (P)	0.033	0.125*	-0.039	0.033	0.088
	(0.045)	(0.073)	(0.056)	(0.057)	(0.088)
Fixed asset share t-1 (P)	0.009	0.021	0.049	-0.250	-0.212
	(0.217)	(0.262)	(0.434)	(0.322)	(0.424)
Retained earnings / total assets t-1(A)	0.585***	0.447**	0.868***	0.535***	0.777**
	(0.137)	(0.176)	(0.220)	(0.155)	(0.305)
Debt ratio t-1 (A)	-0.086	-0.017	-0.175	-0.092	0.001
	(0.077)	(0.103)	(0.116)	(0.089)	(0.169)
Log GDP	0.139***	0.175***	0.088***	0.138***	0.141***
	(0.018)	(0.024)	(0.027)	(0.020)	(0.038)
Log GDP per capita	0.084***	0.094***	0.056	0.084***	0.094*
	(0.024)	(0.033)	(0.035)	(0.028)	(0.053)
Log bank FDI	0.066***	0.043***	0.097***	0.061***	0.100***
	(0.012)	(0.017)	(0.019)	(0.015)	(0.025)
(Weak) contract enforcement	-0.012***	-0.009*	-0.018***	-0.011***	-0.017**
	(0.004)	(0.005)	(0.005)	(0.004)	(0.007)
Constant	8.531***	8.769***	8.319***	8.827***	8.155***
	(0.337)	(0.627)	(0.441)	(0.542)	(0.619)
Observations	3,507	1,796	1,711	2,363	1,052
R^2	0.134	0.142	0.140	0.138	0.157
Cross-sections	864	283	581	537	341

Table 3.8: Determinants of the volume of affiliate sales

This table reports results of a Heckman selection model using the log sales of affiliates of domestic multinational *i* in host country *j* as the dependent variable. (P) = parent- level variables, (A) = affiliate-level variables. State, sector and year fixed effects included. Sector fixed effects included in the selection equation. Standard errors in parentheses. ***, **, * = significant at the 1%, 5%, 10%-level.

	(1) Full sample		(2) Large		(3) Small	
	Outcome	Selection	Outcome	Selection	Outcome	Selection
Log size t-1 (P)	0.316***	0.144***	0.360***	0.126***	0.006	0.140***
	(0.047)	(0.014)	(0.055)	(0.024)	(0.147)	(0.032)
Cost efficiency t-1 (P)	-0.540***	-0.179***	-0.450***	-0.214***	-0.525***	-0.050
	(0.094)	(0.041)	(0.128)	(0.057)	(0.166)	(0.066)
Debt ratio t-1 (P)	0.047	-0.073	-0.152	-0.200**	0.240	0.071
	(0.113)	(0.056)	(0.168)	(0.083)	(0.225)	(0.084)
Log cash flow t-1 (P)	0.005	0.064***	-0.007	0.096***	0.014	0.019
	(0.028)	(0.013)	(0.040)	(0.018)	(0.052)	(0.020)
Fixed asset share t-1 (P)	-0.688***	-0.517***	-0.903***	-0.479***	0.283	-0.448***
	(0.220)	(0.091)	(0.276)	(0.134)	(0.506)	(0.135)
Log GDP	0.090	0.291***	0.210***	0.299***	-0.594**	0.288***
	(0.081)	(0.009)	(0.077)	(0.013)	(0.260)	(0.014)
Log GDP per capita	-0.017	0.037***	-0.047	0.045**	0.068	0.028
	(0.038)	(0.014)	(0.048)	(0.019)	(0.073)	(0.022)
(Weak) contract enforcement	-0.019***	-0.016***	-0.020***	-0.016***	0.014	-0.017***
	(0.006)	(0.003)	(0.008)	(0.003)	(0.017)	(0.004)
Retained earnings /	0.326**		0.471**		-0.062	
total assets t-1 (A)	(0.155)		(0.220)		(0.215)	
Debt ratio t-1 (A)	-0.044		-0.189*		0.070	
	(0.083)		(0.114)		(0.122)	
Log bank FDI	0.090***		0.088***		0.091***	
	(0.015)		(0.019)		(0.022)	
Constant	5.903***	-5.045***	4.183**	-5.461***	16.814***	-4.832***
	(1.704)	(0.572)	(1.960)	(0.518)	(5.285)	(0.664)
Observations	57,672	57,672	24,196	24,196	33,476	33,476
Censored observations	55,373	55,373	22,804	22,804	32,569	32,569
Mill's ratio	0.166	0.166	0.337	0.337	-2.083	-2.083
Standard error	0.322	0.322	0.297	0.297	1.042	1.042
ρ	0.183	0.183	0.345	0.345	-1.000	-1.000

Some parent-level variables such as cost efficiency (negative), size (positive) and fixed asset share (negative) have a consistent impact on both margins.[42] Affiliate's retained earnings have a strong and significant positive impact on the intensive margin, thus confirming the previous finding that distinguishing parent- and affiliate level frictions is important. (Weak) contract enforcement also influences both margins negatively, as expected. Bank FDI has the expected positive impact on the intensive margin.

Finally, splitting the sample into small and large firms confirms that selection into foreign status is affected by financial constraints for the large firms. Market size has a positive and significant impact on the volume of foreign sales of large firms and a negative impact on sales of small firms. This reflects scale economies and the sorting of smaller firms into smaller markets.

3.4.5 Summing up

Comparing our empirical results to the theoretical predictions summarized in Table 3.1, we find that they are more consistent with the scenario of non-binding than with the scenario of binding collateral constraints. Our measure for the parent's internal funds, cash flow, is consistently significant for the extensive margin, but not for the intensive margin. Our measure for the parent's collateral, the debt ratio, is mostly insignificant at both the extensive and intensive margin, the only exception being the Heckman selection equation for large firms and the count model of affiliates where the coefficient of the parent's debt ratio is significantly negative. The fixed asset share as our measure for fixed cost is significantly negative at the extensive and insignificant at the intensive margin, with the exception of the Heckman outcome equation. Size is always significantly positive for the extensive margin, and, in the Heckman outcome equation, also for the intensive margin. Inconsistent with the model, our alternative measure of productivity (cost efficiency) is frequently insignificant or exhibits the wrong sign. A similar observation has been made by Greenaway *et al*. (2007) who find insignificant coefficients for their measure of productivity (TFP) on firm's export choice, but significantly positive coefficients for size.

3.5 Conclusions

Multinationals are large. Earlier literature focuses on differences in productivity across firms as an explanation for this stylized fact. More productive firms find it easier to

[42] Note that results in Table 3.8 are not fully comparable to those in Table 3.7 since we do not include parent fixed effects in Table 3.8 but state, sector, and year fixed effects.

shoulder the fixed cost of foreign entry, thus being more likely to enter new markets. This chapter analyzes the importance of financial constraints as an additional barrier to entry into foreign markets.

We provide a theoretical model and empirical evidence using data on firms' extensive margin of foreign activities (the probability to be a multinational firm) as well as their intensive margin (the volume of affiliate sales across countries). Considering real barriers to entry as captured by size/productivity and entry cost, we find that larger firms and firms with a smaller share of fixed assets are consistently more likely to become multinationals, and these firms also have larger foreign activities. Cost efficiency, in contrast, does not have the expected positive impact.

Considering financial constraints, our empirical results confirm that these constraints matter for foreign expansions. Parents with larger cash flow are more likely to become multinationals and have more affiliates. For the intensive margin, we find a weaker impact of parent-level financial constraints, but a strong positive impact of affiliate's retained earnings. This suggests a financing hierarchy for the intensive margin, with affiliate financing to be the first and parent financing to be the second choice. Furthermore, considering the interaction of real and financial barriers, financial constraints matter more for large firms because these firms are most likely to expand abroad.

The findings of our essay have a number of implications for different literatures. To the literature of multinational firms, we add a mechanism through which productivity and financial constraints interact. Models ignoring financial constraints would predict that enhancement of firm productivity could improve firms' access to foreign markets. Our results suggest that high productivity may be a necessary, but not a sufficient precondition for foreign expansion. Lowering financial constraints might be just as important, as even large and productive firms are hampered in their internationalization strategy by financial constraints.

To the banking literature, we add a mechanism explaining why banks and non-financial firms typically expand into foreign markets in tandem. One reason for the "follow their customer" patterns in the data could be that home-country banks that are active abroad could have comparative advantages over local banks in enforcing credit repayment and in assessing the creditworthiness of FDI projects. This does not ultimately resolve the "follow their customer" question, but the specific interaction between financial and real barriers to entry that we stress may provide the possibility of testing this link more structurally.

Finally, our findings can have implications for the international macroeconomic literature. Essentially, the financial constraints imbedded in our model are similar to financial accelerator mechanisms used to model the interaction between the financial sector and business cycles. In this sense, extensions of our model might provide useful insights into credit channel mechanisms in open economies and the persistence of shocks triggering entry into foreign markets.

3.6 Mathematical appendix

Proof of Proposition 1

We obtain x^* by taking the first-order condition from (3.6) or (3.6a) respectively, setting it equal to zero and solving for the optimal x^*. To see that $x^* \leq x_{\text{FB}}$, note that $\frac{1+\mu z}{1+z} < 1$ if $\mu < 1$, which is required for a positive collateral to be needed. $\pi^* \leq \pi_{\text{FB}}$ follows directly from $x^* \leq x_{\text{FB}}$ and can be shown analytically by checking that $\pi_{\text{FB}} > \pi^*$ whenever $C > 0$.

Q.E.D.

Proof of Proposition 2

Consider first x^*. It is straightforward to see that:

$$\frac{dx^*}{d\beta} > 0, \frac{dx^*}{dp} > 0, \frac{dx^*}{dF} = \frac{dx^*}{dL} = \frac{dx^*}{d\overline{C}} = 0$$

To obtain the remaining comparative statics, we evaluate first:

$$\frac{dz}{d\theta} = \frac{-(1-q)(1-\mu q)\theta - (1-q)(1-\theta)(1-\mu q)}{(1-\mu q)^2\theta^2} = -\frac{(1-q)}{(1-\mu q)\theta^2} = -\frac{z}{\theta(1-\theta)} < 0$$

and

$$\frac{dz}{d\mu} = \frac{(1-q)(1-\theta)q}{(1-\mu q)^2\theta} = \frac{zq}{(1-\mu q)} > 0$$

Using these derivatives, we obtain:

$$\frac{dx^*}{d\theta} = (1+\beta)qp\frac{(1+z)\mu\frac{dz}{d\theta} - (1+\mu z)\frac{dz}{d\theta}}{(1+z)^2} = -(1+\beta)qp\frac{(1-\mu)\overbrace{\frac{dz}{d\theta}}^{(-)}}{(1+z)^2} > 0$$

and

$$\frac{dx^*}{d\mu} = (1+\beta)qp\frac{(1+z)[\mu\frac{dz}{d\mu} + z] - (1+\mu z)\frac{dz}{d\mu}}{(1+z)^2}$$

$$= (1+\beta)qp\frac{(1+z)z - (1-\mu)\frac{dz}{d\mu}}{(1+z)^2} = (1+\beta)qp\frac{z}{(1+z)^2}[1 + z - \underbrace{\frac{(1-\mu)q}{(1-\mu q)}}_{<1}] > 0$$

Consider next the comparative statics for π^*.

$$\frac{d\pi^*}{d\beta} > 0, \frac{d\pi^*}{dp} > 0, \frac{d\pi^*}{dF} < 0, \frac{d\pi^*}{dL} > 0, \frac{d\pi^*}{d\overline{C}} = 0$$

are straightforward to see. To see that $\frac{d\pi^*}{d\theta} > 0$ and $\frac{d\pi^*}{d\mu} > 0$, note that $\frac{dx^*}{d\theta} > 0$ and $\frac{dx^*}{d\mu} > 0$. Using a revealed preference argument, it follows that the profit has to be increasing in these parameters as well.

Q.E.D.

Proof of Proposition 3

We find the constrained optimal choice of $\overline{x}$ by solving the collateral constraint:

$$\overline{C} + F = \frac{[k(\overline{x}) + F - L] - \mu q p \overline{x}}{(1 - \mu q)\theta}$$

for $\overline{x}$. This gives us a quadratic function of $\overline{x}$ which has the following solutions:

$$\overline{x}_{1/2} = (1 + \beta)\mu q p \pm \sqrt{(1 + \beta)^2 \mu^2 q^2 p^2 - 2(1 + \beta)[F - L - (1 - \mu q)\theta(\overline{C} + F)]}$$

Since we are looking at constrained levels of sales that fall short of the second-best level of sales x^*, the solution for the investor is to choose the larger of the two levels of sales.

Q.E.D.

Proof of Proposition 4

Consider first $\overline{x}$. It is straightforward to see that: $\frac{d\overline{x}}{d\beta} > 0, \frac{d\overline{x}}{dp} > 0, \frac{d\overline{x}}{dF} < 0. \frac{d\overline{x}}{dL} > 0, \frac{d\overline{x}}{d\overline{C}} > 0, \frac{d\overline{x}}{d\theta} > 0$

Finally, note that $\frac{d\overline{x}}{d\mu} > 0$, because increasing μ relaxes the collateral constraint. To see this, note that the right-hand side of:

$$\overline{C} + F \geq \frac{[k(\overline{x}) + F - L] - \mu q p \overline{x}}{(1 - \mu q)\theta}$$

decreases in μ, for a given $\overline{x}$. To see this, note that:

$$\frac{d[\frac{[k(\overline{x})+F-L]-\mu qp\overline{x}}{(1-\mu q)\theta}]}{d\mu} =$$

$$\frac{(1-\mu q)\theta(-qp\overline{x}) - (F + k(\overline{x}) - L - \mu qp\overline{x})(-q\theta)}{(1-\mu q)^2\theta^2} =$$

$$-q\theta\frac{p\overline{x} - (F + k(\overline{x}) - L)}{(1-\mu q)^2\theta^2} < 0$$

To see the comparative statics for $\overline{\pi}$ note that they have the same signs as the comparative statics for $\overline{x}$ because they follow from relaxing (or tightening) the constraints on the constrained choice of $\overline{x}$.

Q.E.D.

Chapter 4

On the use of information in repeated insurance markets*

4.1 Introduction

The vast majority of theoretical models on the insurance market are one-shot models of either perfect competition or monopolistic behavior of insurance companies.[43] However, empirical evidence seems to suggest that in reality, the common market structure in most insurance markets is oligopolistic. Concentration indices for the top 5 insurance companies in the non-life business in Europe in 2002 ranged from 27% in Germany to 89% in Finland (Buzzacchi and Valletti (2005)). Concentration measures in the life insurance sector in most developed nations in the 1990s have been constantly high: even in the USA, the least concentrated market, concentration indices for the top 5 insurance companies have been above 25%, while they have been (high) above 50% in Australia, Canada, Japan and the Netherlands (Bikker and van Leuvensteijn (2008)). Market concentration is also reflected in insurance premiums (Dafny *et al.* (2009)).

Recently, numerous empirical studies have attempted to test the predictions of theoretical models of insurance markets regarding the distribution and use of information in insurance markets and its effects on market outcomes.[44] Several empirical results are hard to reconcile with standard theoretical models, and they suggest more work should be devoted

* This chapter is based on joint work with Heiner Schumacher.

[43] Some of the few exceptions are Ania *et al.* (2002) who re-examine the equilibrium non-existence problem of Rothschild and Stiglitz (1976) in a dynamic setting, and Buzzacchi and Valletti (2005) who provide a model of strategic price discrimination in compulsory insurance markets.

[44] See, for example, Chiappori *et al.* (2006) on general testable implications on insurance markets, Chiappori and Salanie (2008) on empirical issues in modeling competition and market equilibrium in insurance markets, Cutler *et al.* (2008) on preference heterogeneity in insurance markets, Finkelstein and Poterba (2006) on testing for adverse selection using unused observables.

to analyzing imperfectly competitive models of insurance markets – or, in the words of Chiappori *et al*. (2006), "there is a crying need for such models".

In this chapter, we analyze a repeated oligopolistic insurance market. The main feature of our model is that insurance companies take into account the impact of pricing decisions on competitors' actions. We think that analyzing repeated interaction is crucial to understand the use of information in insurance markets for the following reasons:

First, our model provides an explanation for the puzzle of "unused observables" that has been tested in several empirical papers, but that has not received interest in the theoretical literature. In theory, profit maximizing insurance companies should exploit any risk-relevant information available to them.[45] However, there is evidence of unused observables in insurance markets, that is information which a) insurance companies collect or could collect, b) is correlated with the risk experience, but c) is not used by companies to set prices. For example, according to Finkelstein and Poterba (2006), the address of the insured person is almost always collected, but seldom used in pricing insurance, although there is a correlation between geographic information and other individual attributes that affects both the demand for insurance and the risk type. They use data on annuity purchases in the UK to illustrate that the information on the annuitant's residential location would help to predict future mortality risk, but that it does not influence the insurance premium. Gender is another example of an unused observable that is usually collected by default, but that is not used for pricing in certain insurance markets, the most prominent example being the the long-term care insurance market and the automotive insurance. In both markets, the expected costs for the insurer differ substantially for men and women.[46] Further empirical evidence on unused observables is provided by Brown and Finkelstein (2007) (gender and place of residence in the U.S. long-term care insurance industry) and Ivaldi (1996) (smoking in the French automobile insurance industry). Finkelstein and Poterba (2006) conclude their article by stating that "a complete understanding of the limited use in pricing of available or collectible risk-related information on insurance buyers remains an open issue".[47]

[45] Under perfect competition, companies will use all information in order to charge the fair premium. A monopolist will use all information in order to maximize profits through price discrimination.

[46] See, for example, Finkelstein and Poterba (2006) or the "Gesamtverband der deutschen Versicherungswirtschaft", (www.gdv.de), the association of German insurance companies.

[47] They mention a number of possible reasons for the existence of unused observables, e.g. regulation or implementation costs, but show that these cannot fully explain the puzzle.

Second, the information available to insurance companies and the correlation with the underlying risk of the insured is subject to constant change. The revolution in information technologies has enabled insurance companies to collect, analyze and make use of large amounts of information. An example of evolving information that has recently received a lot of attention is that of genetic testing (Hoy and Witt (2007), Hoy and Polborn (2000), Rees and Apps (2006)). Correlation like the one between an insured's address and socioeconomic status may change over time as the composition of residents in a certain area changes. It is therefore important to know how insurance companies respond to a constant change in their information about risks.

Third, in an oligopolistic insurance market, the existence of bulk discounts can also be rationalized if companies collude and therefore make positive profits. In competitive insurance markets with asymmetric information, high risk individuals will demand larger quantities of insurance than low-risk individuals. In order for an insurance company to break even, theory predicts that marginal prices should rise with quantity. However, in reality, many insurance companies offer discounts in bulk (Cawley and Philipson (1999) or Chiappori *et al*. (2006)).

In our model, there are two types of individuals who face either high or low risk of damage. There is a finite number of insurance companies. They can distinguish between these risks and compete for customers by setting insurance premia in each period. Companies interact strategically and preconceive the effect of their pricing decision on the prices set by their competitors in subsequent periods. If companies fear a price war after adjusting their prices, they may refrain from doing so. We show that even if insurance companies can distinguish between risk types, equilibria exist in which (1) insurance companies charge the same insurance premium to both risks, and (2) both risk types purchase positive amounts of insurance (however, low risks potentially acquire less insurance than high risks). Thus, we derive an equilibrium with unused observables. Furthermore, if companies make positive profits out of all risks, it may be rational for them to offer bulk discounts.

We then show that the same equilibrium outcome is possible if insurance companies cannot distinguish between high and low risks, i.e. if there is asymmetric information. This renders possible the following explanation for the existence of unused observables: initially, there exists a collusive equilibrium in an insurance market with asymmetric information. Then, after analyzing their data, insurance companies learn how to distinguish between high and low risks. However, they maintain their pricing schedules in order to sustain collusion.

In the next step, we analyze the robustness of our model with regard to two extensions. First, we allow for market entry.[48] Outside firms can enter the market incurring some entry costs and become incumbent firms for the rest of the game. Second, we allow for explicit collusion between firms, i.e. they can agree on charging the profit-maximizing insurance premia for low and high risks. If they can negotiate with each other, companies are likely to exploit their information. We show that if entry costs are neither too high nor too low, there exist equilibria with unused observables in which incumbent companies cannot gain by explicit collusion. The intuition for this result is as follows: if incumbent companies decide to increase their period profits by charging different premia for low and high risks, outside companies can enter the market profitably by making a one-shot gain. If on the other side one incumbent company undercuts the insurance premium of its competitors, it triggers a price war, which wipes out all gains of this deviation. We therefore show that equilibria with unused observables can be robust to explicit collusion and to the threat of market entry. In these equilbria, it does not pay off to use the information about risks as the maximal level of per-period profits can already be attained without this information.

At a technical level, our essay also contributes to the literature of third-degree price discrimination.[49] In the case where insurance companies are able to distinguish between high- and low-risk costumers, they have to decide whether to charge the same or different prices to two groups of costumers who differ in their willingness to pay. This literature has introduced the possibility of firms' competition, but it has not analyzed equilibrium strategies in the repeated game.

The rest of the chapter is organized as follows: The next section outlines the basic model and derives an equilibrium with unused observables. We include explicit collusion and market entry into the model in section 4.3. In section 4.4, we provide further examples of the evolution of the use of information in insurance industries, and discuss welfare and policy implications. The last section concludes.

[48] Bikker and van Leuvensteijn (2008) show empirically that market entry is indeed a relevant phenomenon in insurance markets in the countries analyzed in their study (Canada, Germany, Japan, the Netherlands, and the UK).

[49] See Tirole (1988), chapter 3.2 for a summary of this issue.

4.2 A repeated insurance market

4.2.1 Framework

Time is discrete and denoted by $t \in \{0, 1, ...\}$. The stage game is the simplest version of an insurance market.[50] In each period, there is a continuum of customers of mass 1. These can be the same customers or different ones in each period.[51] Each customer has wealth W in each period, and faces the risk of losing an amount of $d < W$. She may have either a high-risk probability of π_H or a low-risk probability $\pi_L < \pi_H$. Let λ be the fraction of high-risk individuals. All customers have the same von Neuman-Morgenstern utility function $U(W)$.[52] We assume that $U(W)$ is twice continuously differentiable with $U'(W) > 0$ and $U''(W) < 0$.

There are $N > 1$ long-lived risk-neutral insurance companies in the market. Let $I = \{1, ..., N\}$ be the set of insurance companies. First, we will assume that these companies can distinguish between high- and low-risk customers.[53] At a later stage, we will turn to the case of asymmetric information. In each period, each company $i \in I$ offers any positive amount of insurance.[54] Let $\alpha_H^{i,t}$ ($\alpha_L^{i,t}$) be the insurance premium for high-risk (low-risk) individuals offered by company i in period t. If an individual of risk $j \in \{L, H\}$ purchases an insurance cover $D_j \geq 0$ in period t from company i, she pays $D_j \alpha_j^{i,t}$ to the company in this period, regardless of whether damage occurs or not. If the damage

[50] As formalized, for example, in Rees and Wambach (2008).

[51] There may also be a certain in- and outflow of individuals in each period. As long as not all customers are locked into a specific contract for all periods, the results of the model do not change.

[52] This only simplifies the exposition of the model. All of our results would also hold if customer are heterogeneous in their risk aversion (the only thing we need is that customers are risk averse to some extent). Note that in standard models of the insurance market, the assumption of a uniform utility function is not innocuous, see Smart (2000).

[53] An alternative interpretation would be that insurance companies cannot perfectly distinguish risks, but have imperfect information (variables which are imperfectly correlated with risk types) which can be used to categorize risks, as in Hoy (1982).

[54] We thereby assume non-exclusive provision of insurance which is different from most insurance market models, where companies offer price-quantity combinations (such as Rothschild and Stiglitz (1976)). The assumption of non-exclusivity is not crucial for some markets such as the life insurance market (see Polborn *et al.* (2006)). However, for property-liability insurance exclusivity it is more natural (we are indebted to Michael Hoy for pointing out this fact to us). In terms of our model, non-exclusivity is not needed for the results on the insurance market with symmetric information in sections 4.2.2 and 4.3. For the results on the insurance market with asymmetric information in section 4.2.3 it is needed for our characterization of equilibrium outcomes, but not for the existence of collusive equilibria.

occurs, she gets D_j from the company, i.e. $D_j = d$ implies full coverage. We say that company i uses the information about risks in period t if $\alpha_H^{i,t} \neq \alpha_L^{i,t}$.

Customers are not modelled as strategic players: in each period, they purchase the utility maximizing insurance cover from the company that offers at the cheapest premium for their risk. If more than one company has the lowest insurance premium, each customer randomizes with equal probability from which company she buys insurance. The sequence of events in each period t is as follows:

1. Insurance companies announce the insurance premia $\left\{\left(\alpha_L^{i,t}, \alpha_H^{i,t}\right)\right\}_{i \in I}$.

2. Customers purchase insurance D_L and D_H.

3. Nature decides about the occurrence of damage and payoffs are realized.

Now fix

$$\alpha_L^{i,-1} = \alpha_H^{i,-1} = 0 \tag{4.1}$$

for all $i \in I$. For $t \in \{0, 1, ...\}$, we denote by h_t the history of all insurance premia that were charged by all insurance companies up to period t :

$$h_t = \left(\left\{\left(\alpha_L^{i,-1}, \alpha_H^{i,-1}\right)\right\}_{i \in I}, \left\{\left(\alpha_L^{i,0}, \alpha_H^{i,0}\right)\right\}_{i \in I}, ..., \left\{\left(\alpha_L^{i,t-1}, \alpha_H^{i,t-1}\right)\right\}_{i \in I}\right). \tag{4.2}$$

The set of all possible histories at date t will be denoted by H_t. A strategy of company i is an infinite sequence of action functions $\alpha^{i,t}$ for every $t \in \{0, 1, ...\}$, where $\alpha^{i,t}$ determines $\alpha_L^{i,t}$ and $\alpha_H^{i,t}$ as a function of the history h_t:

$$\alpha^{i,t} : H_t \rightarrow \mathbb{R}^2. \tag{4.3}$$

Without loss of generality we concentrate on pure strategies. The strategies of companies determine the sequence of insurance premia

$$\left\{\left\{\left(\alpha_L^{i,t}, \alpha_H^{i,t}\right)\right\}_{i \in I}\right\}_{t=0}^{\infty}. \tag{4.4}$$

From this sequence, we can derive the profit $G^{i,t}$ of company i in period t. Insurance companies discount future gains by δ. The sum of normalized discounted profits of company i is then given by

$$G^i = (1 - \delta) \sum_{t=0}^{\infty} \delta^t G^{i,t}. \tag{4.5}$$

The insurance market is in equilibrium if no company i can increase its profit G^i by choosing unilaterally another strategy.

4.2.2 Equilibria in an oligopolistic insurance market with symmetric information

We start by analyzing the demand for insurance. Assume for a moment that company i offers the lowest insurance premium to individuals with risk $j \in \{L, H\}$ in period t. A customer with risk j purchases the utility maximizing amount of insurance:

$$\tilde{D}_j(\alpha_j^{i,t}) = \arg\max D_j \pi_j U\left(W - d + D_j(1-\alpha_j^{i,t})\right) + (1-\pi_j)\, U\left(W - D_j \alpha_j^{i,t}\right). \tag{4.6}$$

This demand is implicitly given by

$$\frac{U'\left(W - d + \tilde{D}_j(\alpha_j^{i,t})(1-\alpha_j^{i,t})\right)}{U'\left(W - \tilde{D}_j(\alpha_j^{i,t})\alpha_j^{i,t}\right)} = \frac{\alpha_j^{i,t}\,(1-\pi_j)}{\left(1-\alpha_j^{i,t}\right)\pi_j}. \tag{4.7}$$

As U' is continuous, $\tilde{D}_j(\alpha_j^{i,t})$ must also be continuous. The fair insurance premium under which the customer purchases full coverage is given by

$$\alpha_j^f = \pi_j, \tag{4.8}$$

while the highest insurance premium such that the customer is indifferent between purchasing a marginal unit of insurance cover or not is uniquely defined by

$$\alpha_j^{\max} = \frac{\pi_j}{(1-\pi_j)\frac{U'(W)}{U'(W-d)} + \pi_j}. \tag{4.9}$$

For insurance premia $\alpha_j^{i,t} \in \left(\alpha_j^f, \alpha_j^{\max}\right)$ demand is positive and company i earns a positive profit from contracts with individuals of risk π_j. For higher insurance premia, profits are 0, for lower insurance premia, profits are negative. Note that if π_L is sufficiently close to π_H, we have $\alpha_H^f < \alpha_L^{\max}$. Define

$$\tilde{\alpha}_j \in \arg\max \alpha_j \in \mathbb{R}(\alpha_j - \pi_j)\tilde{D}_j(\alpha_j). \tag{4.10}$$

This is well-defined, as $(\alpha_j - \pi_j)\tilde{D}_j(\alpha_j)$ is continuous on the interval $\left[\alpha_j^f, \alpha_j^{\max}\right]$ and therefore attains its maximum.

Assume now that only company i offers the lowest premium for customers of both types, but does not use the information about the insured risk, i.e. $\alpha_L^{i,t} = \alpha_H^{i,t} = \alpha_P^{i,t}$, where $\alpha_P^{i,t}$ is called the "pooling premium". We then have

$$G^{i,t} = \lambda \tilde{D}_H(\alpha_P^{i,t})\left(\alpha_P^{i,t} - \pi_H\right) + (1-\lambda)\,\tilde{D}_L(\alpha_P^{i,t})\left(\alpha_P^{i,t} - \pi_L\right). \tag{4.11}$$

Denote by α_{P0} the pooling premium at which the right-hand side of (4.11) is equal to 0 such that $\alpha_P^{i,t} \in (\alpha_{P0}, \alpha_H^{\max})$ implies positive demand for insurance by at least the high-risk individuals and positive profits from the pooling contract. Note that $\alpha_H^f < \alpha_L^{\max}$ implies $\alpha_L^f < \alpha_{P0} < \alpha_H^f$. We then can state our first main result:

Proposition 1 *(a) For each $\alpha \in (\alpha_{P0}, \alpha_H^{\max})$ there is a $\delta(\alpha) < 1$ such that there exists a subgame-perfect equilibrium in which $\alpha_P^{i,t} = \alpha$ for all companies $i \in I$ and in all periods t if $\delta \geq \delta(\alpha)$. (b) If $\alpha \notin (\alpha_{P0}, \alpha_H^{\max})$, then no such equilibrium exists. (c) We have that* $\lim_{\alpha \downarrow \alpha_{P0}} \delta(\alpha) = 1$.

Proof. (a) Consider the following simple grim-trigger strategy which is played by all companies $i \in I$: charge $\alpha_P^{i,0} = \alpha$. In period $t > 0$, charge $\alpha_P^{i,t} = \alpha$ if and only if $\alpha_L^{l,\tau} = \alpha_H^{l,\tau} = \alpha$, $l \in I$, in all periods $\tau \in \{0, ..., t-1\}$. Otherwise, charge $\alpha_L^{i,t} = \alpha_L^f$ and $\alpha_H^{i,t} = \alpha_H^f$. We employ the one stage deviation principle in order to show that this can be an equilibrium. If at least one company charges the fair insurance premia α_L^f and α_H^f, no other company can make positive profits. Thus, the maximal normalized discounted profit from a deviation of company i in period t is given by

$$G^{i,d} = (1-\delta)\left[\lambda(\tilde{\alpha}_H - \pi_H)\tilde{D}_H(\tilde{\alpha}_H) + (1-\lambda)(\tilde{\alpha}_L - \pi_L)\tilde{D}_L(\tilde{\alpha}_L)\right]. \tag{4.12}$$

The normalized discounted profit from compliance is given by

$$G^{i,c} = \frac{1}{N}\left[\lambda(\alpha - \pi_H)\tilde{D}_H(\alpha) + (1-\lambda)(\alpha - \pi_L)\tilde{D}_L(\alpha)\right]. \tag{4.13}$$

As $\alpha \in (\alpha_{P0}, \alpha_H^{\max})$, this term is positive. Thus, if δ is sufficiently close to unity, we have $G^{i,c} \geq G^{i,d}$. (b) If $\alpha_L^f < \alpha \leq \alpha_{P0}$ and $\alpha_P^{i,t} = \alpha$ for all $i \in I$ and in all periods t, a single company l could increase its normalized discounted profit by charging $\alpha_L^{l,0} = \alpha$, $\alpha_H^{l,0} = \alpha_H^f$ and the fair premia thereafter. If $\alpha \leq \alpha_L^f$ and $\alpha_P^{i,t} = \alpha$ for all $i \in I$ and in all periods t, a single company l could increase its normalized discounted profit by charging $\alpha_L^{l,t} = \alpha_H^f$ and $\alpha_H^{l,t} = \alpha_H^f$ in all periods t. If $\alpha_H^{\max} \leq \alpha$ and $\alpha_P^{i,t} = \alpha$ for all $i \in I$ and in all periods t, a single company l could increase its normalized discounted profit by charging $\alpha_L^{l,0} = \tilde{\alpha}_H$, $\alpha_H^{l,0} = \tilde{\alpha}_L$ and the fair premia thereafter. (c) Assume that an equilibrium exists in which $\alpha_P^{i,t} = \alpha$ for all $i \in I$ and in all periods t. Then the normalized discounted profit for each firm is equal to the term in (4.13). Note that this term converges to 0 as α approaches α_{P0} from above. The maximal normalized discounted profit of a deviating firm can be at least

$$(1-\delta)(1-\lambda)(\alpha_{P0} - \pi_L)\tilde{D}_L(\alpha_{P0}) > 0. \tag{4.14}$$

Thus, we must have $\lim_{\alpha \downarrow \alpha_{P0}} \delta(\alpha) = 1$. ■

In the equilibria of proposition 1, insurance companies fear a price war if they change their insurance premia. Thus, they maintain a pooling premium, which guarantees them positive profits. This situation exhibits the following features:

- Although companies have more detailed information about risks, they do not use it. Thus, we have an equilibrium with unused observables.
- Given that π_L is sufficiently close to π_H, both low- and high-risk individuals purchase positive amounts of insurance. However, there is adverse selection: as we can derive from equation (4.7), low-risk customers purchase less insurance than high-risk customers.

4.2.3 Equilibria in an oligopolistic insurance market with asymmetric information

We now turn to the case where insurance companies cannot distinguish between high- and low-risk individuals and show that a result similar to proposition 1 holds. Because of asymmetric information, each firm $i \in I$ only charges a pooling premium $\tilde{\alpha}_P^{i,t}$ in period t and customers purchase the amount of insurance which maximizes their expected utility from the firm that charges the lowest insurance premium. Fix

$$\tilde{\alpha}_P^{i,-1} = 0 \tag{4.15}$$

for all $i \in I$. For $t \in \{0, 1, ...\}$, we denote by $\tilde{h}_t$ the history of all insurance premia that were charged by all insurance companies up to period t :

$$\tilde{h}_t = \left(\left\{ \tilde{\alpha}_P^{i,-1} \right\}_{i \in I}, \left\{ \tilde{\alpha}_P^{i,0} \right\}_{i \in I}, ..., \left\{ \tilde{\alpha}_P^{i,t-1} \right\}_{i \in I} \right). \tag{4.16}$$

The set of all possible histories at date t will be denoted by $\tilde{H}_t$. A strategy of company i is an infinite sequence of action functions $\tilde{\alpha}^{i,t}$ for every $t \in \{0, 1, ...\}$, where $\tilde{\alpha}^{i,t}$ determines $\tilde{\alpha}_P^{i,t}$ as a function of the history $\tilde{h}_t$:

$$\tilde{\alpha}^{i,t} : \tilde{H}_t \to \mathbb{R}. \tag{4.17}$$

Again, we concentrate on pure strategies. The strategies of companies determine the sequence of insurance premia

$$\left\{ \left\{ \tilde{\alpha}_P^{i,t} \right\}_{i \in I} \right\}_{t=0}^{\infty} \tag{4.18}$$

from which the sum of normalized discounted profits $\tilde{G}^i$ can be calculated as in the last subsection. The rest of the model remains unchanged. Define

$$A_P = \left\{ \tilde{\alpha}_P \mid \tilde{\alpha}_P \in \arg\max \alpha \in \mathbb{R} \lambda (\alpha - \pi_H) \tilde{D}_H(\alpha) + (1 - \lambda)(\alpha - \pi_L) \tilde{D}_L(\alpha) \right\}, \tag{4.19}$$

which is the set of pooling-premia, such that the maximal gain from pooling contracts is attained and

$$\tilde{\alpha}_P^* = \min \left\{ \tilde{\alpha}_P \in A_P \right\}, \tag{4.20}$$

which is the smallest element in this set. We then can derive:

Proposition 2 *Assume that companies cannot distinguish between high- and low risks. For each $\alpha \in [\alpha_{P0}, \alpha_H^{\max})$ there is a $\tilde{\delta}(\alpha) < 1$ such that there exists a subgame-perfect equilibrium in which $\tilde{\alpha}_P^{i,t} = \alpha$ for all companies $i \in I$ and in all periods t if $\delta \geq \tilde{\delta}(\alpha)$. (b) If $\alpha \notin [\alpha_{P0}, \alpha_H^{\max})$, then no such equilibrium exists. (c) If $\alpha > \alpha_{P0}$ is sufficiently small, then $\tilde{\delta}(\alpha) = 1 - \frac{1}{N}$.*

Proof. Observe that one equilibrium of the stage game is $\tilde{\alpha}_P^i = \alpha_{P0}$ for all $i \in I$, therefore it is an equilibrium if $\tilde{\alpha}_P^{i,t} = \alpha_{P0}$ for all $i \in I$ and in all periods t. In this equilibrium, all companies make zero-profits. Assume now that $\alpha \in (\alpha_{P0}, \alpha_H^{\max})$. Consider the following simple grim-trigger strategy which is played by all companies $i \in I$: charge $\tilde{\alpha}_P^{i,0} = \alpha$. In period $t > 0$, charge $\tilde{\alpha}_P^{i,t} = \alpha$ if and only if $\tilde{\alpha}_P^{l,\tau} = \alpha$, $l \in I$, in all periods $\tau \in \{0, ..., t-1\}$. Otherwise, charge $\tilde{\alpha}_P^{i,t} = \alpha_{P0}$. If at least one company charges α_{P0}, no other company can make positive profits. Thus, the maximal normalized discounted profit from a deviation of company i in period t is given by

$$G^{i,d} = (1-\delta)\left[\lambda(\tilde{\alpha}_P^* - \pi_H)\tilde{D}_H(\tilde{\alpha}_P^*) + (1-\lambda)(\tilde{\alpha}_P^* - \pi_L)\tilde{D}_L(\tilde{\alpha}_P^*)\right]. \tag{4.21}$$

The normalized discounted profit from compliance is given by

$$G^{i,c} = \frac{1}{N}\left[\lambda(\alpha - \pi_H)\tilde{D}_H(\alpha) + (1-\lambda)(\alpha - \pi_L)\tilde{D}_L(\alpha)\right]. \tag{4.22}$$

As $\alpha \in (\alpha_{P0}, \alpha_H^{\max})$, this term is positive. Thus, if δ is sufficiently close to unity, we have $G^{i,c} \geq G^{i,d}$. (b) If $\alpha < \alpha_{P0}$ and $\tilde{\alpha}_P^{i,t} = \alpha$ for all $i \in I$ and in all periods t, a single company l could increase its normalized discounted profit by charging $\tilde{\alpha}_L^{l,t} = \alpha_{P0}$ in all periods t. If $\alpha_H^{\max} \leq \alpha$ and $\tilde{\alpha}_P^{i,t} = \alpha$ for all $i \in I$ and in all periods t, a single company l could increase its normalized discounted profit by charging $\tilde{\alpha}_L^{l,0} = \tilde{\alpha}_P^*$ and $\tilde{\alpha}_L^{l,\tau} = \alpha_{P0}$ in all periods τ thereafter. (c) Define

$$G(\alpha) = \lambda(\alpha - \pi_H)\tilde{D}_H(\alpha) + (1-\lambda)(\alpha - \pi_L)\tilde{D}_L(\alpha) \tag{4.23}$$

and observe that $G(\alpha)$ is continuous on the interval $(\alpha_{P0}, \alpha_H^{\max})$ with $G(\alpha_{P0}) = 0$. Thus, there must exist an $\hat{\alpha}$ with $\alpha_{P0} < \hat{\alpha} \leq \tilde{\alpha}_P^*$, such that $G(\alpha)$ strictly increases in the interval $(\alpha_{P0}, \hat{\alpha})$. Consider the same strategy as in part (a) for $\alpha \in (\alpha_{P0}, \hat{\alpha})$. As before, the normalized discounted profit from compliance is then $\frac{1}{N}G(\alpha)$, while the normalized discounted profit from a deviation is $(1-\delta)G(\alpha)$. Consequently, the critical discount factor equals $1 - \frac{1}{N}$. ■

The results of proposition 1 and proposition 2 enable the following interpretation: the arrival of new information about risks does not necessarily change the equilibrium outcome. Suppose that the market is in an equilibrium with asymmetric information in which all insurance companies charge a premium of $\tilde{\alpha} \in (\alpha_{P0}, \alpha_H^{\max})$ in each period (proposition 2 says that this is possible). If in that situation new information about risks arrives (as a result of data-collection, for example), then, by proposition 1, the market can enter an equilibrium with symmetric information and with exactly the same equilibrium outcome (given that δ is sufficiently high). This is especially relevant if collusion must be tacit and companies avoid explicit negotiations. In particular, we have shown that almost every equilibrium outcome under asymmetric information is also an equilibrium outcome under symmetric information.

Whether an equilibrium outcome survives the arrival of new information about risks or not, might depend on the extent of collusion: observe from proposition 1 (c) and proposition 2 (c) that if profits are low under asymmetric information (i.e. if $\tilde{\alpha}_P^{i,t}$ is close to α_{P0} for all i), the critical discount factor is just $1 - \frac{1}{N}$, while it is very close to 1 under symmetric information. Thus, if profits are low, then the arrival of new information potentially triggers a price war and a change of the equilibrium outcome. Compared to that, the critical discount factor equals $1 - \frac{1}{N}$ under both asymmetric and symmetric information if $\tilde{\alpha}_P^{i,t} = \tilde{\alpha}_P^*$ for all i (which can be shown by going through the same steps as in the proof of proposition 2 (c)), i.e. if profits are relatively high.

Whenever insurance companies can make expected profits out of all contracts, it is not difficult to imagine a situation when they do so by granting bulk discounts (instead of linear pricing) to customers, as discovered by Cawley and Philipson (1999). This especially makes sense when firms face administrative fixed costs per contract, such that selling more insurance to some risks increases the expected profit per unit of insurance.

4.3 Explicit collusion and market entry

The equilibria in the last section had a number of attributes that are inconsistent with the results of one-shot models of the insurance market, but consistent with empirical results. However, there remain two important issues:

- If an industry makes profits, we would expect market entry.
- If companies are able to sustain collusion, they should be able to increase their profits even further by using the information about risks, i.e. they may coordinate on an equilibrium in which all firms discriminate between risks and charge $\alpha_H^{i,t} = \tilde{\alpha}_H$ and $\alpha_L^{i,t} = \tilde{\alpha}_L$ in all periods t.[55]

[55] Here we implicitly assume that $\tilde{\alpha}_H \neq \tilde{\alpha}_L$ which is true for most standard utility-functions.

We will deal with both questions in this section and show that the equilibria of proposition 1 still can be robust against market entry and explicit collusion. In all what follows, we will again consider a scenario with symmetric information. Note that explicit collusion is illegal in most legislations and tacit collusion (i.e. collusion without communication between firms) hard to detect.[56] We will not rely on this, but assume that firms can negotiate without being exposed to the danger of punishment.

Denote the set of incumbent companies in each period by I_t, where

$$I_0 = \{1, ..., N\}.$$

In each period t, there is an infinite number of outside firms $k \in \mathbb{N} \backslash I_t$ which can enter the market at cost $c > 0$.[57] These entry costs can be interpreted as installation costs, for example, the costs of acquiring the necessary distribution channels. If an outside company enters the market, it belongs to the set of incumbents in all future periods and can distinguish between high and low risks.[58] As tie-breaking rule we define that a company only enters the market if it can make strictly positive profits. Furthermore, we define:

Definition *An equilibrium is robust against explicit collusion if there is no other weakly pareto superior equilibrium for incumbent companies.*

If an equilibrium in which companies $i \in I_0$ do not use the information about risks is robust against explicit communication, any agreement on adjusting insurance premia to increase profits must result in a decrease of profits for at least one incumbent company, and therefore would not be accepted by this company.[59]

We assume that in each period, outside companies observe the insurance premia charged by incumbent companies and then decide about whether to enter the market or not. Thus,

[56] For a discussion about the difference between explicit and tacit collusion, see Rees (1993).

[57] This also could be insurance companies which offer the same insurance contracts, but at substantially higher rates, such that only a small fraction of uninformed consumers (who do not compare insurance premia, but randomly choose some contract) purchases those contracts.

[58] This particular model of market entry was introduced by Harrington (1989).

[59] Note that robustness against explicit collusion is weaker than (weak) renegotiation proofness. For details about renegotiation proofness, see Mailath and Samuelson (2006), pages 134 - 143. One also could construct weak renegotiation proof equilibria in our setting, however, their structure is not interesting for our purpose.

incumbent companies are Stackelberg leaders and market entry is endogenous as in Etro (2008). The sequence of events now is as follows:

1. Insurance companies announce the insurance premia $\left\{\left(\alpha_L^{i,t}, \alpha_H^{i,t}\right)\right\}_{i\in I_t}$.
2. Outside companies decide whether to enter the market at cost c or not. If a company $k \in \mathbb{N}\backslash I_t$ enters the market, it subsequently sets its insurance premia $\left(\alpha_L^{k,t}, \alpha_H^{k,t}\right)$.
3. Costumers purchase insurance D_L and D_H.
4. Nature decides about the occurrence of damage and payoffs are realized.
5. If a company $k \in \mathbb{N}\backslash I_t$ has entered the market, then $I_{t+1} = I_t \cup \{k\}$.

Clearly, as entry costs are positive, incumbents can price outside companies out of the market. However, if entry costs are small, then per-period profits also must be small. If these profits are generated by charging a pooling premium such that low risks subsidize high risks, then it can be profitable for an outside firm to enter the market and to make a one-shot gain by offering contracts only to low risks. We therefore get:

Proposition 3 *If c is sufficiently small, then in equilibrium all companies that make positive profits in period t use the information about risks in this period.*

Proof. Assume that this is not the case and an incumbent company $i \in I_t$ charges $\alpha_P^{i,t} \in (\alpha_{P0}, \alpha_H^{\max})$ and makes a positive profit in period t. If follows that $\alpha_j^{l,t} \geq \alpha_P^{i,t}$ for all $l \in I_t$ and $j \in \{L, H\}$. If an outside company $k \in \mathbb{N}\backslash I_t$ enters the market, then it earns at least

$$-c + (1-\lambda)(\alpha_{P0} - \pi_L)\tilde{D}_L(\alpha_{P0}), \tag{4.24}$$

by charging $\alpha_L^{k,t} = \alpha_{P0}$ and $\alpha_H^{k,t} = \alpha_H^f$, given that there is no other outside company which enters the market. The term in (4.24) is positive if c is sufficiently low. Therefore, the situation outlined above cannot be an equilibrium outcome if c is sufficiently low. ■

Thus, the equilibria of proposition 1 are not robust against market entry, if entry costs are sufficiently small. However, we do not expect entry barriers to be negligible for insurance markets. If c is sufficiently high, the existence of equilibria with pooling premia might be restored. Define $\tilde{\alpha}_P^*$ as in the last section and denote

$$G^{high} = \lambda(\tilde{\alpha}_P^* - \pi_H)\tilde{D}_H(\tilde{\alpha}_P^*) + (1-\lambda)(\tilde{\alpha}_P^* - \pi_L)\tilde{D}_L(\tilde{\alpha}_P^*), \tag{4.25}$$

$$G^{low} = \max \alpha \in \left[\alpha_L^f, \alpha_H^f\right](1-\lambda)(\alpha - \pi_L)\tilde{D}_L(\alpha). \tag{4.26}$$

G^{high} is the highest period profit from a pooling contract, G^{low} is the highest period profit that can be made by selling contracts only to low risks and by charging a premium in the interval $\left[\alpha_L^f, \alpha_H^f\right]$. For $\pi_H \to \pi_L$, we have $\alpha_H^f \to \alpha_L^f$, such that $G^{low} \to 0$. Thus, if π_H is sufficiently close to π_L, then $G^{high} > G^{low}$. We then can show:

Proposition 4 *Assume that $\delta > 1 - \frac{1}{N}$ and $G^{high} > G^{low}$. If $c \in \left(G^{low}, G^{high}\right)$, then there is a subgame-perfect equilibrium which is robust against explicit collusion and in which $\alpha_P^{i,t} = \alpha$, $\alpha \in \left(\alpha_H^f, \tilde{\alpha}_P^*\right)$, for all incumbent companies $i \in I_0$ in all periods t, while outside firms do not enter the insurance market.*

Proof. Define for $G \in \left(G^{low}, G^{high}\right)$

$$\alpha^G = \min_{\alpha} \left\{ \alpha \in \left(\alpha_H^f, \tilde{\alpha}_P^*\right) \mid \lambda(\alpha - \pi_H)\tilde{D}_H(\alpha) + (1-\lambda)(\alpha - \pi_L)\tilde{D}_L(\alpha) = G \right\}. \tag{4.27}$$

Fix a value $G^* \in \left(G^{low}, G^{high}\right)$. Assume that in each period, incumbent companies play a grim-trigger strategy that also deters entry: Charge $\alpha_P^{i,t} = \alpha^{G^*}$ if and only if $\alpha_L^{l,\tau} = \alpha_H^{l,\tau} = \alpha^{G^*}$, $l \in I_0$, and $I_\tau = I_0$ in all periods $\tau \in \{0, ..., t-1\}$. Otherwise, charge $\alpha_L^{i,t} = \alpha_L^f$ and $\alpha_H^{i,t} = \alpha_H^f$. We show that this strategy can support an equilibrium. If a company $i \in I_0$ undercuts α^{G^*} in period t, then the definition of α^G, the continuity of $\alpha\tilde{D}_j(\alpha)$ and the fact that $G^* > G^{low}$ ensure that $G^{i,t} < G^*$. Given that no outside company ever enters the market, an incumbent company complies to this strategy if

$$\frac{1}{N} G^* > (1-\delta) G^*, \tag{4.28}$$

which is equivalent to

$$\delta > 1 - \frac{1}{N}. \tag{4.29}$$

The tie-breaking rule implies that an outside company will not enter if and only if

$$G^* \leq c. \tag{4.30}$$

Thus, if $c \in \left(G^{low}, G^{high}\right)$ and (4.29) holds, then an equilibrium with no market entry, $G^* = c$, $\alpha_P^{i,t} = \alpha^{G^*}$ for all $i \in I_0$ and all t exists and is robust against explicit collusion. ■

The logic of these equilibria is again simple. As incumbent companies play a grim-trigger strategy, they refrain from changing their pricing schedule. The punishment is also triggered if an outside company enters the market. Therefore, the period profit is limited to entry costs, otherwise it would pay off for an outside company to enter the market and make a one-shot gain. Therefore, incumbent companies cannot coordinate on insurance premia, such that they earn strictly higher profits.

The upper bound on entry costs, G^{high}, ensures that a period profit equal to $\frac{1}{N}c$ per incumbent company can be attained by charging a pooling premium. If entry costs are much higher than G^{high}, incumbent companies can increase their profits by explicit collusion and by using their information about risks. The lower bound, G^{low}, is needed to make sure that no incumbent company can gain by undercutting the premium for customers with small risk if the period payoff is equal to $\frac{1}{N}c$ for each incumbent company. If entry costs are lower, incumbent companies could still deter entry by charging low insurance premia, but they would have to use the information about risks in some periods, otherwise each incumbent company could gain by one-shot deviation. The measure of admissible values of c can be substantial: G^{low} strictly decreases in λ and will be small if π_H is close to π_L, while G^{high} can be large if customers are very risk averse and ready to pay a high risk premium.

The result of proposition 3 remains valid if incumbents use other punishment strategies to deter market entry or deviation from pooling premia. However, the maximal period profit for incumbent companies may decrease. Consider, for example, a tit-for-tat strategy where incumbent companies again start to charge profitable pooling premia after a finite number of periods with zero-profits. Then in an equilibrium with entry-deterrence, the period profit per incumbent company must be lower than $\frac{1}{N}c$. Otherwise, an outside company could enter the market, cover its entry costs by capturing the whole market (as it earns c), and participate in future business profitably after the price war has been finished.

4.4 Discussion

4.4.1 Examples of the evolution of information in insurance markets

Our model interprets the presence of unused observables as a sign of collusion. This is in accordance with experience in the US automotive insurance market where, as long as companies were making extensive profits, contracts were almost not differentiated by risk class (Carter (2005)). However, as profits in the market started to deteriorate in the late 1990s, one insurance company (Allstate), changed the number of pricing categories from 3 to over 1,500. As a consequence, Allstate's return on equity almost doubled in the following two years. However, as the author points out, this strategy might not be of lasting success, as other insurance companies also start to change their pricing system, and a price war in the auto insurance market is on its way.

It seems to be the case that it is often small firms or new entrants who start using a finer risk classification (Finkelstein and Poterba (2006)). Ainslie (2000) provides an example of the U.K. annuity market where new start-up companies were formed to offer impaired annuity products to those individuals in observable poor health. Only under increased competitive pressure did the existing companies follow suit.

The evolution of the use of information in the European Union has followed the evolution of competition in the insurance industry. Before 1994, when the European Commission completed a series of directives in order to remove obstacles to competition, the insurance markets in several European countries such as Germany and Italy were tightly regulated. Considering the use of information in automotive insurance in Germany, risk categories were rather coarse and involved extensive pooling (Rees and Kessner (1999)), while in Italy, companies were even restricted by law to a very limited number of parameters they could use in their pricing schemes (Buzzacchi and Valletti (2005)). After deregulation, as a consequence of increased competition, premiums in automotive insurance have undergone large reductions, and at the same time companies introduced contracts with finer risk categorization, see Rees and Kessner (1999) and Buzzacchi and Valletti (2005). However, in some markets, such as the annuity market in Germany, the companies remain in an equilibrium where contracts are almost not differentiated by risks classes at all.

4.4.2 Welfare and policy implications

The sole existence of unused observables is a signal of anti-competitive behavior in the insurance industry for the regulator. Therefore, the presence of unused observables could be used as a policy tool by competition authorities. Given that customers are aware of their individual risk, equilibria with unused observables are clearly inefficient: as long as customers are not forced to purchase full coverage (for example, by regulation), they will buy too little insurance.

Considering the debate on whether insurance companies should be allowed to gather genetic information or not,[60] there are cases where it might be welfare enhancing if not all information is used to set prices: if customers do not know their individual risk, genetic testing might impose ex-ante a classification risk on potential insurance buyers. However, our analysis shows that there are good reasons for insurance companies not to use genetic information in their pricing schedules: firstly, adjusting pricing schedules

[60] For an in-depth discussion of this issue consult Polborn *et al*. (2006), Strohmenger and Wambach (2000) and the papers cited there.

without coordination with other companies might trigger a price war. Secondly, using more information only makes sense for companies if profits rise. Companies might refrain from using all additional information in their pricing decisions for fear of market entry.

4.5 Conclusion

Recently, several empirical findings have contradicted the predictions of the standard one-shot model of an insurance market. Our model of an oligopolistic insurance market rationalizes the occurrence of two formerly unexplained phenomena, unused observables and bulk discounts.

From the model we derived two explanations why firms may not use all risk-relevant information: first, if firms collude tacitly and available information changes over time, then they possibly refrain from using new information in order to prevent a price war. Second, if there is the threat of market entry, then per-period profits are limited. Whenever firms attain the upper limit of per-period profits, it does not pay off for incumbents to include more information. Under both scenarios, firms can make positive profits out of contracts with high and low risks and therefore may offer bulk discounts to customers.

A number of extensions to our model can be made. We used a very simple model of an insurance market, where in each period a new cohort of customers arrives. Usually, a customer is insured over a longer time horizon and the insurance company can condition premia for this customer on her history of damages. Thus, experience rating could be introduced into the model. Eventually, firms may find it optimal to skip experience rating in order to simplify collusion. Furthermore, one can analyze more complex forms of collusion like, for example, collusion on several different insurance markets as in Bernheim and Whinston (1990). Note that many insurance companies offer various types of insurance. Finally, more empirical work on contracts and collusion in insurance markets would be desirable to investigate the use of information by insurance companies.

Chapter 5

What if everybody had a choice? Using hypothetical choice experiments to analyze the demand for prescription drug insurance

5.1 Introduction

"To say it as plainly as I can, health care reform is the single most important thing we can do for America's long-term fiscal health" (President Barack Obama, speech at the American Medical Association, June 15, 2009).

The USA is the country with the single most expensive health care system in the world (more than 16.2 percent of GDP in 2007) – health care costs per person are almost 50 percent higher than those in the next most costly nation.[61] Real health care costs per capita have been rising at an annual rate of 4.4 percent since the 1980s, and they are projected to rise even faster in the future. However, Americans are not healthier than citizens of other nations – among the twenty developed countries with the highest health care spending, the USA ranks lowest in life expectancy. In fact, the USA is in the group of countries performing particularly poorly in the elasticity of life expectancy of the 15 to 60 years-old with respect to health care spending - together with Botswana, Swaziland, South Africa, Namibia, Zambia, Uganda and Gabon. One of the main reasons for this is uneven access to health care that is caused by the great number of individuals without health insurance, an estimated 47 million U.S. citizens in 2007.

Thus, there is a crying need for health care reform. After Roosevelt, Truman, Nixon, Carter and Clinton, President Obama is the latest American president who has called for such a reform. With the reform, the following features of the present system are supposed to remain in place: First, obtaining health insurance coverage will remain an individual's free choice. Second, the supply of health insurance will remain in the hand of private

[61] This and the following facts about international comparisons of health care spending and outcomes are taken from McFadden *et al.* (2008).

companies. Third, the old system, with employers offering coverage for their employees, the Veterans Administration offering coverage for those (formerly) in the military, and public insurance for the elderly, disabled and poor (Medicare for the elderly and disabled, Medicaid for the poor and Shift for children) will remain in place.

Medicare provides basic health coverage for 43 million elderly and disabled U.S. residents. However, until 2006, prescription drug coverage was not provided by the program, and about a third of the relevant population were without such coverage (Neuman *et al.* (2007)). In 2006, Medicare Part D, a highly subsidized market for prescription drug coverage, was introduced. This introduction has been the largest single expansion in social insurance in the U.S. since 1965. Medicare Part D has implications for the optimal design of the U.S. health insurance systems, because it is similar in many dimensions to the envisaged health care reform.[62]

Lessons from this experiment will be crucial both for deciding whether to introduce universal health care in the USA and for the design of social insurance programs in other countries. Part D also gives important insights into the practicality of Consumer Directed Health Care (CDHC), an approach to achieving efficient allocation of health care resources by confronting consumers with the full marginal cost of the services they use. Further, Medicare Part D can act as a natural experiment of consumer behavior in real-world decision situations that are characterized by complexity, ambiguity and important consequences. Therefore, lively discussion of the consequences of Medicare Part D has taken place in numerous studies in economics and health.[63] Most of these studies have restricted their analysis to the relatively small group of "active deciders": those consumers who had no prescription drug insurance before the introduction of Medicare Part D.[64]

This essay contributes to the discussion on Medicare Part D by using a hypothetical choice experiment that was conducted using a random sample of the relevant population to analyze consumer demand. Thus, the whole potential market is included in the analysis instead of restricting the focus to a small group of consumers. In hypothetical choice experiments, individuals are asked to choose between different commodities whose attributes vary in order to infer the utility associated with these attributes and consumers'

[62] More details on Medicare Part D and its connection to the health care reform are given in section 5.2.

[63] The results of at least some of them are presented in section 5.2.

[64] Many other groups of consumers were also affected by the introduction of Medicare Part D, for example those who had some privately bought or employer sponsored coverage before or those on Medicaid. Many consumers were not given an equally free choice regarding their insurance coverage – as for example the consumers who were eligible for both, Medicaid and Medicare, and who were automatically enrolled and randomly assigned to prescription drug plans. We will call these consumers passive participants in contrast to the active deciders.

willingness-to-pay (WTP). Here, respondents are asked to choose between insurance contracts that differ in their level of coverage. The insurance premia for these contracts are randomly assigned to the respondents. Hypothetical choice experiments are a well-known tool from marketing and psychology that in recent years has been used increasingly for demand estimation by economists. They are also used frequently by health economists to elicit consumer preferences regarding health care.[65] We want to draw attention to the fact that they can be used to elicit the demand of consumer groups whose choices cannot be observed in the actual market.[66] We thereby proceed as follows:

In this chapter, we use the hypothetical choice experiments to elicit WTP for prescription drug insurance with different levels of coverage for all types of consumers. We distinguish between the active deciders and the passive participants (those receiving prescription drug coverage through their employers, the Veterans Administration, private insurance, Medicare Advantage Plans or Medicaid). We show that these groups of consumers differ in many dimensions from the active deciders (for example their WTP for insurance, their income, their risk types and risk attitudes). Therefore, drawing inference from just this group of consumers might lead to misleading conclusions - for example, when analyzing welfare effects of eliminating the coverage gap as suggested by President Obama in his speech to Congress on health care reform on September 10, 2009. As for the active deciders, we observe their actual decisions in addition to their hypothetical choices. We can then estimate a joint model.

For our analysis, we use the Retirement Perspectives Survey (RPS), a dataset unique in providing a random sample of elderly Americans (and thus covering all groups described above) and containing both respondents' actual decisions regarding Medicare Part D as well as a hypothetical choice dataset where respondents are asked to choose between different prescription drug plans and the option of not having coverage at all.

We find that WTP for drug insurance is low for consumers with either low expected drug costs or low income. On the other hand, consumers demand extensive coverage if they are currently in poor health, expect high future drug costs, but also if they are more risk averse. These findings conflict with consumers' real choices where neither health nor socio-economic indicators prove significant. A possible reason for this discrepancy is that the active deciders for whom we observe actual choices are too homogeneous. In order to analyze the impact of socio-economic conditions on insurance demand we need to consider

---

[65] For example Scanlon *et al.* (1997) or Becker and Zweifel (2008). For a review on the literature of applications of discrete choice experiments in health economics see Ryan and Gerard (2003) and Guttmann *et al.* (2009).

[66] More on hypothetical choice experiments and how they can be combined with survey on consumers' actual decisions can be found in section 5.3.

sufficiently heterogeneous consumer groups, which is possible through our hypothetical choice experiment. Taking the whole potential market into consideration, we find that active deciders exhibit a significantly lower WTP than passive participants. Therefore, welfare estimates of the introduction of Medicare Part D based on these consumers alone might underestimate actual welfare.

The chapter is organized as follows: Section 5.2 describes Medicare Part D, the institutional setting, the types of insurance plans offered, and consumer groups. Section 5.3 is on hypothetical choice experiments. Section 5.4 provides a description of our data and some descriptives. Section 5.5 describes our estimation method, and section 5.6 our econometric results. Section 5.7 concludes.

5.2 Medicare Part D

5.2.1 The market

Since its introduction in 1965, Medicare has been providing health insurance for elderly and disabled Americans. In 2008, enrollment was at about 45 million.[67] Individuals are eligible for Medicare if they are U.S. citizens or long-term legal residents of at least 65 years of age and if either they or their spouses have paid Medicare taxes for at least ten years.[68] Medicare is administered by the Centers for Medicare and Medicaid Services (CMS) within the U.S. Department of Health and Human Services.

Medicare Part A provides basic coverage for inpatient hospital stays. Medicare Part B is optional and provides additional care, for example for physician and nursing services and for durable medical equipment. In 1997, the government introduced the possibility to receive care through private health insurance plans which are known as Part C, Medicare+Choice or Medicare Advantage Plans. These plans contract with health care providers, and eligible patients can only receive care through those providers under contract.[69]

Before the introduction of Medicare Part D in 2006, only pharmaceutical treatments administered in a physician's office, in a hospital or other institution were covered by the

[67] There are several websites which provide information on Medicare: www.cms.hhs.gov, www.medicare.gov and www.statehealthfacts.org.

[68] Further, disabled U.S. citizens or those with end stage renal disease are eligible for the program. However, we concentrate on the elderly beneficiaries here who form the vast majority of over 90 percent of beneficiaries.

[69] Sometimes enrollees are given the possibility to use other providers at some extra cost.

program. This was a major drawback of Medicare, because only some Medicare beneficiaries had prescription drug coverage from some other source, while about 30 percent of Medicare beneficiaries had little or no prescription drug coverage (Neuman *et al.* (2007)). This had serious negative consequences: First, medical expenditures placed a major financial burden on the elderly.[70] Second, cost-related non-adherence, i.e. the discontinuation of medication because it is too expensive, was a big concern before the introduction of Part D (Madden *et al.* (2008)).

Consequently, the aims of Medicare Part D were to make drug insurance coverage affordable for the elderly with low incomes, to provide protection against catastrophic drug costs and to reduce cost-related non-adherence.[71]

Another motive was that insurance companies have more bargaining power than single consumers vis-à-vis pharmaceutical companies and this would stop the increase of prescription drug expenditures. Before the introduction of Medicare Part D, the share of prescription drug spending in U.S. health care spending had been rising steadily.[72]

In the designing of the market, the government has taken several measures to overcome problems related to adverse selection and excessive spending. First, the government subsidizes the newly established market so that enrollment should be optimal for almost everybody. Second, the market is organized in a way that is supposed to give companies incentives to pass the reductions they receive from pharmaceutical companies through to consumers.[73] Furthermore, the subsidy is supposed to equalize the portfolio of consumers with regard to their risk types and it provides reinsurance for catastrophic drug benefits.

[70] According to data from the Medical Expenditure Panel Survey, per-person expenditures among Medicare recipients for prescription drugs were equal to $1,789 in 2003, with more than half of this paid out-of-pocket and just about 8 percent paid for by the Medicare program (Duggan and Scott-Morton (2006)). In 2005, about 10 percent of Medicare beneficiaries had catastrophic drug bills of more than $5,600, while the median income in this population was only $15,700 (McFadden *et al.* (2008)).

[71] There exists a subsidy that recipients whose incomes were at or below 135 percent of the poverty line can apply for (the so-called LIS). Estimations for 2008 show that 12.5 million Medicare beneficiaries are eligible for LIS, with 9.4 million actually receiving it. Certain groups of Medicare recipients are automatically enrolled in the subsidy, for example those on Medicaid (The Henry J. Kaiser Family Foundation (2008)).

[72] The increase was first slow starting from 4.5 percent in 1982 and then rapidly accelerating in the 1990s to 10.1 percent by 2005 (Centers for Medicare and Medicaid Services (CMS) (2008)), and about 60 percent of all prescriptions in the U.S. are filed for beneficiaries of Medicare, Medicaid, and other government programs (Duggan and Scott-Morton (2010)).

[73] At the beginning of each period, insurance companies take part in a bidding procedure with their bids representing an estimation of the costs that the company incurs in order to provide basic prescription drug coverage to a representative Part D enrollee. The monthly premium that the plan can charge to its costumers depends on this bid (and therefore, competition for consumers reduces the incentive to increase the bid), while the government subsidy depends on the average national bids.

Third, the late enrollment penalty helps overcoming adverse selection by encouraging individuals to join Medicare Part D earlier in their lives and not only when they have already developed health problems. It has to be paid by everybody who i) was eligible for Medicare Part D during the initial enrollment period (November 15, 2005 to May 15, 2006) or who becomes eligible for the first time during any future open-enrollment period (November 15 - December 31 of each year); ii) does not have creditable coverage (coverage which is considered equivalent to Medicare Part D); and iii) who decides not to enroll in Medicare Part D. It is equal to the months without coverage times one percent of the national base beneficiary premium.

Fourth, the standard plan designed for this market has several features that are supposed to discourage excessive spending: A deductible, copayments and a coverage gap in which the insured has to bear 100 percent of prescription drugs. Further, most plans give incentives to buy cheap drugs such as generics by reducing copayments for these types of drugs.[74]

5.2.2 Market outcomes and prior research on the active deciders

Research on the active deciders has drawn the following lessons from Medicare Part D:

First, by and large, Medicare Part D has been a success in providing a large percentage of the Medicare eligible population with prescription drug coverage. Enrollment rates were above 90 percent in the first year of Medicare Part D (Winter *et al.* (2006)). Those who remained without coverage in 2006 belong to two very different consumer groups: Those in relatively good health and those potentially difficult to reach (Neuman *et al.* (2007)).

Second, the complexity of the market with its many providers and many different products may have resulted in suboptimal choices, especially among the most vulnerable - those with low income, low educational attainment, poor health or some cognitive impairment (Heiss *et al.* (2006) and Duggan *et al.* (2008)). Some research suggests that actual premia are the main driver of consumers' choices, and that too little weight is placed on expected out-of-pocket costs (Abaluck and Gruber (2011)).

Third, substantial adverse selection seems to have occurred, but no moral hazard among the active deciders.[75] Most consumers who remained without prescription drug coverage behaved economically optimal, as their prescription drug needs were low (Winter *et al.*

[74] See section 5.2.3 where the available plans are described in more detail.

[75] The term moral hazard might not be the correct one in the context of health insurance. Individuals might have used too little prescription drugs before, and this might have had adverse consequences for their health. In fact, this was one of the reasons for introducing prescription drug coverage in the first place.

(2006) and Levy and Weir (2009)). However, premia of those plans offering basic Part D coverage (see below) do not seem to have been affected strongly by adverse selection: They were even lower than expected. Only those plans appear to have been affected by adverse selection that offer more extensive coverage that does not benefit from government subsidies, a sign of which were rapidly rising premia and reduced coverage (McFadden *et al*. (2008)).

Several features of Medicare Part D described above make it similar to the envisaged health care reform.[76]

First, it remains an individual's choice whether to enroll in Medicare Part D, whether to remain without coverage, or whether to remain with the already existing coverage, for example from employers or private insurance.

Second, there are subsidies for low-income consumers. The government subsidy for catastrophic costs places a limit on the out-of-pocket expenses of insured individuals. There is no cap on the amount of coverage that someone can receive.

Third, insurance is offered by private companies, but there are government regulations that restrict what can be offered on the market: Contracts offered have to provide a certain minimum coverage as well as coverage for catastrophic events. Individuals cannot be denied coverage because they are high-risk consumers. Because insurance companies buy in great bulks, the government expected that they would be able to use their bargaining power vis-à-vis pharmaceutical companies or health care providers.

Fourth, there are several mechanisms that are supposed to counteract adverse selection and excessive use of care.

Although research on Medicare Part D has already drawn many important conclusions, when analyzing consumer welfare or predicting policy changes, for example the abolition of the coverage gap, it is crucial to make predictions that are valid for the whole population. Therefore, in the following, we describe the features of the supplied plans and the groups of consumers which will be important for the following analysis.

5.2.3 Types of plans

Under Medicare Part D, the plans insurers can offer are standardized. The standard drug benefit, as defined by the Medicare Prescription Drug Improvement and Modernization Act of 2003, is characterized by four main features:[77]

---

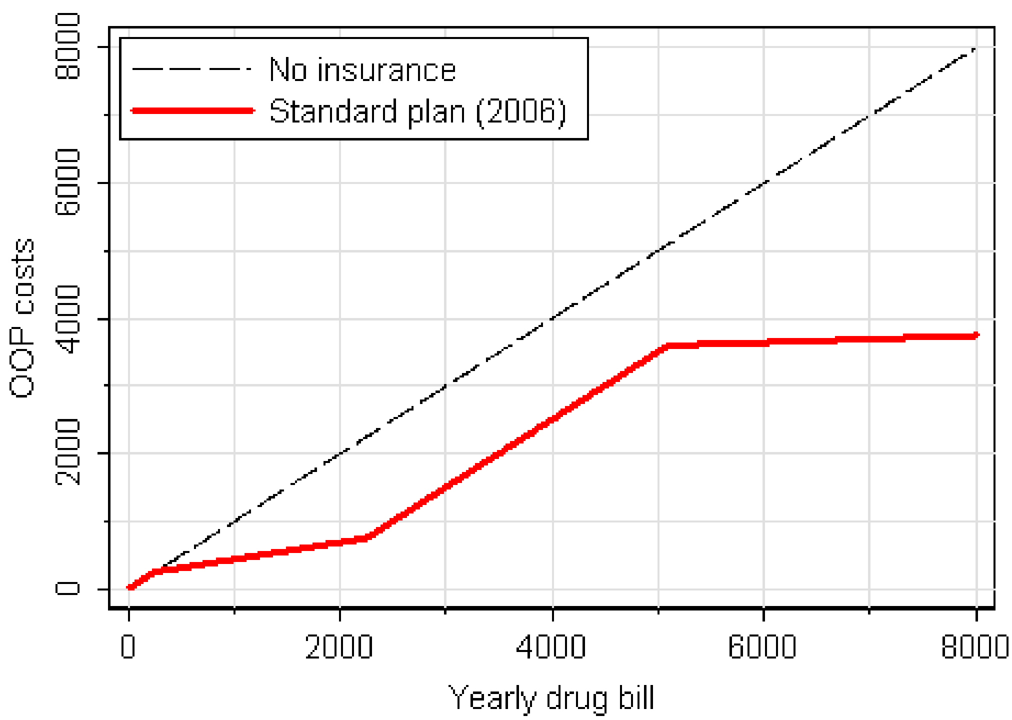
[76] Government plans for the reform can be found at www.healthreform.gov.

[77] Features of the plan have changed slightly over time. These are the features of the plan in 2006, when our first hypothetical choice experiment was conducted.

1. A \$250 (annual) deductible below which the insured have to pay for all costs themselves.
2. An interval of drug spending between \$250 and \$2,250 where the plan covers 75 percent of drug costs.
3. A coverage gap between \$2,250 and \$5,100 where the insured has to bear the full costs (which translates into an out-of-pocket limit of \$3,600).
4. A catastrophic threshold of \$5,100 above which the insurance covers 95 percent of all costs.

Figure 5.1 (from Heiss *et al.* (2009a)) shows the beneficiary's out-of-pocket costs as a function of his yearly drug bills. The plan was designed in order to both reduce moral hazard (through the deductible and the coverage cap) and to provide insurance against catastrophic drug costs.

Figure 5.1: Benefit schedule of the Medicare Part D standard plan



Companies can either offer the standard plans, or plans that offer more extensive coverage, either by having no deductible or by providing coverage in the coverage gap. The CMS classifies the stand-alone prescription drug plans that are available under Medicare Part D in four categories (Heiss *et al.* (2009a)):

- **The standard plan** is the plan with the statutorily defined coverage, deductible, gap, and cost sharing.
- **An actuarially equivalent plan** has the same deductible and gap as the standard plan, but has different cost sharing modalities (such as copayment tiers for preferred drug costs and generic drugs rather than percentage copayment). Actuarial equivalence to the standard plan may be achieved through restrictions in plan formularies, but all approved plans must have formularies that include at least two drugs in each therapeutic category. In 2009, about 34 percent of Part D stand-alone prescription drug plans had the standard deductible (The Henry J. Kaiser Family Foundation (2008)).
- **A basic alternative plan** is actuarially equivalent to the statutory defined benefit, but both the deductible and cost sharing can be altered (most of these plans have no deductible). In 2009, about 55 percent of stand-alone plans had no deductible at all, and 11 percent of plans had a deductible different from the standard one.
- **An enhanced alternative plan** exceeds the defined standard coverage - for example, by offering coverage in the gap for generic drugs only, or both generic and branded drugs. About one quarter of Part D stand-alone plans and one half of Medicare Advantage Plans offered gap coverage in 2008 and 2009. The share of Medicare stand-alone plans with no gap coverage decreased from 85 percent in 2006 to 75 percent in 2009. In Medicare Advantage prescription drug plans this percentage was reduced even more drastically from 72 percent in 2006 to 49 percent in 2009 (The Henry J. Kaiser Family Foundation (2008)).

5.2.4 Groups of consumers

Medicare recipients can belong to different groups regarding their prescription drug coverage: Stand-alone plans, Medicare Advantage Plans or private insurance, coverage through the employer or the Veterans Administration, or coverage through Medicaid. These groups will be described in the following.

Those individuals who had no prescription drug coverage before the introduction of Medicare Part D had to make an active choice to enroll in Medicare Part D - remaining inactive meant choosing to remain uncovered. They could either enroll in Medicare Part D stand-alone plans which cover prescription drugs only, or in Medicare Advantage Plans (called

Medicare + Choice before) where prescription drug coverage is provided as part of overall health care through HMOs.

Some of these plans had covered prescription drugs already before the introduction of Medicare Part D. With the introduction of Medicare Part D, Medicare Advantage Plans were almost forced to offer prescription drug coverage, because their enrollees could not take up Medicare Part D without losing their benefits from outpatient and inpatient care (Levy and Weir (2009)). Further, these plans are subsidized more heavily in order to encourage Medicare recipients to seek more extensive coverage (Duggan *et al.* (2008)).

Thus, Medicare Advantage Plan beneficiaries may belong to either of two groups: They may have had prescription drug coverage before, and this coverage was simply converted into Part D coverage, or they may have chosen prescription drug coverage only with the introduction of Part D. In 2006, about 10.4 million of Medicare recipients enrolled in Part D chose stand-alone coverage, while about 6 million received coverage through Medicare Advantage Plans, including 1.2 million new enrollees. About half a million of those enrolled in Medicare Advantage Plans are recipients of Medicaid (U.S. Departement of Health and Human Services (2006)).

The situation is similar for individuals that had private insurance for their prescription drugs before the introduction of Medicare Part D. A special situation holds for Medigap (or Medicare supplemental) health insurances. These are private supplemental health insurance plans that cover medical expenses that are not, or partially not, covered by Medicare. Since 2006, these plans cannot offer prescription drug coverage to new enrollees.

In order not to crowd out prescription drug coverage offered by employers, there are tax-free subsidies to those employers who provide prescription drug plans that are actuarily equivalent to Medicare Part D.[78] In January 2007, there were 6.9 million Medicare recipients whose coverage was subsidized in this way (Duggan and Scott-Morton (2010)). Alternatively, employers could decide to wrap around Medicare drug coverage. Individuals enrolled in these types of plans are counted under Medicare Advantage Plans.

Those employees who had prescription drug coverage before the introduction of Medicare Part D simply received a letter from their employer that informed them that their prescription drug coverage was creditable when Medicare Part D was introduced.

Veterans already had prescription drug coverage before the introduction of Medicare Part D. As this is considered creditable coverage, there is no need for veterans to sign up

[78] Companies receive 28 percent of covered charges between the deductible and an upper limit of $5,600 for each Medicare-eligible participant.

for Medicare Part D, however, they can do so if they want.[79] In 2006, about 2 million Medicare beneficiaries received their prescription drug coverage through the Veterans Administration, and about one million veterans were enrolled in a Part D plan (U.S. Departement of Health and Human Services (2006)).

Medicare recipients who are also eligible for Medicaid were automatically enrolled in a prescription drug plan with some option to switch plans afterwards. Medicaid recipients do not pay any premium if they choose a plan with average or below average costs. They further have no deductibles, no coverage gaps, and lower copays (Duggan and Scott-Morton (2010)). In 2006, about 6.1 million Medicare and Medicaid recipients were automatically enrolled in prescription drug plans. Additionally, about half a million Medicare and Medicaid recipients received prescription drug coverage through Medicare Advantage Plans (U.S. Departement of Health and Human Services (2006)).

5.3 Hypothetical choice experiments

5.3.1 Real versus hypothetical choice experiments

In this section, we describe two types of data, real and stated preference data, which can be used to analyze consumer preferences.

Since the seminal work of McFadden (1974), economists have used survey data on the micro-level to analyze consumer demand. Each product is defined as a bundle of characteristics, for example in McFadden's famous work on travel demand, each mode of travel is defined by the price and time of travel. Consumer preferences are defined over these characteristics, and the utility consumers obtain from buying a product depends both on the product's and their personal characteristics and tastes which can only be partially observed by the researcher.[80]

In these consumer surveys, respondents were asked about their actual purchasing decisions in real markets. The preferences researchers deduce from this information are called revealed preferences because people reveal their preferences by making their choices. The advantage of deducing preferences from consumers' actual choices is that consumers spend real money facing their actual budget constraints and therefore, the researcher may be confident about making predictions from this type of data.

---

[79] In fact, this may be beneficial for some low-income veterans who are eligible for the Medicare Part D low-income subsidy (Rupper *et al*. (2007)).

[80] Some more information on discrete choice models can be found in section 5.5. A detailed description on the concepts and developments of demand estimation is given in Ackerberg *et al*. (2007).

However, making predictions from observed choices has some important limitations.[81] First, the attributes of products often exhibit little variation in real markets. For example, in many markets, price varies very little over products as well as over time. In a statistical model the researcher might therefore wrongfully deduce that consumer decisions do not respond much to variations in price, when in reality prices do not vary much because price elasticity is very high. Second, when forecasting demand for new products, revealed preference data simply does not exist. The same holds true for certain public goods which are not traded in the market. Third, choices might be observed for a certain group of consumers only - as in the case of Medicare Part D. When drawing conclusions or making predictions, for example what happens if a certain insurance coverage is extended to a larger group of consumers, researchers may want to have variation not only in product, but also in consumer characteristics.

To overcome these problems, individuals have been presented with hypothetical choice tasks. In these tasks, respondents are asked to choose among alternatives from a choice set whose attributes have been defined by the researcher. In hypothetical choice experiments, WTP is inferred from individuals' choices. This allows the researcher to give estimations on the overall utility of the alternatives as well as their attributes.[82] For example, in our data respondents are asked to choose between different insurance contracts whose prices vary randomly. The preferences elicited from this type of experiments are called stated preferences in contrast to consumers' revealed preferences from their real choices. Hypothetical choice experiments have been used extensively in marketing and psychology. Louviere *et al*. (2000) provide a comprehensive guide to the design, implementation and interpretation of stated choice methods. In these experiments, sufficient variation can be created. The choice situation can also be presented to a random sample of the population. Additionally, the researcher can hold fixed everything in the choice situation that he wants to hold fixed, and concentrate only on the product characteristics that he is really interested in. This helps to overcome the well-known problem that prices are correlated with unobserved product quality, which leads to biases in the estimate of consumer price elasticity in real data.

Still, a concern about stated choice data is that consumers might behave differently in the experiment than they do in the real world. As in every experiment, researchers are concerned with the external validity of hypothetical choice experiments. The seminal work in this regard is the paper by Mitchell and Carson (1989). Since then, many studies

[81] Compare chapter 2 of Louviere *et al*. (2000) or chapter 7 of Train (2003).

[82] In contrast, contingent valuation questions ask the respondent directly about their WTP. Avoiding to ask respondents directly about their WTP reduces the risk of strategic manipulation of the answers by the respondents (Brau and Bruni (2008)).

have aimed at testing how WTP stated in hypothetical choice experiments differs from real WTP. Two studies from the health sector are Hanley *et al*. (2003) and Becker and Zweifel (2008). Carson *et al*. (1996) have performed a meta-analysis comparing WTP estimates from stated and revealed preference counterparts, and they find that the stated WTP is slightly lower than real WTP. The issues cited most frequently are that first, the sample drawn for experiments is not representative of the population. For example, a sample of undergraduate students might not be a representative sample or the population for eliciting WTP for public goods. However, the sample of respondents we use in our data is a random sample of the relevant population. Second, when asked for their WTP for public goods, respondents might strategically overstate their WTP or respond in a way that they think is "politically correct". We expect this to hold true more in the public good setting and not in the insurance setting that we will apply the hypothetical choice experiments to.

There are also benefits from combining real and stated preference data, a method that is sometimes also called data enrichment. The basic idea is to create variation in attributes through the hypothetical choice data and to base the hypothetical choice experiments in reality using respondents' real choices.[83]

Figure 5.2 shows what type of choices, real or hypothetical or both, we observe for different consumer groups. For all consumers, we observe their hypothetical choices. For the active deciders, whether or not they decided to take up insurance coverage in Part D, and for those consumers who decided to enroll in Part D, we can observe their actual plan choices. Figure 5.3 illustrates what share the different groups have in our dataset. This gives us different possibilities of using our hypothetical choice experiment:

First, we can use the hypothetical choice experiment to analyze WTP for prescription drug insurance taking into account the whole potential market.

Second, concentrating on the active deciders, we can test for external validity of our hypothetical choice experiment - this is, we can try to answer the question of whether consumers' WTP expressed in our hypothetical choice experiment differs from their actual WTP (Carlsson and Martinsson (2001)).[84]

---

[83] See Train (2003), chapter 7, and Louviere *et al*. (2000), chapter 8. Applications include Morikawa (1989), Ben Akiva and Morikawa (1990), Hensher and Bradley (1993), Adamowicz *et al*. (1997), Hensher *et al*. (1999), Brownstone *et al*. (2000), Bhat and Castelar (2002) and Walker and Ben-Akiva (2002).

[84] Some of the methodological factors concerning the combination of revealed and stated choice experiments are described in section 5.5.

Figure 5.2: Decision tree of consumer groups

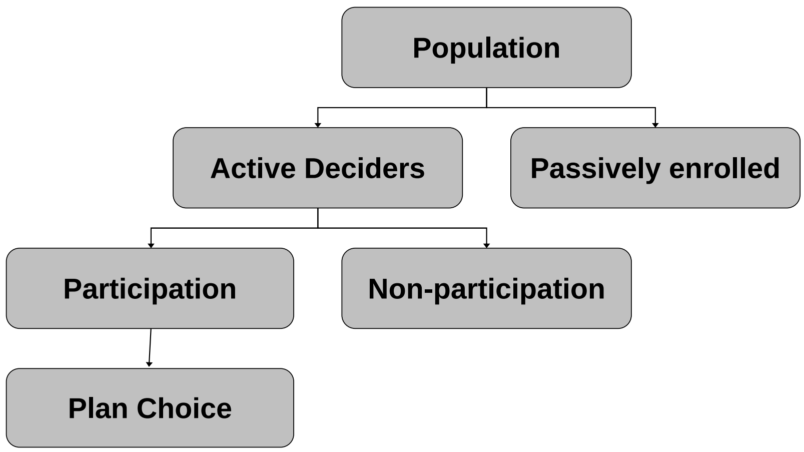
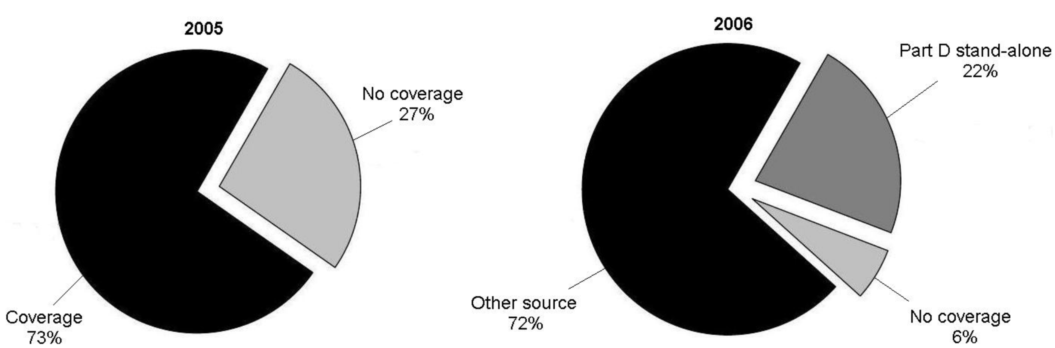


Figure 5.3: Consumer groups in the RPS



Third, we can estimate a joint model of both revealed and stated choices as for the active deciders we observe both types of choices. This can help us to mitigate the weaknesses and make use of the strengths of both types of data.

5.4 Data and descriptives

The Retirement Perspectives Survey (RPS) is a panel dataset of elderly U.S.-citizens that is representative of the U.S. non-institutionalized population in terms of demographics and socio-economic status. It is a research project conducted by Daniel McFadden, Joachim Winter and Florian Heiss, based on a panel of individuals maintained by Knowledge Networks, a commercial survey firm. Participants were provided with web TV hardware which enabled them to answer the internet survey with their TV sets.

There are four waves of the survey:

- RPS-2005, conducted in November 2005
- RPS-2006, conducted in May 2006
- RPS-2007, conducted in March and April 2007
- RPS-2009, conducted in February and March 2009.

The following information was obtained from each respondent:

- Socio-economic characteristics, including household income, age, gender, ethnic group, and education level[85]
- Measures of the respondent's health (such as self rated health, chronic conditions, functional limitations)
- Measures of the respondent's prescription drug use (number of prescription drugs taken, current and expected out-of-pocket costs for prescription drugs, names of prescription drugs taken for some frequent health conditions like chronic pain, high cholesterol or diabetes)
- Information on the respondent's prescription drug insurance coverage
- Information on the respondent's enrollment decisions, knowledge, and opinion regarding Medicare Part D
- Hypothetical choice experiments where respondents were given the choice between insurance contracts with differing levels of coverage (including no coverage) and randomly varying premia

[85] These variables are provided by Knowledge Networks as background variables.

- Some simple measures of risk attitudes and planning horizons.

There are several articles which use the RPS in order to analyze the enrollment decisions of active deciders (Winter *et al*. (2006), Heiss *et al*. (2009a), McFadden *et al*. (2008), Heiss *et al*. (2009b)). Heiss *et al*. (2009a) provide a detailed description of response behavior, selection issues and the application of sampling weights. Table 5.1 and Table 5.2 are taken from Heiss *et al*. (2009b). Table 5.1 shows how the sample of respondents and the response rates developed over time. For our analysis, we are only interested in those individuals eligible for Medicare, therefore we restrict the sample to respondents aged 65 and older in 2006. Table 5.2 shows how the RPS compares to the 2004 Health and Retirement Survey (HRS) in terms of socio-economic characteristics and insurance status. The RPS seems to reasonably mirror the HRS, even more in the weighted samples. However, we will not use weights in our subsequent analysis.

Table 5.1: Sample selection criteria and response rates, RPS 2005-2009

	RPS 2005	RPS 2006	RPS 2007				RPS 2009				
Age selection rule	50+	63+*	64+				64+				
			sub-samples				sub-samples				
Completed RPS 2005		yes**	yes	yes	no		yes	yes	no	no	
Completed RPS 2006			yes	no	no		yes	no	no	no	
Completed RPS 2007							yes	yes	yes	no	
KN members contacted	5'879	2'598	1'704	217	1'250	3'171	1'151	127	789	364	2'430
Completed interviews	4'738	2'137	1'526	165	1'020	2'711	783	77	534	207	1'601
Response rate***	80.6%	82.3%	89.6%	76.0%	81.6%	85.5%	68.0%	60.6%	67.7%	56.9%	65.9%

Notes:
* In addition, RPS 2005 respondents younger than 63 years were contacted for RPS 2006 if they said that they are on Medicare.
** Completion of RPS 2005 was required for this subsample.
*** The response rate is defined as the number of completed interviews as a proportion of the number of KN Panel members contacted.

Definitions and descriptive statistics of the variables used can be found in Table 5.3. Most variables correspond to a question of the survey with the exception of expected drug costs which was created by Winter *et al*. (2006) based on the respondents' prescription drug use and the price they would have paid for them over the counter. The idea is to elicit the expected drug bill for each individual in the case of no insurance.

In 2006, the RPS sample in the relevant age group (aged 65+) consisted of 1,666 respondents. 97 of them had to be excluded from the analysis because they did not answer the questions on prescription drug insurance coverage. 56 percent of our sample are females. Most of the respondents (about 38 percent) are between 65 and 70 years old, 28 percent between 71 and 75, and about 33 percent are older than 75 years. Annual household income was below \$20,000 for about 23 percent, between \$20,000 and \$60,000 for about 58 percent of the sample and above \$60,00 for the remaining 19 percent. About 13 percent

Table 5.2: Descriptive statistics, HRS 2006 and RPS 2006

		HRS 2006		RPS 2006	
		unweigthed	weigthed	unweigthed	weigthed
Gender	Female	57.3%	56.8%	55.8%	57.2%
	Male	42.7%	43.2%	44.2%	42.8%
Race	White	83.5%	89.3%	87.6%	83.3%
	Non-white	16.5%	10.7%	12.4%	16.7%
Age	61 – 70	35.8%	33.7%	39.3%	35.9%
	71 – 80	40.3%	41.6%	46.9%	47.9%
	81 – 90	20.4%	22.0%	12.9%	15.1%
	>90	3.5%	2.7%	0.9%	1.1%
Education	Less than HS	31.5%	28.3%	12.9%	26.1%
	High school	32.6%	33.4%	41.5%	36.5%
	More than HS	36.0%	38.4%	45.6%	37.5%
Income	<$20K	33.2%	31.2%	23.4%	28.9%
	$20K – $60K	46.2%	46.9%	58.2%	52.6%
	>$60K	20.6%	21.9%	18.4%	18.5%
SRHS	excellent	8.5%	9.1%	6.1%	5.6%
	very good	26.3%	27.5%	32.2%	27.8%
	good	31.6%	32.4%	39.5%	41.8%
	fair	23.3%	22.2%	18.1%	19.8%
	poor	10.3%	8.8%	4.0%	4.9%
Number of observations		11399		1666	

of respondents have less than high school education, while about 87 percent have high school education or higher qualification.

As to their health and prescription drug use, in 2006, when asked about the number of different prescription drugs used in the last month, some 10 percent did not use any prescription drugs at all. About 29 percent reported having used one or two different prescription drugs, and about 61 percent of respondents reported having taken more than three different prescription drugs within the last month. Most of the respondents (about 40 percent) rated their health as “good”, 38 percent as “very good” or “excellent”, and only 22 percent said that they were in “fair” or “poor” health. Our estimated prescription drug costs have a skewed distribution with a long tail, where a lot of consumers have quite low costs and very few consumers have very high costs.

Table 5.4 shows how prescription drug coverage in the RPS compared to the total population. Table 5.5 shows how health and demographics differed among the different consumer groups.[86] The most noticeable patterns regarding sociodemographic characteristics are that first, among veterans, the percentage of males is highest. Second, the groups with the lowest incomes are, as expected, those on Medicaid or state programs, while those with employer coverage or private insurance have rather high incomes. The latter groups also tend to have received more education. The fraction of non-whites is largest for those on Medicaid, for veterans or individuals who are covered by their employers.

Regarding health and prescription drug use, there are large differences among consumer groups: Total drug costs in 2005 were much lower for those consumers who were without prescription drug coverage before the introduction of Medicare Part D and who decided to remain without coverage (see also Figure 5.4). The median of estimated annual prescription drug costs for this consumer group is only $94, while it is about $1,567 for the group with the next lowest median, those who were covered by an Medicare HMO. The same holds true for the number of different prescription drugs taken during the last month - almost 40 percent of those remaining without coverage report not to have taken any prescription drugs. Other groups whose prescription drug use in 2005 was relatively low are those consumers with HMO or employer coverage, while consumers with coverage by either Medicaid or the Veterans Administration, but also consumers who bought private description drug coverage, on average took many prescription drugs and had high estimated costs.

Self reported health is likewise far better for those who remained without prescription drug coverage than for any other group (see Figure 5.5). Veterans, beneficiaries eligible for Medicaid or those who received coverage through state programs tend to be in poor health.

5.4.1 The RPS hypothetical choice experiment

In the RPS 2006, a hypothetical choice experiment was conducted in order to elicit the preferences for prescription drug coverage of all consumers, not just the active deciders.[87]

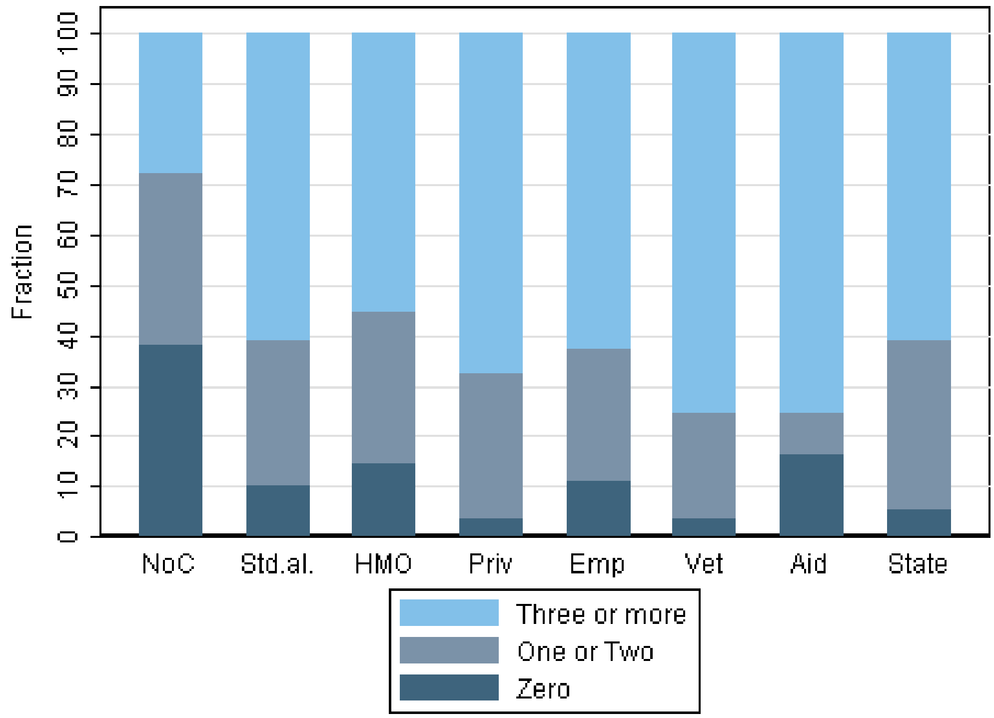
As the RPS focuses on questions on Medicare Part D and as all respondents in the RPS 2006 have already answered the questionnaire in 2005, we expect respondents to be familiar with the questions on insurance and Medicare Part D when taking part in the hypothetical choice experiment.

---

[86] Most differences between the other groups and the active deciders are statistically different at the ten percent level. Results of t-tests are not reported here.

[87] A description of the RPS can be found in section 5.4.

Figure 5.4: Number of different prescription drugs taken in 2005



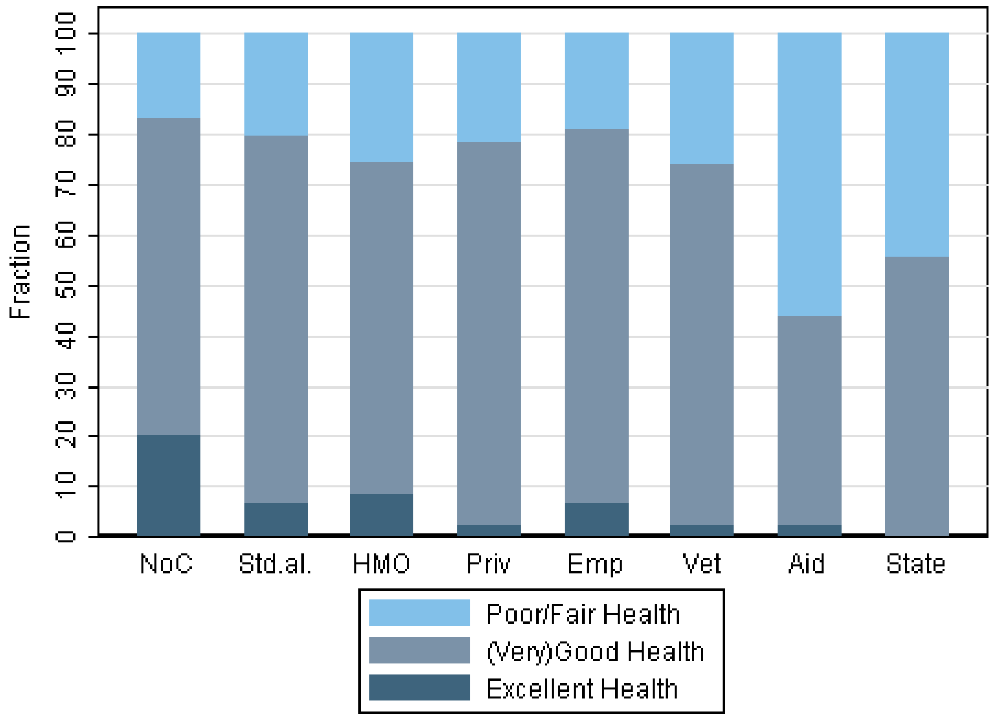
Consumers were provided with a short introduction, in order to place our hypothetical choice experiment in the context of Part D (telling them for example that the same late enrollment premia would apply in the experiment as under the Part D regulation). One part of the introduction for the experiment differed depending on which group of consumers was interviewed. The exact wording of the experiment can be found in the appendix of this chapter.

Then the respondents were given a choice between four alternatives with randomly varying premia:

- **Basic Plan premium:** $ **PBi** *This plan covers all prescription drugs you currently use and most of what you might need in the future. It has a deductible of $250, pays 75 percent of costs above $250 up to $2250, provides no additional benefit until costs reach $5100, and pays 95 percent of costs above that level.*

- **Enhanced Plan premium:** $ **PEi** *This plan is equivalent to the Basic Plan but has no deductible. This means that the 75 percent coverage begins at the first dollar you spend on prescription drugs, up to $2250. Like the Basic Plan, there are no*

Figure 5.5: Self-rated health status in 2005



additional benefits until costs reach \$5100. The Enhanced Plan pays 95 percent of costs above that level.

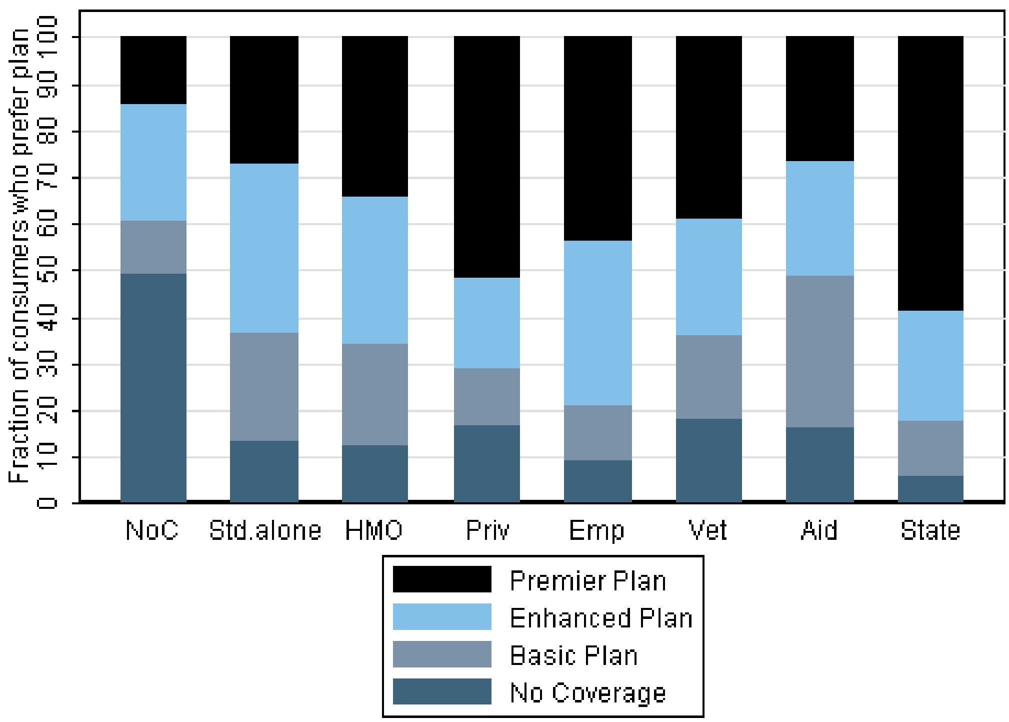
- **Premier plan premium:** \$ **PPi** *This plan is equivalent to the Enhanced Plan, but is does not impose a coverage gap between \$2250 and \$5100. So it pays 75 percent of all costs up to \$5100 and for 95 percent above that amount.*

- **No prescription drug insurance at all**

Each respondent was presented with three different hypothetical choice tasks. In the first round, everybody was presented with the same hypothetical prices, and in the second and third round, prices were randomly assigned to the respondents. The upper panel of Table 5.6 shows the premia that were assigned to the different types of plans in our hypothetical choice experiment, and therefore our hypothetical supply prices. For the first choice, premia were the same for all respondents. These premia closely resemble the premia for all plans available in the market as constructed from Heiss *et al*. (2009a) with the CMS plan finder. Note that the premia for the plans actually chosen by the active deciders in the RPS, and therefore the prices in market equilibrium, are much lower than supply prices, at least for basic and enhanced coverage. For the second and third choice, premia were randomly assigned.

We find large differences regarding the preferred plans when looking at consumer choices in the first hypothetical choice experiment where the prices were the same for everybody: Those consumers who chose to remain without coverage in the real market also show a much greater preference for less or no coverage in the hypothetical choices (see Figure 5.6). The hypothetical choices of those having chosen stand-alone coverage and those who chose HMO coverage are quite similar: A large percentage of consumers would prefer plans that offer more coverage than the Part D standard plan. This tendency is even stronger for those consumers who were covered by their former employers or the Veterans Administration and even more pronounced for those consumers who had private prescription drug insurance or who were covered by state programs. Surprisingly, consumers on Medicaid do not show a comparably low preference for plans with no deductible and gap coverage.

Figure 5.6: Hypothetical choices by consumer groups



5.5 Multinomial Logit Model

The multinomial logit model (MNL), sometimes also called conditional logit model, was developed by Daniel McFadden in the mid 1960s in the context of so-called random utility models (McFadden (1976)). It goes back to the work of Thurstone's "Law of comparative judgment" (1927), Marschak (1960)'s introduction of random utility models and Luce (1959)'s description of the Independence of Irrelevant Alternatives (IIA) axiom.

The idea behind the random utility framework (RUM) is as follows.[88] A decision maker i faces a choice among J mutually exclusive alternatives j. In our hypothetical choice experiment, these alternatives represent different prescription drug insurance contracts. We want to explain what factors drive the individual's choice. Since J is finite, the models describing the individual's choice are called discrete or qualitative choice models in contrast to continuous or quantitative models. The choice set is exhaustive, which means that all possible alternatives are included. In our case, the individual must also be given the alternative to remain without prescription drug insurance.

We assume that the decision maker will choose the alternative with the highest (indirect) utility U_{ij}.[89] Utility depends on characteristics

- that vary over alternatives only, for example the premium of the insurance contract if it is unconditional on the insured's characteristics
- that vary over the decision makers only, for example income or risk aversion
- that vary over both the decision maker and the alternatives, for example expected drug costs when a specific insurance contract is chosen.

Some of these characteristics, the x_{ij}, are observed by the researcher. $V_{ij} = V(x_{ij})$ represents the observed or representative part of utility. Other characteristics of either the alternatives or the decision maker are known by the decision maker himself, but unobservable to the researcher. The influence of all unobservable factors is combined in an error term e_{ij}. Therefore

$$U_{ij} = V_{ij} + e_{ij}, \tag{5.1}$$

or, stated differently, the unobservable is simply the difference between the true utility U_{ij} and the representative utility V_{ij}. Therefore, the distribution of e_{ij} will always depend on the specification that the researcher chooses for the observed part of utility.

[88] See Train (1986) and Train (2003).

[89] Utility that has been maximized under a budget constraint.

The particular choice model depends on the distribution that the e_{ij} are assumed to follow. The assumption for the logit model is that the e_{ij} are independently, identically distributed (i.i.d.) extreme value type I with density function

$$f(e_{ij}) = exp(-e_{ij}) * exp(-exp(-e_{ij})) \tag{5.2}$$

and cumulative distribution function

$$F(e_{ij}) = exp(-exp(e_{ij})). \tag{5.3}$$

Knowing that the decision maker will choose the alternative j that yields the highest utility and knowing the distribution of the unobserved error terms, the probability that decision maker i will choose alternative j can be expressed as:

$$\begin{aligned} P_{ij} &= Prob(U_{ij} > U_{ik} \ \forall k \neq j) && (5.4)\\ &= Prob(V_{ij} + e_{ij} > V_{ik} + e_{ik} \ \forall k \neq j) && (5.5)\\ &= Prob(e_{ik} - e_{ij} < V_{ij} - V_{ik} \ \forall k \neq j). && (5.6) \end{aligned}$$

This choice probability does not mean that each decision maker's choice is random - in fact, from the decision maker's point of view, his choice is deterministic: He will choose the alternative that yields the highest utility. The choice is random only for the researcher who does not observe all relevant characteristics. It is random in the sense that if the researcher observes a sample of decision makers with the same observed characteristics, he will still not be able to perfectly predict their choices because he does not observe some of the relevant factors of this decision. If the sample size goes to infinity, then the share of people choosing alternative j will converge to P_{ij}.[90]

Note that the individual's choice depends on the differences in the utilities of the different alternatives. Therefore, each decision maker's choice depends on the joint (cumulative) distribution of the differences in unobserved utility between all alternatives. He will only choose alternative j if the random variables $e_{ik} - e_{ij} \ \forall k \neq j$ are jointly below the known value $V_{ij} - V_{ik} \forall k \neq j$. In the logit model, the $e_{ik} - e_{ij}$ follow a logistic distribution with distribution function

$$F(e_{ik} - e_{ij}) = \frac{exp(e_{ik} - e_{ij})}{1 + exp(e_{ik} - e_{ij})} \ . \tag{5.7}$$

[90] There exist also other interpretations of this randomness, as for example bounded rationality or quixotic factors. A survey respondent in a stated choice experiment might for example simply make his cross at a random alternative (Train and Wilson (2009)).

The choice probabilities then become

$$P_{ij} = \frac{exp(V_{ij})}{\sum_k exp(V_{ik})} \tag{5.8}$$

where the observed part of utility, V_{ij}, is usually specified as a linear function of the observed characteristics and a vector of unknown parameters, β, that has to be estimated

$$V_{ij} = \beta' x_{ij}. \tag{5.9}$$

Note that as only differences in utility matter, only characteristics that vary over alternatives can affect the choice probabilities. The influence of characteristics that vary over the decision makers can only be identified by interacting them with variables that vary over the attributes.

Estimating discrete choice models requires some type of normalization because utility is a cardinal variable which means that the level of utility is not identified. Neither adding a constant to each alternative's utility, nor multiplying utility by a constant will change the observed choices. Therefore, some normalization is necessary when estimating the unknown parameters β. As $U_{ij} = V_{ij} + e_{ij}$, multiplying utility by a constant λ means multiplying the variance of the error term by λ^2. Normalization is therefore usually done by normalizing the variance of the error term. In the case of the logit model, the variance is normalized to $\pi^2/6$. This means that the estimated βs are estimates of the "real" βs, divided by λ which is defined by the (unknown) variance σ^2 of the unobserved factors:

$$\sigma^2 = \frac{\lambda^2 \, \pi^2}{6}. \tag{5.10}$$

In other words, the "true" βs cannot be identified separately from σ^2, and when we compare the coefficients from two data sources, we will never know whether differences result from differences in the true parameters or the variance of unobserved factors.

If the model contains alternative-specific constants c_j, these constants will represent the mean of all unobserved factors per alternative, and the normalized error

$$\epsilon_{ij} = \frac{e_{ij} - c_j}{\lambda} \tag{5.11}$$

is distributed i.i.d. extreme value with mean zero and variance $\pi^2/6$.

The parameters of the MNL model are usually estimated by maximum likelihood. The joint likelihood (over all individuals) of the observed choices is maximized with respect to the unknown parameters.

One important feature of the multinomial logit model is the IIA property (Luce (1959)). In the logit model, the choice between two alternatives j and k is *independent of irrelevant alternatives* in the sense that the ratio of the probability of choosing alternative j to the probability of choosing alternative k is independent of all other alternatives and their attributes. This property is caused by the relatively strict assumption that the errors are independent of each other, which means that the unobserved parts of utility of the alternatives are not related to each other. Expressed in terms of elasticities of substitution, when one desirable characteristic of one alternative changes and this leads to an increase in the probability that this alternative will be chosen, the result will be a proportionate decrease in the probabilities that all other alternatives are chosen. This might not be realistic if some alternatives are more similar than others. However, the residuals in the logit models depend on the researcher's specification of representative utility. Therefore, the IIA will hold if heterogeneity in the attributes and the decision makers is attributed to observed characteristics. In other words, the logit model can capture heterogeneity due to observed characteristics, while it cannot capture variation due to unobserved characteristics.

Several tests of the IIA property were suggested by McFadden (McFadden (1978) and McFadden (1987)), Hausman and McFadden (1984), and Train *et al*. (1989). The test we will use in this chapter is based on the test developed in Hausman and McFadden (1984). The idea is to estimate the model twice: Once with the full set of alternatives, and once with a subset of alternatives (where IIA is assumed to hold). The parameter estimates of the full sample are consistent and efficient under the null hypothesis that IIA holds, but inconsistent if it fails. The parameter estimates from the restricted sample are consistent, but inefficient under IIA, and consistent even if IIA fails. Therefore, a standard Hausman specification test can be used. If the variance-weighted difference of the two estimates is too large compared with the critical values of a χ^2 distribution, then the null hypothesis of IIA is rejected.

There are some methodological issues concerning stated and revealed preference models in the context of logit models (see Train (2003), chapter 7, and Louviere *et al*. (2000), chapter 8): While the coefficients in respondents' utility functions should be the same in both types of data, unobserved factors will differ in stated and revealed preference situations. We would expect that the real choices of respondents are affected by many more unobserved factors than stated choices. For the real choices, these might include unobserved attributes of both the alternatives or the decision maker and measurement error, while for the stated choices, there might be unobserved attributes of the decision maker, factors that are specific to the experimental design and variables that are relevant for the choice situation, but that have not been included in the experiment (Bhat and Castelar (2002)).

Therefore, when estimating a joint model of the stated and revealed preferences, we want to allow for different scale factors λ^{SP} and λ^{RP} (Morikawa (1989) and Louviere *et al*. (2000)), assuming that the true utility parameters are the same for the two data sets. As the scale factors are unobserved and can never be identified within one source of data, it is the convention to normalize λ^{RP} to unity such that λ^{SP} represents the stated preference *relative to* the revealed preference scale factor. One possibility to jointly estimate both the model parameters and the (relative) scale factor is using a nested logit model (Bradley and Daly (1992), Hensher and Bradley (1993) and Louviere *et al*. (2000)). The nested logit model generalizes the MNL by relaxing the IIA assumption. In particular, alternatives are grouped into subsets or nests, and the variances of the error terms (and therefore the scale factors) are allowed to vary across nests. Therefore, artificially creating two nests for each decision made, one that contains the alternatives from the stated and one that contains the alternatives from the revealed preferences, allows estimating both the βs and the relative scale parameters.

5.6 Econometric results

Let the indirect utility that consumer i obtains from the insurance contract j be

$$U_{ij} = V(a_j, p_j, y_i, s_i, g_i) + e_{ij} \tag{5.12}$$

where a_j are the attributes of the insurance contract of which we explicitly state p_j, the contract's premium; y_i is the income of the insured; s_i are socio-economic conditions; and g_i is the group of the population that the consumer belongs to, which might either just reflect a difference in the status quo of the insurance[91] or be some proxy for other unobserved characteristics of the respondent.

The attributes of the insurance contracts are (compare also Table 5.3):

- **Insurance**: The contract provides prescription drug coverage with copayments, the deductible and the coverage gap equal to the Part D standard benefit.
- **No deductible**: The contract does not have the \$250 deductible of the Part D benefit.
- **Gap coverage**: The contract provides coverage also in the coverage gap.
- **Premium**: Monthly premium in \$.

We will proceed with our analysis as follows: We will first analyze the hypothetical choices from a random sample of the whole potential market. Next, we will analyze the real choices

[91] Compare Becker and Zweifel (2008).

of the active deciders. Finally, we will estimate a joint model combining both real and hypothetical choices.

Table 5.7 shows the MNL estimates of consumers' hypothetical choices. Each consumer is presented with three choices, and thus we use standard errors clustered by the individual. The standard Hausman/ McFadden IIA test (see section 5.5) cannot be performed in this case, because the assumption that the MNL estimator is efficient under the null hypothesis is violated with clustered standard errors. We instead implement Stata's suest version of the Hausman test (StataCorp (2007)). We assume that it is most likely that the IIA assumption will be violated for the no coverage alternative. The p-value for the rejection of the IIA assumption is at 8.3 percent, so we cannot reject IIA at either the 1 or 5 percent level, but at the 10 percent level, we can. Thus, for a first analysis, we stick with the MNL model and the IIA assumption.

Our main interest lies in estimating consumers' WTP for drug insurance with different levels of coverage where WTP is defined as the amount of premium increase that exactly offsets the increase of an attribute by one unit (or in the discrete case, the amount of premium that exactly offsets being provided with the discrete attribute versus not being provided with it), so that total utility remains unaffected:

$$WTP = -(\beta_{attribute}/\beta_{premium}). \tag{5.13}$$

The ratio of two coefficients can be interpreted directly in the MNL model as the unobserved scale factors cancel out in this case. For illustration, Table 5.8 contains the estimates of consumers WTP for selected equations. Our basic regression (1) shows that the WTP estimated from the hypothetical choice experiments is \$41.80 for basic, \$66.57 for enhanced and \$94.32 for premium coverage.

Regressions (2) and (3) show that WTP for basic insurance coverage is low for consumers with either low expected drug costs or low income. By contrast, consumers demand extensive coverage if they are currently in poor health, expect high future drug costs, but also if they are more risk averse. These findings conflict with consumers' real choices where neither health nor socio-economic indicators prove significant. A possible reason for the discrepancy of regression results on hypothetical and real data is that the active deciders are too homogenous, and that we need to consider sufficiently heterogeneous consumer groups in order to analyze patterns in these characteristics.

Hypothetical WTP in our base regression is somewhat higher than WTP for all observable plans and a lot higher than WTP for the plans actually demanded by RPS respondents (compare Table 5.6). In part, this is due to the fact that we observe real choices for the active deciders only, and the hypothetical choices confirm that WTP is significantly lower for the active deciders than for the passive participants (compare regression (4)). WTP of the active deciders is at \$30.30 for basic, \$48.91 for enhanced and \$63.33 for premium coverage in the hypothetical choices which is much closer to actual market behavior. Therefore, welfare estimates on the introduction of Medicare Part D based on the active deciders only will significantly underestimate welfare for the whole potential market.

Table 5.9 shows the MNL estimates of the real market decisions of the active deciders. WTP estimated from the decisions of the active deciders is insignificant for prescription drug insurance per se (compare Table 5.8). Consumers are willing to pay \$5,19 for coverage without deductible and an additional \$8.30 for gap coverage. These estimates seem unreasonably low. This might be due to three reasons: First, 101 of our 469 active deciders are consumers who decide to remain without prescription drug coverage. For this group, WTP should be indeed below the lowest observed supply price. Second, insignificant WTP for certain product attributes might be due to the fact that there is a high correlation between attributes and prices. Third, for consumers who expect not to need any prescription drugs and therefore have zero expected costs, WTP for insurance, whether basic or more extensive, is either negative or insignificant. Only consumers with high drug costs exhibit significant and positive WTP. Fourth, we observe more product attributes in the real market than in the hypothetical market. For example, gap coverage is either provided for generics only or for both generics and brand-name drugs (see regression (3)). In fact, WTP for gap coverage for generics only is not significant, while it is \$25.78 if both generics and brand-name drugs are covered. Further, consumers' WTP is lower for plans with drug tiers and higher for plans with a mail-order option, and WTP decreases for each top 100 drug that is either uncovered or only covered after authorization (See regression (4)). We will mitigate the outlined shortcomings of a separate analysis of hypothetical and real choices by estimating a joint model using both types of data.

As described in section 5.5, combining stated and revealed preference models allows us first, to make use of the whole potential market, second, to create variation in attributes through the hypothetical choice experiments and third, to base respondents' decisions in reality by using the real choices of the active deciders. As described in section 5.5, we can use a nested logit model in order to allow for different scale factors in the stated and revealed preference model (see Table 5.10). We restrict the scale parameter of the real choices to be one. Stata does not report the scale parameter itself, but the dissimilarity parameter $1/\lambda$. First, note that the estimate of the hypothetical relative to the real dissimililarity parameter is equal to 0.72. As we can reject the hypothesis that it is equal to unity (unity lies outside of the confidence interval), we can reject that the scale parameters and therefore the unobserved variances in our real and hypothetical choices are equal.

$$\frac{1/\lambda_{SP}}{1/\lambda_{RP}} = \lambda_{RP}/\lambda_{SP} = 0.72, \tag{5.14}$$

and

$$\frac{\sigma^2_{SP}}{\sigma^2_{RP}} = \left[\frac{1/\lambda_{SP}}{1/\lambda_{RP}}\right]^2 = 0.72^2 \approx 0.52. \tag{5.15}$$

Therefore, the variance of the hypothetical data is about 52 percent of the variance of the real preference data.

Estimated WTP from the combined choices is at \$35.39 for the Part D standard plan, \$51.84 for enhanced and \$72.39 for premium coverage. Note that these estimates are on the one hand much more sensible than the unreasonably low estimates from consumers' real choices. This might be due to the variation in the premia which we created in our experiment. On the other hand, they are somewhat lower than the estimates from the hypothetical choices. Therefore, combining real and stated preference data can help researchers to make more reasonable predictions on how consumer welfare has been affected by a policy change. Further, we can include consumer groups in our analysis whose actual choices cannot be observed, but who have still been affected by the reform. This is especially important when we want to take into account the effect of socio-economic characteristics on demand, because taking the whole market into consideration, we observe much more variation in these characteristics. Both make our results more generalizable in order to predict the effect of policy changes in other markets.

5.7 Conclusion

This essay has contributed to the discussion about Medicare Part D by using a hypothetical choice experiment that was conducted using a random sample of the relevant population to analyze consumer demand. Thus, the whole potential market has been included in the analysis instead of restricting the focus to a small group of consumers whose actual choices we can observe, the so-called active deciders.

We have found that willingness-to-pay for drug insurance is low for consumers with either low expected drug costs or low income. By contrast, consumers demand extensive coverage if they are currently in poor health, expect high future drug costs, but also if they are risk averse. With the exception of drug costs, none of these variables prove significant when using the actual decisions of the active deciders only, because this group is too homogenous in their characteristics.

Further, demand for prescription drugs differs significantly between the active deciders and other consumer groups who have also been affected by the reform. WTP of the passive participants is significantly higher than that of the active deciders. Therefore, welfare estimates of the introduction of Medicare Part D taking into account the active deciders only might be too low.

As we observe actual choices for one group of consumers, we have estimated a joint model using both real and hypothetical choices, thereby making use of the strengths and mitigating the weaknesses of both types of data. The estimates of WTP of the joint model seem to be most realistic regarding the prices we actually observe in the market.

So far, hypothetical choice experiments have mainly been used to create variation in product attributes. We want to draw attention to the fact that they can be used to elicit the demand of consumer groups whose choices cannot be observed in the actual market. This becomes important when making predictions about the impacts of policy changes, especially when we want to analyze how demand is affected by consumers' socio-economic conditions.

As lessons from the introduction of Medicare Part D will be crucial both for deciding whether to introduce universal health care in the USA and for the design of social insurance programs in other countries, analyzing how consumers behave in this market is highly policy relevant. There has been substantial debate on the question whether the U.S. government should have engaged itself further in individuals' health care decision, both in a regulatory and a financial way. Therefore, researchers have been interested how consumers' welfare has been affected by the introduction of Medicare Part D. In particular, it is important to find out whether the most vulnerable groups of consumers have been reached by the reform. Our findings can help to make the findings from the literature more generalizable to other populations and markets, because we do not restrict our focus to a small group of consumers.

5.8 Appendix

5.8.1 The RPS hypothetical choice experiments

In 2006, the introduction for individuals without prescription drug coverage was:

> *"At the end of this year, you will be able to make new choices about your prescription drug coverage for the next year."*

The introduction for those with Part D plans (whether stand-alone or HMO/ Medicare+Choice) was:

> *"At the end of this year, you will be able to make new choices about your prescription drug coverage for the next year. You may stay in your current plan, you may switch to another plan, or you may even elect to unsubscribe and not select any plan."*

The introduction for those covered through their employer or union, the Veterans Administration, private insurance or some other source was:

> *"Even though you have prescription drug insurance from other sources, we would like to know what your choices would be if the only coverage you could get would be through a Part D plan."*

The second part of the introduction was then the same for everybody:

> *"We are now going to show you some plans that have realistic features and premiums. We are interested in what plan you would choose if these were your only options.*
>
> *Specifically, on each of the following pages we will show you three plans that differ in coverage and premiums.*
>
> *On each page, please report which of these options is the most attractive and which is the least attractive. You will always have the option to choose none of these three plans and thus have no prescription drug coverage (but then you would have to pay higher premiums if you enroll later, according to current Medicare Part D regulations."*

The actual choice task was as follows:

> *"Please consider a situation in which you would have no prescription drug coverage from any other source. Imagine that these were the only three prescription drug plans that you could choose from. You can also choose not to have coverage at all."*

- **Basic Plan premium:** $ **PBi** *This plan covers all prescription drugs you currently use and most of what you might need in the future. It has a deductible of $250, pays 75 percent of costs above $250 up to $2250, provides no additional benefit until costs reach $5100, and pays 95 percent of costs above that level.*

- **Enhanced Plan premium:** \$ **PEi** *This plan is equivalent to the Basic Plan but has no deductible. This means that the 75 percent coverage begins at the first dollar you spend on prescription drugs, up to \$2250. Like the Basic Plan, there are no additional benefits until costs reach \$5100. The Enhanced Plan pays 95 percent of costs above that level.*

- **Premier plan premium:** \$ **PPi** *This plan is equivalent to the Enhanced Plan, but is does not impose a coverage gap between \$2250 and \$5100. So it pays 75 percent of all costs up to \$5100 and for 95 percent above that amount.*

- **No prescription drug insurance at all.**

Table 5.3: Variable description and descriptive statistics

	No coverage	Part D stand-alone	Part D HMO	Private	Employer	Veterans
Observations	94	352	226	83	573	139
share	6.06%	22.68%	14.56%	5.35%	36.92%	8.96%
Total drug costs in $ (2005)						
mean	1411.29	2766.98	2517.73	3268.08	2390.29	2908.36
1st quartile	0.00	834.99	685.39	1110.94	690.91	1204.36
median	93.79	1989.67	1567.16	2246.36	1878.80	2424.45
3rd quartile	1492.48	3335.78	3272.15	4056.73	3373.03	3940.83
Drug costs (2005)						
costs=0	39.36%	9.94%	14.60%	3.61%	11.34%	3.60%
0<costs<=250	15.96%	6.53%	7.08%	8.43%	7.16%	6.47%
250<costs<=1000	9.57%	8.81%	11.50%	10.84%	9.77%	10.07%
1000<costs<=2250	20.21%	32.39%	30.53%	27.71%	29.49%	23.02%
2250<costs<=5100	8.51%	29.26%	24.34%	33.73%	31.24%	43.17%
costs>5100	6.38%	13.07%	11.95%	15.66%	10.99%	13.67%
Number of prescriptions (2005)						
0	38.30%	9.94%	14.60%	3.61%	11.17%	3.60%
1 to 2	34.04%	29.26%	30.09%	28.92%	26.18%	20.86%
3+	27.66%	60.80%	55.31%	67.47%	62.65%	75.54%
Number of prescriptions (2006)						
0	31.51%	7.05%	10.84%	5.33%	7.97%	7.09%
1 to 2	36.99%	31.54%	29.06%	32.00%	29.48%	26.77%
3+	31.51%	61.41%	60.10%	62.67%	62.55%	66.14%
Self reported health (2005)						
excellent	20.21%	6.55%	8.41%	2.41%	6.46%	2.16%
(very) good	62.77%	72.93%	65.93%	75.90%	74.69%	71.94%
poor/fair	17.02%	20.51%	25.66%	21.69%	18.85%	25.90%
Self reported health (2006)						
excellent	14.89%	5.68%	6.19%	3.61%	6.11%	4.32%
(very) good	69.15%	73.86%	69.91%	73.49%	73.12%	71.22%
poor/fair	15.96%	20.45%	23.89%	22.89%	20.77%	24.46%
Age group (2006)						
<=70	32.98%	42.90%	34.51%	22.89%	42.06%	30.94%
70<age<=75	42.55%	24.72%	34.07%	27.71%	25.65%	31.65%
>75	24.47%	32.39%	31.42%	49.40%	32.29%	37.41%
Gender						
male	35.11%	38.64%	32.30%	45.78%	42.58%	92.09%
female	64.89%	61.36%	67.70%	54.22%	57.42%	7.91%
Educational level						
more than high school	38.30%	42.33%	38.94%	53.01%	53.23%	43.17%
high school or less	61.70%	57.67%	61.06%	46.99%	46.77%	56.83%
Income						
<=20K	30.85%	28.41%	36.73%	21.69%	12.04%	30.22%
20K-60K	58.51%	55.97%	48.23%	59.04%	63.00%	56.83%
>60K	10.64%	15.63%	15.04%	19.28%	24.96%	12.95%
Risk						
pay less than expected value	59.34%	52.03%	54.17%	55.42%	51.78%	52.94%
pay expected value	19.78%	22.38%	14.81%	22.89%	19.57%	19.12%
pay more than expected value	20.88%	25.58%	31.02%	21.69%	28.65%	27.94%
Hypotheticals: preferred plan						
no coverage	49.45%	13.45%	12.27%	16.46%	9.11%	17.91%
basic	10.99%	23.10%	21.82%	12.66%	11.79%	17.91%
enhanced	25.27%	36.26%	31.82%	18.99%	35.54%	25.37%
premium	14.29%	27.19%	34.09%	51.90%	43.57%	38.81%

Table 5.4: Prescription drug coverage among Medicare beneficiaries in 2006

	RPS	Total population	
Medicare advantage	20%	14%	(6.0m)
Part D stand-alone	22%	24%	(10.4m)
No coverage	6%	10%	(4.4m)
Employer	37%	24%	(10.4m)
Veterans	9%	5%	(2.0m)
Medicaid	3%	14%	(6.1m)
Other source	3%	8%	(3.4m)

Total market data taken from Department of Health & Human Services in Kaiser Family Foundation (2006).

Table 5.5: Descriptive statistics by consumer group

	No coverage	Part D stand-alone	Part D HMO	Private	Employer	Veterans
Observations	94	352	226	83	573	139
share	6.06%	22.68%	14.56%	5.35%	36.92%	8.96%
Total drug costs in $ (2005)						
mean	1411.29	2766.98	2517.73	3268.08	2390.29	2908.36
1st quartile	0.00	834.99	685.39	1110.94	690.91	1204.36
median	93.79	1989.67	1567.16	2246.36	1878.80	2424.45
3rd quartile	1492.48	3335.78	3272.15	4056.73	3373.03	3940.83
Drug costs (2005)						
costs=0	39.36%	9.94%	14.60%	3.61%	11.34%	3.60%
0<costs<=250	15.96%	6.53%	7.08%	8.43%	7.16%	6.47%
250<costs<=1000	9.57%	8.81%	11.50%	10.84%	9.77%	10.07%
1000<costs<=2250	20.21%	32.39%	30.53%	27.71%	29.49%	23.02%
2250<costs<=5100	8.51%	29.26%	24.34%	33.73%	31.24%	43.17%
costs>5100	6.38%	13.07%	11.95%	15.66%	10.99%	13.67%
Number of prescriptions (2005)						
0	38.30%	9.94%	14.60%	3.61%	11.17%	3.60%
1 to 2	34.04%	29.26%	30.09%	28.92%	26.18%	20.86%
3+	27.66%	60.80%	55.31%	67.47%	62.65%	75.54%
Number of prescriptions (2006)						
0	31.51%	7.05%	10.84%	5.33%	7.97%	7.09%
1 to 2	36.99%	31.54%	29.06%	32.00%	29.48%	26.77%
3+	31.51%	61.41%	60.10%	62.67%	62.55%	66.14%
Self reported health (2005)						
excellent	20.21%	6.55%	8.41%	2.41%	6.46%	2.16%
(very) good	62.77%	72.93%	65.93%	75.90%	74.69%	71.94%
poor/fair	17.02%	20.51%	25.66%	21.69%	18.85%	25.90%
Self reported health (2006)						
excellent	14.89%	5.68%	6.19%	3.61%	6.11%	4.32%
(very) good	69.15%	73.86%	69.91%	73.49%	73.12%	71.22%
poor/fair	15.96%	20.45%	23.89%	22.89%	20.77%	24.46%
Age group (2006)						
<=70	32.98%	42.90%	34.51%	22.89%	42.06%	30.94%
70<age<=75	42.55%	24.72%	34.07%	27.71%	25.65%	31.65%
>75	24.47%	32.39%	31.42%	49.40%	32.29%	37.41%
Gender						
male	35.11%	38.64%	32.30%	45.78%	42.58%	92.09%
female	64.89%	61.36%	67.70%	54.22%	57.42%	7.91%
Educational level						
more than high school	38.30%	42.33%	38.94%	53.01%	53.23%	43.17%
high school or less	61.70%	57.67%	61.06%	46.99%	46.77%	56.83%
Income						
<=20K	30.85%	28.41%	36.73%	21.69%	12.04%	30.22%
20K-60K	58.51%	55.97%	48.23%	59.04%	63.00%	56.83%
>60K	10.64%	15.63%	15.04%	19.28%	24.96%	12.95%
Risk						
pay less than expected value	59.34%	52.03%	54.17%	55.42%	51.78%	52.94%
pay expected value	19.78%	22.38%	14.81%	22.89%	19.57%	19.12%
pay more than expected value	20.88%	25.58%	31.02%	21.69%	28.65%	27.94%
Hypotheticals: preferred plan						
no coverage	49.45%	13.45%	12.27%	16.46%	9.11%	17.91%
basic	10.99%	23.10%	21.82%	12.66%	11.79%	17.91%
enhanced	25.27%	36.26%	31.82%	18.99%	35.54%	25.37%
premium	14.29%	27.19%	34.09%	51.90%	43.57%	38.81%

	Medicaid	State program	Part D unspecified	Other source	Total
Observations	49	18	14	21	1569.00
share	3.16%	1.16%	0.90%	1.35%	1.00
Total drug costs in $ (2005)					
mean	3383.00	3526.57	1340.41	1851.85	2554.30
1st quartile	1110.94	685.39	122.77	93.79	685.39
median	2726.72	1388.86	898.17	1602.46	1878.80
3rd quartile	5109.45	4797.22	1574.03	2674.51	3338.37
Drug costs (2005)					
costs=0	16.33%	5.56%	21.43%	23.81%	12.43%
0<costs<=250	2.04%	11.11%	7.14%	14.29%	7.52%
250<costs<=1000	2.04%	11.11%	21.43%	4.76%	9.69%
1000<costs<=2250	16.33%	27.78%	42.86%	23.81%	28.68%
2250<costs<=5100	36.73%	27.78%	0.00%	23.81%	29.38%
costs>5100	26.53%	16.67%	7.14%	9.52%	12.30%
Number of prescriptions (2005)					
0	16.33%	5.56%	21.43%	23.81%	12.30%
1 to 2	8.16%	33.33%	35.71%	28.57%	27.21%
3+	75.51%	61.11%	42.86%	47.62%	60.48%
Number of prescriptions (2006)					
0	15.22%	0.00%	11.11%	26.32%	9.65%
1 to 2	13.04%	12.50%	44.44%	21.05%	29.39%
3+	71.74%	87.50%	44.44%	52.63%	60.96%
Self reported health (2005)					
excellent	2.08%	0.00%	7.14%	9.52%	6.83%
(very) good	41.67%	55.56%	85.71%	80.95%	71.09%
poor/fair	56.25%	44.44%	7.14%	9.52%	22.08%
Self reported health (2006)					
excellent	2.04%	5.56%	0.00%	9.52%	6.12%
(very) good	55.10%	50.00%	85.71%	80.95%	71.83%
poor/fair	42.86%	44.44%	14.29%	9.52%	22.05%
Age group (2006)					
<=70	48.98%	22.22%	35.71%	28.57%	38.37%
70<age<=75	30.61%	33.33%	21.43%	14.29%	28.36%
>75	20.41%	44.44%	42.86%	57.14%	33.27%
Gender					
male	42.86%	27.78%	28.57%	42.86%	44.04%
female	57.14%	72.22%	71.43%	57.14%	55.96%
Educational level					
more than high school	36.73%	33.33%	35.71%	28.57%	45.70%
high school or less	53.27%	66.67%	64.29%	71.43%	54.30%
Income					
<=20K	61.22%	61.11%	28.57%	33.33%	25.05%
20K-60K	32.65%	38.89%	57.14%	47.62%	56.79%
>60K	6.12%	0.00%	14.29%	19.05%	18.16%
Risk					
pay less than expected value	58.33%	58.82%	57.14%	42.86%	53.13%
pay expected value	22.92%	17.65%	14.29%	23.81%	19.78%
pay more than expected value	18.75%	23.53%	28.57%	33.33%	27.09%
Hypotheticals: preferred plan					
no coverage	16.33%	5.88%	14.29%	26.32%	14.56%
basic	32.65%	11.76%	42.86%	26.32%	17.44%
enhanced	24.49%	23.53%	28.57%	36.84%	32.26%
premium	26.53%	58.82%	14.29%	10.53%	35.74%

Table 5.6: Market shares and premia of prescription drug plans in hypothetical and real choices in 2006

Monthly premia in $					
Hypothetical market		**Basic**	**Enhanced**	**Premium**	
First choice	Fixed premium	30.79	37.88	50.33	
Second and third choice	Lowest premium	15.39	18.94	25.16	
	Highest premium	40.02	49.25	65.43	
Real market		Basic	Enhanced	Generics	Generics and Brand
All available plans (Average premia)		30.75	37.92	48.13	61.88
Plans actually chosen in the RPS (Average)		17.00	26.60	46.10	60.80
Market shares in percent					
Hypothetical Market		**Basic**	**Enhanced**	**Premium**	
	All respondents	20.0	31.9	48.2	
	Only those with Part D	22.4	31.6	46.0	
First choice	All respondents	20.4	37.8	41.8	
	Only those with Part D	26.7	41.9	31.4	
Prediction (Price available plans)	All respondents	20.1	32.0	45.7 *	
Prediction (Price chosen plans)	All respondents	24.8	37.0	38.1 *	
Prediction (Price available plans)	Only those with Part D	22.8	31.8	45.4 *	
Prediction (Price chosen plans)	Only those with Part D	30.9	38.8	30.3 *	
Real Market				Generics	Generics and Brand
All available plans 2006		34.0	50.6	12.9	2.5
Plans actually chosen in the RPS		36.3	54.3	4.8	4.6

* Price is the weighted sum of generics and generics and brand coverage, with the weights given by the market shares

Source: The prices and market shares of supplied plans are taken from Heiss, McFadden, Winter (2009)

Table 5.7: Multinomial logit analysis of hypothetical choices

	(1)	(2)	(3)	(4)	(5)
Plan attributes					
Reference group: no coverage					
Premium	-0.0263***	-0.0286***	-0.0298***	-0.0301***	-0.0311***
Insurance	1.100***	1.233***	1.487***	0.913***	1.256***
No deductible	0.652***	0.766***	0.618***	0.561***	0.459**
Gap coverage	0.730***	0.435***	0.384***	0.435***	0.200
Real coverage					
Reference group: active deciders					
Passive x insurance				0.456***	0.413**
Passive x no deductible				0.249*	0.230
Passive x gap coverage				0.324***	0.262**
2005 drug costs					
Reference group: medium drug costs					
No costs x insurance		-0.571***	-0.604***		-0.581***
No costs x no deductible		-0.603***	-0.593***		-0.580***
No costs x gap coverage		0.186	0.185		0.200
High costs x insurance		0.149	0.195		0.193
High costs x no deductible		0.102	0.174		0.170
High costs x gap coverage		0.472***	0.406***		0.400***
Socio-economic variables					
Female x insurance			-0.123		-0.107
Female x no deductible			0.160		0.167
Female x gap coverage			-0.112		-0.110
Low income x insurance			-0.223		-0.197
Low income x no deductible			-0.245*		-0.227
Low income x gap coverage			-0.239*		-0.219*
Higher education x insurance			-0.116		-0.144
Higher education x no deductible			0.0652		0.0532
Higher education x gap coverage			0.136		0.122
SHRS poor/fair x insurance			-0.116		-0.151
SHRS poor/fair x no deductible			-0.170		-0.182
SHRS poor/fair x gap coverage			0.381***		0.371***
Age > 75 x insurance			-0.0322		-0.0352
Age > 75 x no deductible			0.116		0.118
Age > 75 x gap coverage			-0.111		-0.108
Risk averse x insurance			0.100		0.114
Risk averse x no deductible			0.223		0.228
Risk averse x gap coverage			0.333***		0.334***
Observations	6262	4604	4531	4604	4531

* denotes p<.1, ** denotes p<.05, and *** denotes p<.01 for a two-sided t-test (clustering by respondent)

Table 5.8: Willingness-to-pay for insurance attributes by consumer group

	Insurance	No deductible	Gap coverage
Hypothetical choices			
Basic MNL	41.80***	24.77***	27.75***
Passive	45.41***	26.88***	25.16***
Active	30.30***	18.61***	14.42***
Real choices			
Basic MNL	0.91	5.19***	8.30***
Combined model			
Nested logit	35.39***	16.45***	20.55***

* denotes p<.1, ** denotes p<.05, and *** denotes p<.01 for a two-sided t-test.

Table 5.9: Multinomial logit analysis of revealed choices

	(1)	(2)	(3)	(4)
Plan attributes				
Premium	-0.0938***	-0.0951***	-0.0977***	-0.0792***
Insurance	0.0851	0.0595	0.178	0.356
No deductible	0.487***	0.743***	0.494***	0.896***
Gap coverage	0.779***	0.234		
Gap coverage (generics)			0.305	0.158
Gap coverage (brand-name drugs)			2.519***	1.785***
No costs x insurance		-1.323***		
No costs x no deductible		-0.724*		
No costs x gap coverage		-21.61***		
High costs x insurance		1.110***		
High costs x no deductible		-0.424*		
High costs x gap coverage		1.121***		
Drug tiers				-0.668***
Mailorder				0.567*
Top 100 drugs uncovered				-0.0891***
Top 100 drugs with authorization				-0.0598***
Observations	469	469	469	469

* denotes p<.1, ** denotes p<.05, and *** denotes p<.01 for a two-sided t-test.

Table 5.10: Nested logit analysis of combined stated and revealed choices

	Coefficient	SD	95% Confidence intervall	
Plan attributes				
Premium	-0.034***	0.004	[-0.043,	-0.026]
Insurance	1.208***	0.152	[0.911,	1.505]
No deductible	0.561***	0.065	[0.435,	0.688]
Gap coverage	0.701***	0.084	[0.536,	0.866]
Dissimilarity parameters				
τ_hypothetical	0.720	0.079	[0.566,	0.875]
τ_real	1			

LR test for IIA (τ =1): $\chi^2(1)$=11.64 Prob>χ^2=0.0006

Log likelihood = -9766.1763

Wald $\chi^2(4)$=93.73 Prob>χ^2=0.0000

* denotes p<.1, ** denotes p<.05, and *** denotes p<.01 for a two-sided t-test.

List of Figures

List of Tables

Bibliography

Abaluck, J. T. and J. Gruber (2011): Choice inconsistencies among the elderly: Evidence from plan choice in the Medicare Part D program. *American Economic Review*, forthcoming.

Ackerberg, Benkard, Berry, and Pakes (2007): Econometric tools for analyzing market outcomes. *Handbook of Econometrics*, 6, 4172–4271.

Ackerman, R. (19/09/2008): Venezuela seizes Cemex. *Forbes.com*, www.forbes.com.

Adamowicz, W., J. Swait, P. Boxall, J. Louviere, and M. Williams (1997): Perceptions versus objective measures of environmental quality in combined revealed and stated preference models of environmental valuation. *Journal of Environmental Economics and Management*, 32, 65–84.

Aggarwal, R. and N. A. Kyaw (2004): Internal capital markets and capital structure choice of U.S. multinationals' affiliates. *mimeo*.

Ainslie, R. (2000): Annuity and insurance products for impaired lives. *Paper presented to the Staple Inn Actuary Society.*

Andersson, T. (1991): *Multinational Investment in Developing Countries: A Study of Taxation and Nationalization*. Routledge, Oxfordshire.

Ania, A. B., T. Troger, and A. Wambach (2002): An evolutionary analysis of insurance markets with adverse selection. *Games and Economic Behavior*, 40(2), 153–184.

Asiedu, E. and H. Esfahani (2001): Ownership structure in foreign direct investment projects. *Review of Economics and Statistics*, 83(4), 647–662.

Banerjee, A. (2005): New development economics and the challenge to theory. *Economic and Political Weekly*, 40, 4340–4344.

Bayraktar, N., P. Sakellaris, and P. Vermeulen (2005): Real versus financial constraints to capital investment. *European Central Bank. Working Paper*, 566, Frankfurt a. M.

Beck, T., A. Demirgüc-Kunt, and R. Levine (2000): A new database on financial development and structure. *World Bank Economic Review*, 14, 597–605.

Becker, K. and P. Zweifel (2008): Age and choice in health insurance: Evidence from a discrete choice experiment. *The Patient*, 1(1), 27–40.

Bellone, F., P. Musso, L. Nesta, and S. Schiavo (2008): Financial constraints and firm export behavior. *Department of Economics (University of Trento) Working Paper No. 0816.*

Ben Akiva, M. and T. Morikawa (1990): Estimation of switching models from revealed preferences and stated intentions. *Transportation Research A*, 24A, 485–495.

Berle, A. and G. Means (1932): *The Modern Corporation and Private Property.* Macmillan, New York.

Berman, N. and J. Héricourt (2010): Financial factors and the margins of trade: Evidence from cross-country firm-level data. *Journal of Development Economics*, forthcoming.

Bernard, A. B., J. Eaton, J. B. Jenson, and S. Kortum (2003): Plants and productivity in international trade. *American Economic Review*, 93(4), 1268–1290.

Bernard, A. B. and J. B. Jensen (1995): Exporters, jobs and wages in U.S. manufacturing: 1976-1987. *Department of Economics (MIT) Working Paper No. 95-7.*

Bernheim, B. D. and M. D. Whinston (1990): Multimarket contact and collusive behavior. *RAND Journal of Economics*, 21(1), 1–26.

Bhat, C. and S. Castelar (2002): A unified mixed logit framework for modelling revealed and stated preferences: Formation and application to congestion pricing analysis in the San Francisco Bay Area. *Transportation Research Part B*, 36, 593–616.

Bikker, J. A. and M. van Leuvensteijn (2008): Competition and efficiency in the Dutch life insurance industry. *Applied Economics*, 40(16), 2063–2084.

Bradley, M. and A. Daly (1992): Estimation of logit choice models using mixed stated preference and revealed preference information. *6th International Conference on Travel Behaviour.*

Brau, R. and M. Bruni (2008): Eliciting the demand for long-term care coverage: A discrete choice modelling analysis. *Health Economics*, 17, 411–433.

Brealey, R. A. and S. C. Myers (2000): *Principles of Corporate Finance.* McGraw-Hill, New York.

Brown, J. R., S. M. Fazzari, and B. C. Petersen (2009): Financing innovation and growth: Cash flow, external equity, and the 1990s R&D boom. *Journal of Finance*, 64(1), 151 –185.

Brown, J. R. and A. Finkelstein (2007): Supply or demand: Why is the market for long-term care insurance so small? *Journal of Public Economics*, 91, 1967–1991.

Brownstone, D., D. Bunch, and K. Train (2000): Joint mixed logit models of stated an revealed preferences for alternative-fuel vehicles. *Transportation Research B*, 34, 315–338.

Brusco, S. and F. Panunzi (2005): Reallocation of corporate resources and managerial incentives in internal capital markets. *CEPR Discussion Paper No. 2532.*

Buch, C., I. Kesternich, A. Lipponer, and M. Schnitzer (2010): Exports versus FDI revisited: Does finance matter? *Deutsche Bundesbank Discussion Paper 03/2010 and CEPR Discussion Paper*, 7839.

Buckley, A. (2003): *Multinational Finance*, 5th edition. Financial Times Press, New Jersey, NJ.

Buettner, T., M. Overesch, U. Schreiber, and G. Wamser (2009): Taxation and capital structure choice - Evidence from a panel of German multinationals. *Economics Letters*, 105 (3), 309–311.

Buzzacchi, L. and T. Valletti (2005): Strategic price discrimination in compulsory insurance markets. *Geneva Papers on Risk and Insurance Theory*, 30(1), 71–97.

Cameron, C. A. and P. K. Trivedi (2005): *Microeconometrics: Methods and Applications.* Cambridge University Press, New York.

Carlsson, F. and P. Martinsson (2001): Do hypothetical and actual marginal willingness to pay differ in choice experiments? *Journal of Environmental Economics and Management*, 41, 179–192.

Carson, R. T., N. E. Flores, K. M. Martin, and J. L. Wright (1996): Contingent valuation and revealed preference methodologies: Comparing the estimates for quasi-public goods. *Land Economics*, 72, 80–99.

Carter, A. (2005): Telling the risky from the reliable. *Business Week*, 57.

Cawley, J. and T. Philipson (1999): An empirical examination of information barriers to trade in insurance. *American Economic Review*, 89(4), 827–846.

Centers for Medicare and Medicaid Services (CMS) (2008): National health expenditure data.

Chaney, T. (2005): Liquidity constrained exporters. mimeo.

Chiappori, P.-A., B. Jullien, B. Salanie, and F. Salanie (2006): Asymmetric information in insurance: General testable implications. *RAND Journal of Economics*, 37(4), 783–798.

Chiappori, P.-A. and B. Salanie (2008): Modeling competition and market equilibrium in insurance: Empirical issues. *American Economic Review: Papers and Proceedings*, 98(2), 146 –150.

Cutler, D. M., A. Finkelstein, and K. McGarry (2008): Preference heterogeneity and insurance markets: Explaining a puzzle of insurance. *American Economic Review: Papers and Proceedings*, 98(2), 157–162.

Dafny, L., M. Duggan, and S. Ramanarayanan (2009): Paying a premium on your premium? Consolidation in the U.S. health insurance industry. *NBER Working Paper No. w15434.*

Demsetz, H. (1983): The structure of ownership and the theory of the firm. *Journal of Law and Economics*, 26, 375−390.

Desai, M. A., C. F. Foley, and J. R. Hines (2004): A multinational perspective on capital structure choice and internal capital markets. *Journal of Finance*, 59(6), 2451−2487.

Desai, M. A., C. F. Foley, and J. R. Hines (2008): Capital structure with risky foreign investment. *Journal of Financial Economics*, 88 (3), 534−553.

Deutsche Presse Agentur (27/09/2007): Simbabwe plant Enteignung ausländischer Firmen. *Published in Netzeitung.de.*

Duggan, M., P. Healy, and F. Scott-Morton (2008): Providing prescription drug coverage to the elderly: America's experiment with Medicare Part D. *Journal of Economic Perspectives*, 22(4), 69−92.

Duggan, M. and F. Scott-Morton (2006): The distortionary effects of government procurement: Evidence from Medicaid prescription drug purchasing. *Quarterly Journal of Economics*, 32 (1), 1−30.

Duggan, M. and F. Scott-Morton (2010): The effect of Medicare Part D on pharmaceutical prices and utilization. *American Economic Review*, 100(1), 590−607.

Eaton, J. and M. Gersovitz (1983): Country risk: economic aspects. In R. J. Herring (Ed.), *Managing International Risk*, 75−108. Cambridge, UK: Cambridge University Press.

Eaton, J. and M. Gersovitz (1984): A theory of expropriation and deviations from perfect capital mobility. *Economic Journal*, 94(16-40s), 16−40.

Eichengreen, B. (2009): The last temptation of risk. *The National Interest*, May/ June.

Etro, F. (2008): Stackelberg competition with endogenous entry. *Economic Journal*, 118(532), 1670−1697.

Fama, E. and M. Jensen (1983): Separation of ownership and control. *Journal of Law and Economics*, 26, 301−325.

Finkelstein, A. and J. Poterba (2006): Testing for adverse selection with "unused observables". *NBER Working Paper No.12112.*

Greenaway, D., A. Guariglia, and R. Kneller (2007): Financial factors and exporting decisions. *Journal of International Economics*, 73(2), 377−395.

Guttmann, R., R. Castle, and D. G. Fiebig (2009): Use of discrete choice experiments in health economics: An update of the literature. *CHERE Working Paper No. 2009/2, University of Technology, Sydney.*

Hanley, N., M. Ryan, and R. Wright (2003): Estimating the monetary value of health care: Lessons from environmental economics. *Health Economics*, 12, 3–16.

Harrington, J. E. (1989): Collusion and predation under (almost) free entry. *International Journal of Industrial Organization*, 7(3), 381–401.

Harrison, A. E., I. Love, and M. Mcmillan (2004): Global capital flows and financing constraints. *Journal of Development Economics*, 75(1), 269–301.

Hausman, J. and D. McFadden (1984): Specification test for the multinomial logit model. *Econometrica*, 52, 1219–1240.

Heiss, F., D. McFadden, and J. Winter (2006): Who failed to enroll in Medicare Part D, and why? Early results. *Health Affairs*, 25, 344354.

Heiss, F., D. McFadden, and J. Winter (2009a): Mind the gap! Consumer perceptions and choices of Medicare Part D prescription drug plans. In D. Wise (Ed.), *Research Findings in the Economics of Aging*, Chicago University Press, Chicago, IL.

Heiss, F., D. McFadden, and J. Winter (2009b): Regulation of private health insurance markets: Lessons from enrollment, plan type choice, and adverse selection in Medicare Part D. *NBER Working Paper No.15392.*

Helpman, E., M. J. Melitz, and S. R. Yeaple (2004): Export versus FDI. *American Economic Review*, 94, 300–316.

Hensher, D. and M. Bradley (1993): Using stated response data to enrich revealed preference discrete choice models. *Marketing Letters*, 4, 39–152.

Hensher, D., J. Louviere, and J. Swait (1999): Combining sources of preference data. *Journal of Econometrics*, 89, 197–221.

Hill, C. A. (1998): How investors react to political risk. *Duke Journal of Comparative and International Law*, 8, No 2.

Howell, L. D. (2008): *Political Risk Assessment: Concept, Method, and Management.* The PRS Group, East Syracuse, NY.

Hoy, M. (1982): Categorizing risks in the insurance industry. *Quarterly Journal of Economics*, 97(2), 321–336.

Hoy, M. and M. Polborn (2000): The value of genetic information in the life insurance market. *Journal of Public Economics*, 78, 235–252.

Hoy, M. and J. Witt (2007): Welfare effects of banning genetic information in the life insurance market: The case of BRCA1/2 genes. *Journal of Risk and Insurance*, 74(3), 523–546.

Huizinga, H., L. Laeven, and G. Nicodeme (2008): Capital structure and international debt shifting. *Journal of Financial Economics*, 88(1), 80−−118.

Inderst, R. and C. Laux (2005): Incentives in internal capital markets: Capital constraints, competition, and investment opportunities. *RAND Journal of Economics*, 36(1), 215−228.

Ivaldi, M. (1996): Sinistralité des jeunes conducteurs novices, incitations a la sécurité routiére et tarifications de l'assurance automobile. *IDEI Report, Toulouse*.

Janeba, E. (2000): Tax competition when government lack commitment: excess capacity as a countervailing threat. *American Economic Review*, 90, 1508−1519.

Javorcik, B. S. and S.-J. Wei (2009): Corruption and cross-border investment in emerging markets: Firm-level evidence. *Journal of International Money and Banking*, forthcoming.

Jensen, M., W. Meckling, S. Field, and T. Park (1976): Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics (JFE)*, 3(4).

Jones, A., N. Rice, T. Bago d'Uva, and S. Balia (2007): *Applied Health Economics.* Routledge, Oxfordshire.

Kaplan, S. N. and L. Zingales (1997): Do investment-cash flow sensitivities provide useful measures of financial constraints? *The Quarterly Journal of Economics*, 112, 169−215.

Keynes, J. (1940): On a method of statistical business-cycle research. A comment. *Economic Journal*, 50(197), 154−156.

Kobrin, S. J. (1980): Foreign enterprise and forced divestment in ldcs. *International Organization*, 41, 609−638.

Konrad, K. and K. E. Lommerud (2001): Foreign direct investment, intra-firm trade and ownership structure. *European Economic Review*, 45(3), 475−−494.

Kraus, A. and R. H. Litzenberger (1973): A state-preference model of optimal finance leverage. *Journal of Finance*, 28, 911−922.

Levy, H. and D. Weir (2009): Take-up of Medicare Part D: Results from the Health and Retirement Study. *Journal of Gerontology: Social Sciences*, 65B(4), 492−501.

Lipponer, A. (2006): *Microdatabase Direct Investment - MiDi. A brief guide.* Frankfurt: Bundesbank Working Paper.

Louviere, J., D. Hensher, and J. Swait (2000): *Stated Choices Methods.* Cambridge University Press, Cambridge.

Luce, D. (1959): *Individual Choice Behaviour*. John Wiley and Sons, New York.

Maddala, G. (2001): *Introduction to Econometrics*, 3 edition. John Wiley and Sons, New York.

Madden, J., A. Graves, F. Zhang, A. Adams, B. Briesacher, D. Ross-Degnan, J. Gurwitz, M. Pierre-Jacques, D. Safran, G. Adler, and S. Soumerai (2008): Cost-related medication nonadherence and spending on basic needs following implementation of Medicare Part D. *JAMA*, 299(16), 1922–1928.

Mailath, G. and L. Samuelson (2006): *Repeated Games and Reputations*. Oxford University Press, Oxford.

Manova, K. (2010): Credit constraints, heterogeneous firms, and international trade. *NBER Working Paper No.14531.*

Markusen, J. (2002): *Multinational Firms and the Theory of International Trade*. MIT Press, Cambridge, MA.

Marschak, J. (1960): Binary choice constraints and random utility indicators. *Cowles Foundation Discussion Paper No.74.*

Mayer, T. and G. I. Ottaviano (2008): The happy few: The internationalisation of European firms. *Intereconomics: Review of European Economic Policy*, 43(3), 135–148.

McConnell, J. and H. Servaes (1990): Additional evidence on equity ownership and corporate value. *Journal of Financial Economics*, 27, 595–612.

McFadden (1974): The measurement of urban travel demand. *Journal of Public Economics*, 3(4), 303–328.

McFadden, D. (1976): The revealed preferences of a government bureaucracy: Empirical evidence. *Bell Journal of Economics*, 7(1), 55–72.

McFadden, D. (1978): Modeling the choice of residential location. In A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull (Eds.), *Spatial Interaction Theory and Planning Models*, 75–96.

McFadden, D. (1987): Regression-based specification tests for the multinomial logit model. *Journal of Econometrics*, 34, 63–82.

McFadden, D., J. Winter, and F. Heiss (2008): Consumer-directed health care: Can consumers look after themselves? *Swiss Journal of Economics and Statistics*, 144(III), 285–307.

Melitz, M. (2003): The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6), 1695–1725.

Mintz, J. and A. J. Weichenrieder (2005): Taxation and the financial structure of German outbound FDI. *CESifo Working Paper No. 1612.*

Mitchell, R. and R. Carson (1989): *Using Surveys to Value Public Goods: The Contigent Valuation Method.* RFF Press, Washington, D.C.

Mookherjee, D. (2005): Is there too little theory in development economics? *EPW Commentary.*

Morck, R., A. Shleifer, and R. W. Vishny (1988): Management ownership and market valuation. An empirical analysis. *Journal of Financial Economics*, 20, 293–315.

Morikawa, T. (1989): *Incorpotating stated preference data in travel demand analysis.* PhD thesis, Department of Civil Enginerring, MIT.

Mueller, T. and M. Schnitzer (2006): Technology transfer and spillovers in international joint ventures. *Journal of International Economics*, 68(2), 456–468.

Neuman, P., M. Strollo, S. Guterman, W. Rogers, A. Li, A. Rodday, and D. Safran (2007): Medicare prescription drug benefit progress report: Findings from a 2006 national survey of seniors. *Health Affairs*, 26, 630–643.

Novaes, W. and S. Werlang (2005): Political risk and capital structure choice of foreign subsidiaries: An empirical analysis. *mimeo.*

Polborn, M. K., M. Hoy, and A. Sadanand (2006): Advantageous effects of regulatory adverse selection in the life insurance market. *Economic Journal*, 116(508), 327–354.

Redding, S. (2008): Empirical approaches to international trade. In S. N. Durlauf and L. E. Blume (Eds.), *New Palgrave Dictionary of Economics.* Palgrave Macmillan, Basingstoke, second edition.

Rees, R. and P. Apps (2006): Genetic testing, income distribution and insurance markets. *CHERE Working Paper No.2006/3, University of Technolgy, Sydney.*

Rees, R. and E. Kessner (1999): Regulation and efficiency on European insurance markets. *Economic Policy*, 14(29), 365–397.

Rees, R. and A. Wambach (2008): The microeconomics of insurance. *Foundations and Trends in Microeconomics*, 4(1-2), 1–163.

Romero, S. (18/05/2008): Chavez orders increased nationalization in venezuela. *San Francisco Chronicle*, www.sfgate.info.

Ross, S. A. (1977): The determination of financial structure: the incentive-signalling approach. *The Bell Journal of Economics*, 8, 23–40.

Rothschild, N. and J. E. Stiglitz (1976): Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics*, 90(4), 630–649.

Rupper, R., B. D. Bair, B. C. Sauer, J. R. Nebeker, J. Shinogle, and M. Samore (2007): Out-of-pocket pharmacy expenditures for veterans under Medicare Part D. *Medical Care*, 45, 77–80.

Russ, K. (2007): The endogeneity of the exchange rate as a determinant of FDI: A model of money, entry, and multinational firms. *Journal of International Economics*, 71(2), 267–526.

Ryan, M. and K. Gerard (2003): Using discrete choice experiments to value health care programmes: Current practice and future research reflections. *Applied Health Economics and Health Policy*, 2, 55–64.

Scanlon, D., M. Chernew, and J. Lave (1997): Consumer health plan choice: Current knowledge and future directions. *Annual Review of Public Health*, 18, 507–528.

Schnitzer, M. (1999): Expropriation and control rights: A dynamic model of foreign direct investment. *International Journal of Industrial Organization*, 17, 1113–37.

Schnitzer, M. (2002): Debt vs. foreign direct investment: The impact of sovereign risk on the structure of international capital flows. *Economica*, 69, 41–67.

Scott, J. H. (1976): A theory of optimal capital structure. *The Bell Journal of Economics*, 7, 33–54.

Short, H. and K. Keasey (1999): Managerial ownership and the performance of firms: Evidence from the U.K. *Journal of Corporate Finance*, 5, 79–101.

Smart, M. (2000): Competitive insurance markets with two unobservables. *International Economic Review*, 41(1), 153–169.

StataCorp (2007): *Stata Base Reference Manual Vol.3 Release 10.* Stata Press, College Station, TX.

Stein, J. C. (2003): Agency, information and corporate investment. In G. Constantinides, M. Harris, and R. Stulz (Eds.), *Handbook of the Economics of Finance*. Elsevier, Amsterdam.

Stiebale, J. (2008): Do financial constraints matter for foreign market entry? A firm-level examination. *Ruhr Economic Paper No. 0051.*

Strohmenger, R. and A. Wambach (2000): Adverse selection and categorical discrimination in the health insurance markets: The effects of genetic tests. *Journal of Health Economics*, 19(2), 197–218.

The Associated Press (01/05/2008): Chavez enteignet grössten stahlkonzern venezuelas. *Published in NZZ Online*, www.nzz.ch.

The Henry J. Kaiser Family Foundation (2008): *Medicare Part D prescription drug plan (PDP) availability.* Technical report.

Thomas, J. and T. Worrall (1994): Foreign direct investment and the risk of expropriation. *Review of Economic Studies*, 61, 81–108.

Tirole, J. (1988): *The Theory of Industrial Organization.* MIT Press, Cambridge and London.

Train, K. (1986): *Qualitative Choice Analysis.* MIT Press, Cambridge, MA.

Train, K. (2003): *Discrete Choice Methods with Simulation.* Cambridge University Press, Cambridge, MA.

Train, K., M. Ben-Akiva, and T. Atherton (1989): Consumption patterns and self-selecting tariffs. *Review of Economics and Statistics*, 71, 62–73.

Train, K. and W. Wilson (2009): Monte Carlo analysis of SP-off-RP data. *Journal of Choice Modelling*, 2, 101–117.

U.S. Departement of Health and Human Services (2006): *News Release: Over 38 Million people with Medicare now receiving prescription drug coverage.* Technical report.

Voice of America (11/03/2008): Zimbabwe says no expropriation of foreign companies. *VOA News*, www.voanews.com.

Walker, J. and M. Ben-Akiva (2002): Generalized random utility model. *Mathematical Social Sciences*, 43(3), 303–343.

Winter, J., R. Balza, F. Caro, F. Heiss, B.-H. Jun, R. Matzkin, and D. McFadden (2006): Medicare prescription drug coverage: Consumer information and preferences. *Proceedings of the National Academy of Sciences of the United States of America*, 103(20), 7929 –7934.

Zellner, A. (1962): An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association*, 57, 348–68.

Curriculum Vitae

Since September 2008	Research and Teaching Assistant Chair of Prof. Dr. Monika Schnitzer
October 2005 - May 2010	Ph.D. Programme in Economics Munich Graduate School of Economics Ludwig-Maximilians-Universität, Munich
September 2007 - March 2008	OECD, Paris Directorate for Employment, Labour and Social Affairs Traineeship with the Carlo Schmid Programme
April 2001 - January 2006	Studies of Economics (Dipl. Volksw.) Ludwig-Maximilians-Universität, Munich, and Universidad Carlos III, Madrid (ERASMUS)
September 1999 - March 2001	Studies of Financial Mathematics Technische Universität, München, and Universität Ulm
June 1999	Abitur Irmgardis-Gymnasium, Köln