Essays on Natural Experiments in Behavioral Finance and Trade

Inaugural-Dissertation
zur Erlangung des Grades
Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

2009

vorgelegt von

Wolf Nicolas Sauter

Referent:	Prof. Dr. Joachim Winter
Koreferent:	Prof. Dr. Thiess Büttner
Promotionsabschlussberatung:	10. Februar 2010

To my parents

Acknowledgments

The development and completion of this thesis would not have been possible without the help and the encouragement of numerous people. First of all, I would like to thank my supervisor Prof. Joachim Winter, whose teachings, uninterrupted support, and guidance made this dissertation possible. I also thank Prof. Thiess Büttner who kindly accepted to be my co-supervisor and Prof. Florian Englmaier for being my third examiner.

I thank all participants at the Empirical Research Workshop and my colleagues at the *Seminar für Empirische Wirtschaftsforschung* at LMU Munich, who provided substantial feedback and many helpful ideas during the research process. I am particularly indebted to Florian Heiss for extensive feedback on the first chapter of this dissertation, for his teachings and for making his Stata code publicly available. Also, I am grateful to Jan Walliser, who provided important inputs on the elaboration of the second chapter. I especially thank Amelie Wuppermann for her assistance in proofreading the final draft of my study.

The stimulating environment at the Munich Graduate School of Economics (MGSE) provided me with important input throughout my dissertation. For their help and support, I thank particularly Prof. Rady, Ms Buchmayr, Ms Bierprigl, and Ms Szantone-Sturm. Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged.

My greatest gratitude, however, goes to my parents for their support and encouragement and to Caroline for her enduring love and patience when my attention was focused on this dissertation.

Contents

List of Tables

List of Figures

Preface

The eventual goal of scientific research is to identify causal relationships between different classes of objects in order to enrich our understanding of the world. To accomplish this goal, natural scientists use repeated experiments to test their explanations of causal relations in nature. The validity of their theories is then inferred empirically by inductive reasoning, based on the implicit assumption that natural laws do not change.

Social sciences, and economics in particular, face more difficulties in identifying causal relationships due to human nature. People adapt their behavior to their living circumstances. Therefore, 'laws of human behavior' cannot immediately be inferred from repeated observations of individuals, since each time an observation is made, the individual conditions his or her behavior on past experiences and behaves differently in the future. For instance, a person faced with the offer to buy the same stocks in two subsequent periods will not automatically take the same decision, even if his expectations about future returns do not change. In the second period, his behavior will depend on the stock market development and the choices made in the first period (see chapter 1). Hence, omitting experiences individuals made in the past would bias the results.

A further challenge for the identification of causal effects in economics is that most real world choices are interdependent. For example, observing wealthy individuals investing into stocks does not allow concluding that wealth increases the likelihood of stock ownership, because individuals could also be wealthy because they invest into stocks.

During the last years, experiments became increasingly popular among economists to solve both the omitted variables problem and the problem of endogeneity. Three types of experiments can be distinguished: (i) laboratory experiments, (ii) controlled field studies or randomized field experiments, and (iii) natural experiments.

The use of lab experiments amplified due to the growing interest in issues such as economic cooperation, trust, and neuroeconomics, e.g. Gürerk, Irlenbusch, Rockenbach (2006) and Fischbacher, Kosfeld, Fehr (2005). In such experiments, a certain treatment is randomly assigned to one group of individuals in order to compare their economic actions to an untreated control group within the artificial environment of the lab. By construction, such virtual experiments exclude feedback effects and reduce potential bias from omitted factors. However, it is unclear if the results from the lab translate into real world decisions, because people are likely to behave differently if they make decisions that affect their own lives and their own money (Gul, Pesendorfer, 2005).

An alternative to lab experiments are controlled field experiments, which also randomize treatments but do so in real world applications. Development economists particularly are increasingly trying to identify the efficiency of their measures by randomizing the allocation of, for instance, aid or political leadership across the population (Gertler, 2004; Chattopadhyay, Duflo, 2004). Average effects on people's behavior can then be consistently estimated by comparing behavior before and after the allotment under relatively weak identifying assumptions. However, Deaton (2009) forcefully argues that consistent statistical interference does not guarantee that such an estimate has a meaningful interpretation, as it only says something about 'what works' but nothing about 'why something works'. Also, there are substantial moral constraints to randomized field experiments. In the worst case, scarce resources would not be given to those with the most urgent needs or the best leadership skills but would merely be distributed randomly across the population for the sake of statistical identifiability.

Natural experiments have the same straightforward identification approach as controlled field studies do. Yet they are not subject to moral constraints. A natural experiment occurs when some feature of the real world is randomly changed in a way that allows using the exogenous variation due to this change in order to study causal effects of an otherwise endogenous explanatory variable. For instance, an unanticipated tax reform that only affects one group of investors facilitates the identification of a causal link between taxation and portfolio choices (see chapters 2 and 3). Thus, one reason for the popularity of natural experiments in economic research is that they typically offer an intuitive interpretation of the underlying identifying assumptions and enable a broader audience to check their consistency compared to purely statistical identification approaches.

This dissertation consists of five self-contained chapters that are contributions from research in two areas: behavioral finance and international trade. Each chapter has its own introduction, references, and appendix and the five parts can be read independently from each other. Still, to some extent, the five chapters can be subsumed under the common theme of "natural experiments", as they all make use of changes in real world environments in order to identify causal effects.

Chapters 1, 2, and 4 use the German reunification in 1990 as a natural experiment to study subsequent savings choices. Reunification was unanticipated and prompted substantial changes in economic behavior among East Germans. Recent studies exploiting this natural experiment are Frijters, Haisken-DeNew, Shields (2004), Alesina and Fuchs-Schündeln (2007), and Redding and Sturm (2008). Also, post-reunification changes in savings preferences have been studied for East German households. Building on the fact that job choices were exogenously restricted before reunification, Fuchs-Schündeln and Schündeln (2005) find that self-selection of risk-averse individuals into low-risk occupations in West Germany relative to the East is economically important and decreases aggregate precautionary wealth holdings significantly. Fuchs-Schündeln (2008) also shows that East Germans have higher savings rates than West Germans and that this East-West gap is increasing in age.

In contrast to the above-mentioned studies, chapters 1, 2, and 4 study portfolio choices rather than changes in the total wealth or savings of households. Moreover, no comparison takes place between East and West German households. Rather, the unique natural experiment of German reunification allows me to eliminate different potential sources of bias, which are common to other studies, in order to identify causal effects within East Germany only.

Chapter 1 analyzes the non-participation puzzle among East German households from 1990 to 2006. The non-participation puzzle states that a large proportion of households in industrialized countries does not own securities despite an expected return premium on risky assets (Guiso, Haliassos, Jappelli, 2003). The most widely accepted explanation for this puzzle is that transaction costs deter households from investing small amounts of money. The fact that East German investors of all age groups did not have any prior experience with securities before 1990 helps to identify the importance of different transaction costs motives such as trading costs, habit persistence, and investment experience on participation decisions. Habit persistence and lack of investment experience appear as the main reasons for the widespread non-participation, whereas pure trading costs motives have been largely overestimated in the past. There is no evidence that historical market returns, which investors experienced over their life-cycles, affect their risk-taking behavior. These findings are particularly relevant for policy-makers and financial consultants seeking to establish household portfolios as an additional pillar of the social security system. In order to increase the acceptance of private savings plans, they will have to override households' habit persistence and reservation with regard to risky assets. Hence, the high degree of habit persistence in savings decisions strengthens the case for policies that raise the awareness for securities. Also, the prevalent habit persistence advocates the increased use of savings plans which do not require people to opt in but rather allow opting out of private retirement savings schemes.

Chapter 2 studies two natural experiments to identify the effects of tax incentives and bequest motives on life-insurance demand. An unanticipated tax reform in 2000 halved the tax exemption limit for capital income in Germany. We document that the demand for life insurance reacted strongly to this change. This result points out that standard tax revenue estimates, which assume that current investors would stick to their asset choices if capital taxation were introduced, may be misleading. Governments need to account for changes in investment behavior due to tax reforms. With regard to bequest motives, we analyze the demand for life insurance in the former German Democratic Republic. Relative to market-based economies, the socialist GDR can be viewed as an experimental institutional setting where life-insurance demand was not influenced by tax considerations, which allows us to isolate bequest motives, while controlling for life-cycle and precautionary motives. We find a significantly higher ownership probability among households with children and a high regard for the family, confirming strong bequest motives in life-insurance demand.

Chapter 3 presents additional evidence for the importance of taxation in households' investment decisions. A difference-in-difference analysis shows that a tax reform in Germany which revoked the tax exemption of life insurance returns triggered a significant increase in demand prior to the reform.

Chapter 4 uses the natural experiment of German reunification to study the adoption of East Germans to building society contracts (BSCs) after 1990. A striking feature of research on financial innovation is the relative dearth of empirical studies analyzing the adoption of households to new financial products. This chapter analyzes (i) who uses BSCs to save for a house purchase, and (ii) how long it takes after reunification until future savers start saving. We find that households with close ties to their families in the West invest earlier, suggesting the presence of information asymmetries in the adoption processes. There is also evidence that households trade-off long-term savings goals for short-term consumption, because initial car owners are more likely to invest into BSCs and to do so earlier. Life insurance appears to be a substitute for BSCs.

Chapter 5 analyzes differences in the structure of intra-Canadian trade that are due to language barriers between French and English speakers. While the existence of a language barrier in trade has been documented in numerous empirical studies (Rose, 2000; Anderson and van Windcoop, 2004; Fidrmuc and Fidrmuc, 2008), these studies remain silent about the channel through which language impacts trade. Using bilingual regions in Canada to study the effect of language commonality on trade, I test for one specific mechanism that can explain the existence of a language barrier to trade.

Specifically, I ask if those industries that require more cross-border communication in order to export their products trade more between Canadian provinces that know each other's language(s). Identification comes from the fact that some Canadian regions introduced minority-language-friendly legislation in the 1970s, which is uncorrelated with trade patterns in 2001. There is robust evidence that trade in industries with a need to communicate directly (orally) with importers increases with the probability that people in another province speak the same language. This finding reveals a potential disadvantage for minority language regions in services trade and might also help to explain part of the observed specialization across industries and sectors in international trade.

In conclusion, the five chapters of this dissertation provide new insights in the economic analysis of the areas of behavioral finance and international trade, studying unique natural experiments that permit the identification of causal relationships. Chapter 1 emphasizes the presence of substantial persistence in investment habits and analyzes the impact of investment experience on the probability that households own risky assets. The policy reforms studied in chapters 2 and 3 provide compelling evidence for the importance of tax considerations in households' savings choices. Chapter 4 shows that social networks and consumption-savings trade-offs determine the adoption process of investors to new financial products. Finally, chapter 5 investigates the importance of language barriers in trade flows, exploiting the unique language composition of Canada. All five chapters are part of a broader research agenda that aims at identifying causal effects in applied economic research. Due to their straightforward interpretability, natural experiments appear well-suited to form a key pillar of this effort in the future.

References

ALESINA, ALBERTO and NICOLA FUCHS-SCHÜNDELN (2007): "Good-Bye Lenin (or Not?): The Effect of Communism on People's Preferences," *American Economic Review*, **97**, 1507-1528.

ANDERSON, JAMES E. and ERIC VAN WINCOOP (2004): "Trade Costs," *Journal of Economic Literature*, **42:3**, 691-751.

CHATTOPADHYAY, RAGHABENDRA and ESTHER DUFLO (2004): "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India," *Econometrica*, **72**, 1409-1443.

DEATON, ANGUS (2009): "Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development," NBER Working Paper 14690.

FEHR, ERNST; URS FISCHBACHER and MICHAEL KOSFELD (2005): "Neuroeconomic Foundations of Trust and Social Preferences," *American Economic Review - Papers and Proceedings*, **95**, 346-351.

FIDRMUC, JAN and JARKO FIDRMUC (2008): "Foreign Languages and Trade," CEPR Discussion Paper 7228.

FUCHS-SCHÜNDELN, NICOLA (2008): "The Response of Household Saving to the Large Shock of German Reunification," *American Economic Review*, **98**, 1798-1828.

FRIJTERS, PAUL; JOHN P. HAISKEN-DENEW and MICHAEL A. SHIELDS (2004): "Money Does Matter! Evidence from Increasing Real Income and Life Satisfaction in East Germany Following Reunification," *American Economic Review*, **94**, 730-740.

GERTLER, PAUL (2004): "Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA's Control Randomized Experiment," *American Economic Review*, **94**, 336-341.

GUL, FARUK and WOLFGANG PESENDORFER (2005): "The Case for Mindless Economics," mimeo, Princeton University.

GÜRERK, ÖZGÜR; BERND IRLENBUSCH and BETTINA ROCKENBACH (2006): "The Competitive Advantage of Sanctioning Institutions," *Science*, **312. no. 5770**, 108-111.

REDDING, STEPHEN J. and DANIEL M. STURM (2008): "The Costs of Remoteness: Evidence from German Division and Reunification," *American Economic Review*, **98**, 1766-1797.

ROSE, ANDREW K. (2001): "One Money, One Market: The Effect of Common Currencies on Trade," *Economic Policy*, **30:1**, 7-46.

Chapter 1

Tearing Down the Wall: (Non-)Participation and Habit Persistence in East German Securities Markets

1.1 Introduction

Non-participation in securities markets has been an increasingly important issue to economists and policy-makers alike. So far there is no consensus as to why a large proportion of households in industrialized countries does not own securities despite an expected return premium on risky assets (Guiso, Haliassos, Jappelli, 2003). Yet policy-makers considering to add an individual-account component to pension systems need to know whether non-participation is a deliberate decision by informed citizens or mere unawareness of the risk and return characteristics of stocks and bonds. While the first would advocate the use of private pension plans, ignorance or financial illiteracy could justify state intervention into retirement savings choices. A better understanding of non-participation behavior might also help to explain the size of the equity premium (Mankiw and Zeldes, 1991).

The most widely accepted explanation for the non-participation puzzle is that transaction costs deter households from entering the market (Haliassos, 2008). Such transaction costs might be the pecuniary costs of trading. Mulligan and Sala-i-Martin (2000) show that small brokerage fees might suffice to explain part of the observed non-participation. However, an increasing number of studies points toward behavioral explanations that focus on non-monetary transaction costs. Brunnermeier and Nagel (2007) find strong persistence in households' holdings of risky assets, which indicates the presence of substantial opportunity costs to overcome inertia. Other non-monetary costs could be related to (i) the lack of knowledge about the risk and return characteristics of certain assets that falls with the accrual of ownership experience (Agarwal et al., 2008) or (ii) individual risk preferences, which were shaped by one's life-cycle experience of bull and bear markets (Malmendier and Nagel, 2007). Yet despite their theoretical importance and empirical persuasiveness, no paper has so far tried to disentangle habit persistence and investment experience from pure trading costs motives in households' portfolio choices.

German reunification provides a quasi-experimental setting to study savings adjustments among East German households (Fuchs-Schündeln and Schündeln, 2005; Fuchs-Schündeln, 2008). In contrast to Anglo-Saxon and Scandinavian countries, Germany does not have a sophisticated equity culture (Calvet, Campbell, Sodini, 2007). Only 33.2% of households owned securities in 2006. Participation is even lower in the East. Sixteen years after reunification, only 20.4% of East German households own stocks and 9.9% hold bonds.[1] The fact that the German Democratic Republic (GDR) barred private ownership of securities offers a unique opportunity to analyze dynamic participation behavior after reunification, because (i) I can control for individual ownership histories between 1990 to 2006, since households had no prior experience with securities, and (ii) there is no initial conditions problem that typically plagues random effects estimators in binary choice models.

The estimated effect of monetary transaction costs on non-participation is much smaller than previously assumed. Rather, non-participation seems to be largely driven by strong persistence in unobserved savings habits and risk preferences. This is in accordance with Alesina and Fuchs-Schündeln (2007), who show that communism left a long-lasting mark on people's preferences for individual risk taking and state intervention in East Germany. I also find that the probability of owning risky assets increases significantly once households gain investment experience. However, there is little indication that macro-experiences matter for investment decisions. These results substantiate the case for policies that raise the awareness for securities and reduce indolence, e.g. default options in retirement accounts (Choi, Laibso, Madrian, 2006).

Related work coming closest to my paper is by Vissing-Jørgensen (2002a). Using the 1984 and 1989 waves of the Panel Study of Income Dynamics (PSID), she finds that the likelihood of participation in 1989 is 31.8% higher for households that have already participated in 1984. However, a dynamic probit estimator is potentially bi-

[1] Author's calculation based the GSOEP.

ased in her two-period panel and the longitudinal dimension is too short to test for persistence in preferences.[2] Alessie and Hochguertel (2002) estimate state dependence among Dutch stock investors, using the CentER Saving panel survey. In their sample, lagged ownership increases the probability of current ownership by 28 percentage points. This estimate is potentially upward biased, as the initial conditions problem is ignored in their estimation strategy. Using the same data, Alessie, Hochguertel, van Soest (2004) estimate a bivariate dynamic probit model in order to explain the dynamics of stock and mutual fund ownership. Yet they do not control for potential autocorrelation in the error structure, which leads to overestimation of the state dependence coefficient, which is 26% for both stocks as well as mutual funds. Finally, Muñoz (2006) estimates a multinominal probit model that distinguishes between persistence in individual heterogeneity and state dependence, using five waves of the Bank of Italy Survey of Household Income and Wealth. She finds strong persistence of individual preferences over time, yet she also ignores the initial conditions problem.

The remainder of the paper is structured as follows: Section 1.2 reviews the existing theoretical explanations for non-participation in securities markets. The identification strategy is laid out in section 1.3. In section 1.4, the estimation approach is described. The data is presented in section 1.5. Estimation results are discussed in section 1.6 before I conclude.

1.2 Theoretical Background

One of the most surprising puzzles of individual asset allocation decisions is the low proportion of households holding stocks and bonds, as documented in numerous studies, e.g. Mankiw and Zeldes (1991), Porterba and Samwick (1995), Bertraut and Haliassos (1995), Rosen and Wu (2004). Various explanations for this puzzle have been suggested. Haliassos and Michaelides (2003) argue that borrowing constraints and limited ability to sell short could solve the puzzle. Other explanations focus on widespread ignorance of securities markets (Guiso and Jappelli, 2005), lack of trust in securities (Guiso, Sapienza, Zingales, 2005), lack of social interaction with other investors (Hong, Kubik, Stein, 2004), or tax laws (Bergstresser and Poterba, 2004). However, the prevailing explanation is that transaction costs discourage small investors from entering the market.

In order to better understand the transaction costs explanation of the non-

[2]For thorough discussions of the bias of dynamic discrete choice models in short panels, see Heckman (1981), Hsiao (1986), Lee (1997).

participation puzzle, I recap the essence of the model by Vissing-Jørgensen (2002a). Consider a household i that lives for multiple periods. Consumption at date t is denoted by C_{it}. Financial wealth is denoted as Y_{it}. The return on the riskless asset held from date t to date $t+1$ is denoted as $r_{z,t+1}$. The stochastic net return on the securities portfolio is denoted by $r_{s,i,t+1}$. At time t, the household chooses fraction ψ_{it} of risky securities in the total portfolio. A household reveals that its certainty equivalent rate of return, $r^{ce}_{s,i,t+1}$, is larger than $r_{z,t+1}$ if it chooses to hold securities, i.e. if $\psi_{it} > 0$. The household optimizes the present value of its lifetime utility from consumption subject to the growth of its initial wealth endowment less consumption:

$$V_t(Y_t) = \max_{C_{it},\psi_{it}} \left\{ U(C_{it}) + \beta V_{t+1}\Big((Y_{it} - C_{it})\big(1 + \psi_{it} r^{ce}_{s,i,t+1} + (1-\psi_{it}) r_{z,t+1}\big)\Big)\right\}, \quad (1.1)$$

where $V_t(Y_t)$ denotes the value function and β the discount factor. Based on the above setup, Samuelson (1969) and Merton (1969) show that households should take positions in all existing assets, and non-participation as well as exit and entry are not observed.

1.2.1 Pecuniary Transaction Costs

Vissing-Jørgensen (2002a), Paiella (2001), and Haliassos and Michaelides (2003) argue that relatively small fixed costs could justify the observed patterns of non-participation. Let household i face fixed pecuniary entry and exit costs of F_{it}. Such costs could for instance be brokerage commissions or the bid-ask spread. Exit costs could include potential re-entry costs, brokerage commissions, or a capital gains tax, which German investors face when they hold asset for less than one year. Let the optimal share of risky investments, $\psi^*_{it} > 0$, be either exogenously given or determined independently of the optimal consumption path, as for example in the case of an isoelastic utility function. Then household i will only buy securities if the present expected value of its consumption from the risky portfolio, incurring entry costs F_{it}, is larger than the value from the riskless portfolio:

$$\begin{aligned} &\max_{C_{it}} \left\{ U(C_{it}) + \beta V_{t+1}\Big((Y_{it} - F_{it} - C_{it})\big(1 + \psi^*_{it} r^{ce}_{s,i,t+1} + (1-\psi^*_{it}) r_{z,t+1}\big)\Big)\right\} \quad (1.2) \\ &\qquad > \max_{C_{it}} \left\{ U(C_{it}) + \beta V_{t+1}\Big((Y_{it} - C_{it})(1 + r_{z,t+1})\Big)\right\}. \end{aligned}$$

For this inequality to be true, $r^{ce}_{s,i,t+1}$ needs to be sufficiently greater than $r_{z,t+1}$. By a similar argument, a household will only exit if

$$\max_{C_{it}} \Big\{ U(C_{it}) + \beta V_{t+1}\Big((Y_{it} - F_{it} - C_{it})(1 + r_{z,t+1}) \Big) \Big\} \tag{1.3}$$

$$> \max_{C_{it}} \Big\{ U(C_{it}) + \beta V_{t+1}\Big((Y_{it} - C_{it})\big(1 + \psi^*_{it} r^{ce}_{s,i,t+1} + (1 - \psi^*_{it}) r_{z,t+1}\big) \Big) \Big\},$$

which will only occur if $r^{ce}_{s,i,t+1}$ is sufficiently smaller than $r_{z,t+1}$. Thus, fixed transaction costs impose an entry barrier for current non-participants and an exit barrier for current participants. This makes it more likely that households will not change their participation status between period $t-1$ and period t, which induces first-order state dependence in the participation decision, i.e. current participation is positively correlated with past participation. Testing for positive state dependence in households' securities holdings could therefore provide a sufficient justification for the existence of pecuniary transaction costs (Vissing-Jørgensen, 2002a).

1.2.2 Habit Persistence

Pecuniary trading costs, however, are not the only possible explanation for state dependence in securities market participation decisions. Habit persistence would also yield behavior that was consistent with first-order state dependence in the participation decision.

Habit persistence might stem from household specific historical characteristics that adapt only slowly over time. If a household is used to put money aside in a savings book, it might continue in this pattern and refrain from investing in risky securities even if monetary costs do not impose any barrier. Reluctance to rebalance the portfolio could also result from differences in the elasticity of intertemporal substitution, which is closely linked to the degree of risk aversion through the parameter ρ in a utility function with constant relative risk aversion $U(C_{it}) = \frac{C_{it}^{1-\rho}-1}{1-\rho}$. Much evidence documents significant heterogeneity in this elasticity (Vissing-Jørgensen, 2002b). Empirical studies point out that households refrain from rebalancing their portfolios in each period. Guiso, Haliassos, Jappelli (2002) report that, conditional upon ownership, the age profile of asset shares is nearly flat in most countries. This indicates that people do not rebalance their wealth portfolio very frequently. Ameriks and Zeldes (2001) find that, in spite of different professional advice and in contrast to standard portfolio choice models, most households choose a particular portfolio of assets and stick to it even when their circumstances change. Similar observations have been reported by

Samuelson and Zeckhauser (1988).

1.2.3 Investment Experience

Another cause of state dependence in participation decisions could arise from increasing financial sophistication of investors. Calvet, Campbell, and Sodini (2007, 2009) show that wealthy and educated Swedish investors outperform the market by 4.3 percent and diversify more actively into risky assets. Similarly, Agnew, Balduzzi, and Sunden (2003) document that the wealthy and educated own a larger share of risky assets and rebalance their retirement accounts more actively. While wealth and education proxy the investor's degree of financial literacy, these measures exclude the possibility that ownership experiences could also foster investors' financial sophistication. To the best of my knowledge, no study has so far been able to account for the full ownership history of households.

The East German data allow me to distinguish two different channels through which ownership experiences could affect participation. First, I control for the number of years a household owned securities, which is a measure for asset-specific learning. Alessie, Hochguertel, van Soest (2004) argue that asset-specific learning may increase familiarity and awareness of the risk and return characteristics of the assets one owns. Also, Vissing-Jørgensen (2003) shows that experienced investors are less likely to overestimate stock returns and are more cautious with regard to inflation forecasts. This indicates that individual investment experience is likely to affect estimates of $r^{ce}_{s,i,t+1}$. If the accuracy of return expectations increases with the experience of the investor, experienced households will be less likely to leave the market due to false expectations. In line with this idea, Agarwal et al. (2008) estimate that the quality of financial decision-making is a concave function of age with a peak at an age of 53.

The second measure adds up the market returns for the years in which a household owned securities, which proxies the macro-experience with asset markets that each household made over its life-cycle. King and Leape (1987) point out that the presence of age and cohort patterns, which are typical for household portfolios, might be due to the fact that households accumulate information about investment opportunities over time. Malmendier and Nagel (2007) show that people who grew up during a time of bear markets are less likely to invest into risky assets over their life-cycle. For risk averse households, $r^{ce}_{s,i,t+1}$ is smaller than the expected net return of the risky portfolio $E\big(\psi_{it}r_{s,i,t+1} + (1-\psi_{it})r_{z,t+1}\big)$. Thus, risk averse households will underestimate the certainty equivalent return of the risky portfolio and stay out of the market for subsequent periods. Not controlling for such experience effects could thus induce state dependence in the estimates that does not stem from pure trading costs.

The essential implication of the above reasoning is that both monetary as well as non-monetary transaction costs would cause state dependence in ownership decisions. The theoretical considerations suggest the following strategy to quantify the importance of each type of costs. First, I explain participation in the securities market by estimating probit regressions with a binary dependent variable, which is unity if a households owns one or more securities in a given year and zero otherwise. Using a standard dynamic random effects model, I then test for state dependence in the participation decision. Second, I compare this estimate with an estimate from a regression that also controls for habit persistence (or investment experience). If the new estimate is significantly smaller than before, this would suggest that state dependence is rather caused by habit persistence (or investment experience) than by pecuniary transaction costs. Finally, I compare both estimates to a regression that controls for habit persistence as well as investment experience.

1.3 Identification Strategy

For identification, I rely crucially on three features of the institutional environment in East Germany. First, I reason that each household's individual ownership experience can be observed. Second, I argue that the unobserved effects are uncorrelated with initial securities ownership. Third, I assume that the explanatory variables are weakly exogenous to the participation decision, allowing a causal interpretation of the estimates.

A unique feature of East German portfolios is that the full ownership history of each household can be tracked down. In 1990, households of all age groups without any prior experience were confronted with securities for the first time. At the macro level, age and cohort patterns should therefore not reflect any experience effects, because everybody faced the same market conditions during his life under a capitalist system. Trends that are common to all investors will be captured via year dummies. At the individual level, however, there might be strong effects from macro-experiences, depending on whether investments were made during a good or a bad year on the market.

Second, I do not face the initial conditions problem common to dynamic binary choice models, as securities ownership was not possible before reunification. East Germans could only put their money aside in savings accounts, savings books and life insurance with a unitary interest rate of 3.25% (Dabbert, 1992; Schwarzer, 1999). Private ownership of companies was barred. Although transitional ownership was not forbidden until 1972, private ownership during this period typically included the state as shareholder. Similarly, bond markets did not develop in the GDR due to limited issuing as well as a unitary interest rate. Foreign investments were blocked because borders between the GDR and the West were closed from 1961 onward. Hence, a new data generating process started with German reunification in 1990.

Finally, most lifetime decisions were made under the communist system, so that post-reunification incomes are sufficiently exogenous to the participation decision. In the GDR, incomes of university graduates were only 15 percent higher than those of blue-collar workers, compared to 70 percent in West Germany. Sectoral wage differences were also very low, and full employment was constitutionally guaranteed (Fuchs-Schündeln and Schündeln, 2005). Additionally, educational choices were restricted by state intervention. Only a certain quota of students was permitted to attend the last two years of high school, which were necessary to enter university. Access to higher education required membership in the GDR youth organization (*FDJ*), and political opinions in accordance with official party positions. Children from working class families had preferential access. Based on these circumstances, I draw two conclusions. First, the financial rewards for different qualifications in the reunified country were not taken into account when career decision were made. Second, self-selection into careers with low risks concerning income fluctuations is unlikely, given the small income differences and the absence of unemployment risks. Thus, educational and occupational characteristics of households are not likely to be correlated with their preferences for risky assets.

1.4 Empirical Models

In this section, I introduce the benchmark dynamic discrete-choice model for panel data with time-constant random effects, which has been widely used in the literature, e.g. Miniaci and Weber (2002). I then proceed to introduce time-varying heterogeneity in this model, which is theoretically more convincing, because it allows to disentangle true state dependence from habit persistence in unobserved preferences. The empirical part in section 1.6 will also show that, empirically, persistence in unobserved effects is the preferred assumption.

1.4.1 Dynamic Random Effects Probit

Let S_{it}^* denote a latent variable which represents the desired level of risky securities of household $i = 1, \ldots, N$ at time $t = 1, \ldots, T$. Even if the desired level of risky assets is not known, the probability of securities market participation $Pr(S_{it} = 1)$ can be estimated, where

$$S_{it} = \begin{cases} 1 & \text{if } S_{it}^* > 0 \\ 0 & \text{if } S_{it}^* \leq 0. \end{cases} \tag{1.4}$$

Assuming that participation is independent over time, the joint probability is given by $Pr(S_{i1}, \ldots, S_{iT}) = \prod_{t=1}^{T} Pr(S_{it})$. Under these assumptions, participation probabilities could be studied using standard cross-section discrete choice models. However, the independence assumption fails in presence of state dependence, which implies that the probability of participation is higher for previous participants than non-participants, i.e. $Pr(S_{it} = 1|S_{i,t-1} = 1) > Pr(S_{it} = 1|S_{i,t-1} = 0) \neq Pr(S_{it})$. As the presence of pecuniary transaction costs would suggest that past ownership modifies behavior today, a dynamic specification is needed:

$$S_{it}^* = \gamma S_{i,t-1} + \mathbf{x_{it}}\beta + c_i + \epsilon_{it}, \tag{1.5}$$

where c_i is an unobserved time-invariant individual effect and ϵ_{it} an i.i.d. error term. $\mathbf{x_{it}}$ is a vector of strictly exogenous variables, $\mathbf{x_i} = (x_{i1}, \ldots, x_{iT})$. In order to test for true state dependence, it is important to control for the individual effect c_i, because even if the true $\gamma = 0$, $Pr(S_{it} = 1|S_{i,t-1}, \mathbf{x_i}) \neq Pr(S_{it} = 1|\mathbf{x_i})$ due to the presence of c_i.

In short panels, there is also the well-known initial conditions problem due to the need to integrate out c_i in order to approximate the joint distribution $f(\mathbf{S_i}|\mathbf{x_i}, c_i)$, where $\mathbf{S_i} = (S_{i1}, \ldots, S_{iT})$. Several approaches are available to deal with the initial conditions, e.g. Heckman (1981), Wooldridge (2005), Honoré and Tamer (2006). However, in this application, a new process starts with the first sampling period in 1990. This will greatly facilitate modeling the autocorrelation structure of the errors compared to, for instance, Hyslop (1999), who estimates an AR(1)-RE model in the context of labor market participation. Hence, the distribution of c_i is independent of the initial value S_{i0} (which is zero for all individuals),

$$g(c_i|\mathbf{x_i}, S_{i0}) = g(c_i|\mathbf{x_i}). \tag{1.6}$$

I relax the assumption that $(S_{i1}, \ldots, S_{iT})$ are independent conditional on $(\mathbf{x_i}, c_i)$ using a Mundlak (1978) specification of the random effect

$$c_i|\mathbf{x_i} \sim N(\delta_0 + \overline{\mathbf{x}}_\mathbf{i}\delta_\mathbf{1}, \sigma^2), \tag{1.7}$$

where σ^2 is the conditional variance of c_i. While maintaining the strict exogeneity assumption on $\mathbf{x_{it}}$ conditional on c_i, this specification allows for correlation between c_i and the average $\mathbf{x_i}$. Two features of the Mundlak specification should be noted. First, all time dummies must be dropped from (1.7) and a constant may only appear in (1.7) to avoid collinearity. Second, the coefficients of time-invariant regressors are a composite of the relevant elements of $(\beta + \delta_\mathbf{1})$ and cannot be separately identified. [3]

1.4.2 Dynamic Random Effects Probit with an AR(1) Error Component

To see why state dependence might be caused by persistence in unobserved effects, consider the same model as before, yet with time-varying (c_{it}) instead of time-constant (c_i) individual heterogeneity

$$S^*_{it} = \gamma S_{i,t-1} + \mathbf{x_{it}}\beta + c_{it} + \epsilon_{it}. \tag{1.8}$$

Assume the AR(1) structure for the unobserved characteristics

$$c_{it} = \rho c_{i,t-1} + a_{it}, \tag{1.9}$$

where $a_{it}|\mathbf{x_i} \sim N(\delta_0 + \overline{\mathbf{x}}_\mathbf{i}\delta_\mathbf{1}, (1-\rho^2)\sigma^2)$, following the Mundlak specification above. Due to the discrete and continuous nature of $S_{i,t-1}$ and $c_{i,t-1}$ respectively, this specification of the correlation structure across time assures that γ and ρ are identified. Note that this model corresponds to a random effects model with $c_i = c_{it}$ for all t in the case of $\rho = 1$, and a standard probit model if $\sigma = 0$. The correlation between c_{it} and c_{is} conditional on $\mathbf{x_i}$ is $\rho^{|t-s|}$. In presence of (decaying) habit persistence, it follows that $Pr(S_{it} = 1|S_{i,t-1}, \mathbf{x_i}, c_i) \neq Pr(S_{i,t}|\mathbf{x_i}, c_i)$ as long as $0 < \rho < 1$, even if the true $\gamma = 0$. Thus a probit model with time-constant random effects does not allow studying true

[3]Chamberlain (1980) allowed for a more flexible specification by using the vector of all explanatory variables across all time periods $\mathbf{x_i}$ instead of $\overline{\mathbf{x}}_\mathbf{i}$. I prefer the Mundlak (1978) specification, as it conserves on parameters. However, the model does not specify a complete model for the unobserved effects and may therefore be sensitive to misspecification. Both approaches are identical if $\delta_\mathbf{11} = \cdots = \delta_\mathbf{1T}$.

state dependence if persistence in individual (unobserved) effects can be suspected.

Integrating out the full latent process $c_{i1}, \ldots, c_{iT}$, the likelihood contribution of the observations on cross-sectional unit $i = 1, \ldots, N$ for the model in (1.8) and (1.9) can be written as

$$\begin{aligned} L_i = Pr(\mathbf{S_i}|\mathbf{x_i}, S_{i,t-1}) &= Pr(S_{i1}|\mathbf{x_i})Pr(S_{i2}|\mathbf{x_i}, S_{i1}) \ldots Pr(S_{iT}|\mathbf{x_i}, S_{i1}, \ldots, S_{i,T-1}) \quad (1.10) \\ &= \int \cdots \int \Big(\prod_{t=1}^{T} Pr(S_{it}|\mathbf{x_i}, S_{i,t-1}, c_{it}) \Big) g(c_{i1}, \ldots, c_{it}|\mathbf{x_i}) dc_{i1} \ldots c_{iT}, \end{aligned}$$

since $S_{i0} = 0$. Although this T-dimensional integral does in general not have an analytic solution, it can be approximated numerically using simulation techniques, which yield a consistent estimator if the number of replications R rises sufficiently with N (Hajivassiliou and Ruud, 1994). While this holds asymptotically, Lee (1997) shows that the accuracy given a finite number of replications often worsens with the dimension of the integral. Given the high time-series dimension of the data here ($T = 17$), joint simulation over the T-dimensional integral cannot be expected to work well. Therefore, I apply Heiss' (2008) random effects estimator, which reduces the dimensionality via a sequential nonlinear Kalman-filter. Based on (1.10), the outcome probabilities conditional on past values are approximated at each t for the one-dimensional integral

$$Pr^*_{it} \approx Pr(S_{it}|\mathbf{x_i}, S_{i1}, \ldots, S_{i,T-1}) = \int Pr(S_{it}|\mathbf{x_i}, S_{i,t-1}, c_{it}) g(c_{it}|\mathbf{x_i}, S_{i1}, \ldots, S_{i,T-1}) dc_{it}. \quad (1.11)$$

Sequentially approximating (1.11) for all $t = 1, \ldots, T$ and updating of the conditional probability distribution function $g^*(c_{i,t+1}|\mathbf{x_i}, S_{i1}, \ldots, S_{i,T-1})$ now allows to approximate the joint likelihood contribution $Pr(\mathbf{S_i}|\mathbf{x_i}, S_{i,t-1}) = Pr^*_{i1} Pr^*_{i2} \ldots Pr^*_{iT}$. The outcome probability in the first time period, $Pr(S_{i1}|\mathbf{x_i})$ is estimated unconditionally of past values of S_{it}. This approximation is consistent, because the initial conditions can be neglected here. For details of the algorithm, see Heiss (2007). Gauss-Hermite quadrature with 30 nodes is used for simulations. Sensitivity of the results to the choice of nodes is checked for both models.

In order to estimate the average partial effects of this non-linear estimator, it is necessary to calculate the average partial effect over the distribution of individual heterogeneity (Wooldridge, 2002). For given values of $\mathbf{x_t} = \mathbf{x^0}$ and $S_{t-1} = S^0_{-1}$, the average partial effects for the model with time-constant heterogeneity, $E[E[\Phi(\gamma S^0_{-1} +$

$\mathbf{x^0}\beta + c_i)|\mathbf{x_i}]]$, can be estimated by

$$\frac{1}{N}\sum_{i=1}^{N}\Phi\Big((1+\hat{\sigma^2})^{-\frac{1}{2}}(\hat{\gamma}S^0_{-1}+\mathbf{x^0}\hat{\beta}+\hat{\delta_0}+\overline{\mathbf{x}}_{\mathbf{i}}\hat{\delta_{\mathbf{0}}})\Big). \tag{1.12}$$

Using iterated expectations with respect to the distribution of c_{it} for the model with time-varying heterogeneity $E[E[\Phi(\gamma S^0_{-1}+\mathbf{x^0}\beta+c_{it})|\mathbf{x_i}, c_{i,t-1}]]$, the average partial effect can be consistently estimated by

$$\frac{1}{N}\sum_{i=1}^{N}\Phi\Big((1+(1-\hat{\rho}^2)\hat{\sigma^2})^{-\frac{1}{2}}(\hat{\gamma}S^0_{-1}+\mathbf{x^0}\hat{\beta}+\hat{\delta_0}+\overline{\mathbf{x}}_{\mathbf{i}}\hat{\delta_{\mathbf{0}}})\Big). \tag{1.13}$$

1.5 The Data

The data come from the German Socioeconomic Panel (GSOEP), which is a longitudinal survey of private households in Germany.[4] The subsample covering the territory of the former GDR started in 1990 and contains 1988 initial responses. As I am interested in dynamic behavior, I only use those observations for which data is available for all 17 years. Due to attrition, this leaves me with 1078 East German households that are observed from 1990 to 2007. At the end of section 1.6, I report results of tests for systematic attrition, but cannot find any evidence for selection bias. If questions are not at the household level, I use information about the household head (e.g. age, education).[5]

Each year households are asked about the assets in their savings portfolio. The question reads: "Did you or any other person in your household possess one or more of the following investments last year? a) savings book, b) savings agreement with building society, c) life insurance, d) securities [stocks, bonds, mortgage bonds,...], e) business capital, f) no, none of these investments." These variables are one if households hold a certain asset, and zero otherwise. Two things are noteworthy about response category d) of this question. First, the question gives information about security holdings in general, while detailed information about bond and stock holdings cannot be inferred. Similar to Guiso, Japelli, Terlizzese (1996), I therefore define security holdings as a

[4]The Add-On package PanelWhiz for Stata (http://www.PanelWhiz.eu) has been used for extracting the data. See Haisken-DeNew and Hahn (2006) for details. The PanelWhiz generated DO file to retrieve the data used here is available from me upon request. Any data or computational errors are my own.

[5]The GSOEP is the only panel data set that allows to study East German households over time. The Bundesbank Income and Consumer Survey (EVS) is only conducted every five years for a cross-section of households. For a comparison of both surveys, see Börsch-Supan and Eymann (2002).

broad set of risky assets, compared to save investments such as savings books, building society contracts, or life insurance. Only from 2001 onward, the questionnaire asks for bond and stock holdings separately. Second, the information is reported with a one year lag. So the 1990 survey asked for securities holdings in 1989. Although individuals could not legally own securities in the GDR, 13 households actually reported to do so in 1989. However, these positive responses are likely to be erroneous, since only two of these households indicated to hold securities a year later (which might also be newly bought securities). I therefore set all observations before 1990 to zero.

Starting with zero security holdings in 1990, figure 1.1 depicts that ownership rates adjusted very quickly to those in West Germany. From 1993 onward, ownership rates in East and West Germany move parallel, although they are lower in the East probably due to higher unemployment risk and lower average incomes. Similar to other industrialized countries ownership increased during the 1990s. In the East German case, the main explanatory factors for this increase might be portfolio adjustments after reunification, increases in real income, and the three waves of privatization of Deutsche Telekom (1996, 1999, 2000), which boosted stock market participation among average households. After the burst of the dotcom bubble, ownership rates did slightly decay to 31% in 2006. In addition, figure 1.1 depicts that stocks made up about 2/3 of all securities investments since 2000. This indicates that those households reporting to hold securities indeed had much riskier portfolios than those relying on savings books.[6]

The GSOEP does not provide information with regard to portfolio composition. The only available time series data that give an indication of portfolio shares across different assets are from the Bundesbank (2008), which reports the allocation of private wealth in Germany. Table 1.1 reports the total wealth invested into different asset classes between 1991 and 2006 in Germany. There is a clear trend that less and less wealth is held in savings books and savings accounts, although 33.2% of private German wealth was still held on savings accounts in 2006. Particularly, investments in securities increased strongly from 29.2% in 1991 to 35.0% in 2006, but also life insurance and pension funds became more and more important during the 1990s. Among securities, the share of wealth invested in savings certificates fell from 17.0% in 1991 to 4.7% in 2006. Similarly, bonds became less attractive. Stock ownership peaked around 2000, whereas mutual funds account for the largest share among securities in 2006 (33.3%).

Each households' accumulated ownership experience is calculated at each point in time as the one-period lagged sum of years during which the household owned securi-

[6]German banking laws require mandatory insurance of deposits. For *Sparkassen* (public savings banks), which administer a large share of private savings, deposit insurance is unlimited. But also most private banks take part in voluntary guarantee schemes, which implicates for the average investor that all of his deposits are secure.

ties, i.e. $exp_{it} = \sum_{\tau=1}^{t} S_{i,\tau-1}$. Using the lagged sum is particularly convenient, because it ensures that exp_{it} is uncorrelated with the unobserved preference for holding securities, given that lagged ownership and persistence in unobserved preferences are control for. Table 1.1 presents the total ownership experience that the households in the sample gained during the period from 1990 to 2006. Although 14 households held securities in every year since reunification, the majority of households gained no more than two years of experience. 32.6% of all households did not participate at all, which is in stark contrast to standard lifetime portfolio choice models.

The second proxy for investment experience builds on the work by Malmendier and Nagel (2007), who calculate the average return history of the stock market during each investor's lifetime. Yet age-related returns differentials should be insignificant in East Germany, since all households accumulated the same macro-experience after reunification. Rather, returns should matter for those years in which a household owned securities. The macro-return experience of household i is therefore defined as

$$rexp_{it} = \begin{cases} \exp(\frac{1}{exp_{it}} \sum_{\tau=1}^{t} \ln(1 + S_{i,\tau-1} \sum_{k=1}^{K} w_{k,\tau-1} R_{k,\tau-1}) - 1) & \text{if } exp_{it} > 0 \\ 0 & \text{if } exp_{it} = 0. \end{cases}$$

w_k is the share of asset $k = \{Stocks, Bonds\}$ in total annual savings, as calculated from table 1.1. Average annual returns on stocks and bonds, R_k, are taken from the Bundesbank, which provides data on the performance of the DAX stock market index as well as average interest rates on savings deposits in Germany.

To get an impression of the average ownership duration, figure 1.2 depicts the average share of new entrants in a given year who invest into securities for at least the three following years. As expected, investment duration is longer if households enter before markets surge. However, households who entered at the peak of the DAX index in 2000 have similar investment durations as those entering during the previous years. One reason could be that many unexperienced investors, which entered in 2000, hoped they could recover their losses due to the dotcom crash. The share of long-term investors has declined since then.

In figure 1.3, market returns are combined with individual investment durations. Although this does not provide information on individual returns, it gives an indication of the average returns that households experienced. For each year, the figure depicts the average market return that market entrants experienced, given the duration of their investment. For instance, a household that entered in 1999 experienced market returns of around -10 percent, given the average investment duration. Thus, the figure rejects the idea that households benefited from the equity premium, because investors that entered during bullish market years did not have the stamina to invest for long enough as to materialize potential long-run gains.

Table 1.2 presents annual transition dynamics and table 1.3 transitions between 1995 and 2005. This gives a partial view of changes in securities holdings, as only transitions of households that sell all their securities or enter the securities market are shown. Table 1.2 reads as follows. Of those households not holding securities last period, 89.5% do also not hold securities this period, whereas 10.5% bought securities. Persistence among those who owned securities last period is very high (77.2%). Yet the exit rate (22.8%) is higher than the entry rate (10.5%). Similarly, table 1.3 reports that 67.6% of those owning securities in 1995 also participated in 2005. Both tables show that households change their investment choices relatively infrequently. It is unlikely that such long-term persistence can be explained with state dependence from year to year only. Rather, there might be persistence in unobserved household preferences over time.

Figure 1.4 depicts the annual share of households that enter and those that stay out of the market. Households that enter had no securities last period but hold securities this period. The entry rate is around 10% during the entire period, except for a jump in 2000. Similarly, the share of those staying out is relatively high, which indicates reluctance to change the status quo (Ameriks and Zeldes, 2001). On the other hand, a declining exit rate is the main reason for the observed increase in ownership rates during the 1990s (figure 1). This also indicates strong habit persistence among households, because state dependence alone would not lead to declining exit rates.

Figure 1.5 reports (head of household) age and cohort patterns for securities ownership rates. I use 5-year intervals to define age-in-1990 cohorts with an initial age between 18 and 25 in 1990 for the youngest cohort, and those older than 75 for the oldest cohort. The figure gives raw ownership rates for the unbalanced panel of all cohorts in all waves. Each cohort curve consists of 17 points that indicate the average age and ownership rate at the time of the interview. While all cohorts have zero holdings in 1989, ownership rates in 1990 start already at different levels. These jumps between cohort curves indicate that there are age and/or cohort effects. Specifically, cohorts aged 55 and older at reunification have a very distinct ownership pattern compared to younger cohorts. Ownership rates seem to oscillate around a flat average. This indicates presence of cohort and time effects, but potential absence of age effects for the older cohorts. Yet there are age and/or time effects among the younger cohorts, as their cohort curves are not horizontal. However, age, cohort, and time effects cannot be identified without further assumptions.

Summary statistics of additional covariates are reported in table 1.4. Based on the annual income distributions, I partition household incomes into five quintiles (taking the poorest quintile as the reference category). In theory, non-financial income could affect securities ownership positively as well as negatively, depending on the correlation between securities returns and non-financial income. Specifically, perfect correlation between both would allow to perfectly hedge non-financial income streams. In practice, however, German households do not invest a significant amount of wealth into derivatives (see table 1.1), so that a positive effect can be suspected. With regard to financial income, I construct a measure of financial wealth based on the question: "How high was the income from interest and dividends in the last year, all in all? (If not known exactly, please estimate the amount using this list: a) less than 500 DM, b) 500 to 2000 DM, c) 2000 to 5000 DM, d) 5000 to 10000 DM, e) 10000 DM and over.)" Due to potential endogeneity, I use lagged financial wealth and include this covariate only for sensitivity checks. Additional control variables for the household head include age, martial status, labor force status, education, dummies for the number of children (one, two, or more than two), and an indicator for house ownership.

1.6 Estimation Results

Before moving on to the estimation results, it is worth remembering that the high persistence in participation has so far almost exclusively been modeled as either time-constant unobserved heterogeneity and/or state dependence. Table 1.6 gives an impression of the intertemporal correlation pattern over several years. The ownership decision is modeled as a function of socio-economic characteristics of the household as well as lagged participation outcomes. Most strikingly, the coefficients of all lags are significantly different from zero and get smaller the higher the order of the lag. This observation contrasts to the assumptions of the models that have been used so far. First, a random effects model would imply equal predictive power of all lags, which cannot be confirmed by the estimates. Second, a model that exhibits state dependence in the participation decision would imply that, after controlling for the first lag, all other lags have no significant predictive power. This is in contradiction to the estimates. Finally, the combination of state dependence with a (time-constant) random effects framework would imply higher predictive power of the first lag and equal predictive power of the other lags. However, a Wald test strongly rejects the hypothesis that lags two to five have equal predictive power with a $\chi 2$-statistic of 48.42 (0.000).

These results indicate that the models that are commonly used to study ownership decisions cannot sufficiently capture the correlation structure in the data. As long as habit persistence and investment experience are not controlled for, the lagged dependent variable picks up different sources of intertemporal correlation in the standard random effects dynamic probit model. In order to see how large this bias is, it will be helpful to compare estimates from a random effects model with time-constant heterogeneity to those from a random effects model that explicitly models habit persistence and/or controls for experience effects.

The estimates reported in table 1.7 show that the random effects specifications with time-varying heterogeneity are clearly preferred to their time-constant alternatives. The estimated ρ in columns (3) and (5) is clearly different from 1.[7] Also, a likelihood ratio test on the full vector of parameters rejects the hypothesis that model (2) is preferred to model (3). Similarly, model (4) is rejected against (5). Moreover, the Akaike and Bayesian information criteria favor those specifications that allow for autocorrelation in the error term. According to both criteria, the preferred specification is (5), which allows for state dependence, habit persistence and controls for investment experiences. As $\overline{\mathbf{x}}_\mathbf{i}$ is jointly significant, a Mundlak specification is needed to account for correlation between random effects and covariates in oder to get unbiased estimates of β. In all specifications, the lagged dependent variable is strongly significant, substantiating the view that monetary transaction costs cause true state dependence in participation decisions. However, the estimated degree of state dependence varies considerably between different specifications.

Table 1.8 reports average partial effects of the degree of state dependence. I estimate that last period participants have a 28.5 percentage points higher probability of participating in the current period, than those that did not participate last period. This estimate is of similar size as those of other (time-constant) random effects specifications, e.g. Vissing-Jørgensen (2002a) or Alessie, Hochguertel, van Soest (2004). Yet once I allow for habit persistence, the estimated degree of state dependence falls to only 8 percentage points in model (3) (5.2 percentage points if experience is controlled for too). So the estimated degree of state dependence is only three quarters as large as in the case of time-constant heterogeneity. The reason for this huge difference is that the correlation of the error term appears to decay over time (table 1.6). But specifications (2) and (4) constrain the random effect to be constant, so that the lag of the dependent variable takes up all autocorrelation. Once this restrictive assumption on the random effect is relaxed, the coefficient of the lagged dependent variable only captures the pecuniary transaction costs effect. The degree of habit persistence in model (5) is $\hat{\rho} = 0.86$, which corresponds to a persistence of $0.86^{17} = 8.5\%$ of the initial preferences till 2006.

[7] The necessary test is to compare $\arctan(\rho) = 15$ against the alternative that $\arctan(\rho) \neq 15$.

In contrast to Agarwal et al. (2008), Malmendier and Nagel (2007), and Ameriks and Zeldes (2001), who take age and cohort effects as an indication for the importance of investment experience in portfolio choices, I find little evidence for pronounced age and cohort patterns. This is in accordance with the fact that all households started off with zero investment experience in 1990. However, the accumulated ownership experience has a strong impact on subsequent investment decisions. Table 1.8 reports for specification (5) that one year of ownership experience increases the probability to invest into securities by 4 percentage points on average. This strong impact corroborates the idea that asset-specific learning affects households' savings decisions. However, macro-experiences, as captured by $rexp_{it}$, are insignificant in all specifications. This calls into doubt that the market returns that individuals experience over their life-cycle affect their risk taking behavior.

Table 1.8 also reports average partial effects of non-financial income on participation. In specification (5), households in the top income quintile have a 4.0 percentage points higher probability of staying in the market than those in the bottom quintile, supporting the view that entry costs are less important for richer households (Vissing-Jørgensen, 2003).

In table 1.9, I also allow for nonlinear terms of exp_{it} and $rexp_{it}$ and test in how far financial wealth affects the participation decision. Models (6) and (7) indicate the presence of diminishing returns to ownership experience. $rexp_{it}$ is still insignificant. The wealth dummies are neither individually nor jointly significant, but their averages are. These findings are in line with Brunnermeier and Nagel (2008), who show that changes in wealth do not affect the risk taking behavior of households in the PSID. While my estimates also reject that changes in wealth affect participation, they indicate that the degree of risk aversion is related to lifetime financial wealth.

I also check the sensitivity of the estimates to expectations about financial and employment risks. In each year, household heads are asked if they are severely worried about their own financial situation. In 27% of the responses there was indication of financial worries, which might also be a sign of borrowing constraints among these households. Similarly, 16% of the respondents indicated that they were severely worried about the security of their jobs. However, both variables do not significantly explain participation, which gives me confidence that my specification yields unbiased estimates.

Finally, I test for potential attrition bias by applying an inverse probability weighted (IPW) pooled probit estimator (Wooldridge, 2002).[8] To implement this estimator I di-

[8]An IPW estimator cannot be applied to random effects models. However, the partial maximum likelihood properties of the pooled probit estimator imply that differences to random effects estimates

vide all initial respondents into those that participated in each survey until 2007 (the balanced sample) and those that dropped out earlier. Given the initial characteristics of the household and its participation status in 1990, I estimate a probit model for being in the balanced sample and compute inverse probability weights for each observation. This regression includes an additional dummy indicating politically interested households. As required in a selection on observables approach, this indicator is significantly correlated with the probability to participate in the survey repeatedly as well as the probability to own securities (Fitzgerald, 1998). Unweighted estimates from a pooled probit model for the participation decision can then be compared to estimates from an IPW probit, as reported in table 1.10. The estimates show that attrition does not impose a severe problem. Particularly, the coefficients of the lagged dependent variable as well as the experience dummies show little difference between unweighted and IPW regressions.

1.7 Conclusion

The widespread non-participation in securities markets is one of the most surprising puzzles of standard portfolio theory. I analyze portfolio decisions of East German households after German reunification. A descriptive analysis documents that (i) only few households in the sample own risky assets between 1990 and 2006, (ii) most households change their portfolios infrequently, and (iii) increases in ownership rates during the 1990s are mainly due to fewer exits, while the percentage of new entries remains relatively constant over the entire period. These stylized facts indicate that household behavior exhibits a high degree of persistence that cannot merely be explained by state dependence from year to year.

Compared to previous attempts to study dynamic participation behavior in securities markets, the peculiar institutional environment in East Germany offers several advantages. First, I can estimate the effect of investors' financial sophistication from their newly gained investment experience on participation dynamics, since households of all age groups had no prior experience with securities in 1990. This allows me to disentangle habit persistence and investment experience from pecuniary transaction costs motives in participation decisions. Second, the initial conditions of this dynamic process are truly exogenous thanks to the fact that markets for securities did not exist in the former GDR.

I find that habit persistence and lack of investment experience are the main reasons

are, asymptotically, negligible.

for the observed non-participation, whereas the importance of pecuniary transaction costs has been largely overestimated in past studies. In other words, East Germans' tastes for risky assets developed only slowly over time. Contrary to Malmendier and Nagel (2007), Agarwal et al. (2008), and King and Leape (1987), who take age and cohort patterns as an indication for the importance of life-cycle effects in risk preferences, there is little evidence that past market returns affect investment decisions of East German households (for which the full ownership history is known). Similarly, changes in financial wealth do not affect risk taking behavior, which is in contrast to theory (Brunnermeier and Nagel, 2008), whereas average wealth levels impact investment choices.

However, the data do not allow an analysis of the welfare consequences of non-participation, as I only observe participation choices. Based on Swedish data, Calvet, Campbell, and Sodini (2007) estimate that poor and poorly educated households would potentially invest in underperforming assets. This could explain why they prefer to abstain from stock markets altogether. In contrast to Sweden, where 62 percent of households own stocks, only 20.4 percent do so in East Germany. It is thus unlikely that education and income alone could explain such rampant non-participation sixteen years after reunification. Additionally, the peculiar environment in East Germany provides an ideal natural experiment to study participation behavior. Yet the findings have a wider applicability, since inertia in investment decisions have been documented for several countries. Also, participation in the East is similar to West Germany, albeit at a lower level due to inexperience with securities, different preferences, lower incomes, and higher unemployment risks in the East.

The findings are essential for understanding portfolio choices of households in general. It has been argued that the non-participation puzzle is closely related to the equity premium puzzle, since the correlation between consumption and stock returns is higher for stockholders than for non-stockholders (Vissing-Jørgensen, 2002b). Yet the theoretical literature has so far mainly focused on habit persistence in consumption patterns, e.g. Constantinides (1990), Møller (2009). The incorporation of habit formation in households' preferences for risky assets may help to explain the equity premium in life-cycle portfolio choice models.

My results are particularly relevant for policy-makers and financial consultants seeking to establish household portfolios as an additional pillar of the social security system. In order to increase the acceptance of private savings plans, they will have to override households' habit persistence and reservation with regard to risky assets. Once they have overcome these inhibitions, households are likely to take advantage of the full range of available assets to put aside their savings. Hence, my results strengthen the case for policies such as default options, which do not require people to opt in but rather allow opting out of private retirement savings schemes. This conclusion differs from previous studies in so far as they argued that non-participation is merely a rational response to high entry costs. However, the East German experience demonstrates that it is primarily the wall in people's minds to adopt to new circumstances, which explains non-participation behavior. This is what is truly puzzling about the non-participation puzzle.

References

Agarwal, Sumit; John C. Driscoll; Xavier Gabaix and David Laibson (2008): "The Age of Reason: Financial Decisions Over the Lifecycle," mimeo, NYU Stern School of Business.

Alesina, Alberto and Nicola Fuchs-Schündeln (2008): "Good Bye Lenin (Or Not?): The Effect of Communism on People's Preferences," *American Economic Review*, **97**, 1507-1528.

Alessie, Rob; S. Hochguertel and A. van Soest (2004): "Ownership of Stocks and Mutual Funds: A Panel Data Analysis," *The Review of Economics and Statistics*, **86**, 783-796.

Alessie, Rob and Stefan Hochguertel (2002): "Household Portfolios in the Netherlands," in *Household Portfolios*, eds. Luigi Guiso, Michael Haliassos, and Tullio Jappelli. MIT Press: Cambridge, 341-388.

Ameriks, John and S. Zeldes (2001): "How Do Household Portfolio Shares Vary with Age?," TIAA-CREF Institute Working Paper No. 6-120101.

Agnew, Julie; Pierluigi Balduzzi and Annika Sunden (2003): "Portfolio Choice and Trading in a Large 401(k) Plan," *American Economic Review*, **93**, 193-215.

Bergstresser, Daniel and James Poterba (2004): "Asset Allocation and Asset Location: Household Evidence from the Survey of Consumer Finances," *Journal of Public Economics*, **88**, 1893-1916.

Börsch-Supan, Axel and Angelika Eymann (2002): "Household Portfolios in Germany," in *Household Portfolios*, eds. Luigi Guiso, Michael Haliassos, and Tullio Jappelli. MIT Press: Cambridge, 291-340.

Brunnermeier, Markus K. and Stefan Nagel (2008): "Do Wealth Fluctuations Generate Time-Varying Risk Aversion? Micro-Evidence on Individuals Asset Allocation," *American Economic Review*, **98**, 713-736.

Calvet, Laurent E.; John Y. Campbell and Paolo Sodini (2009): "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes," *Journal of Political Economy*, **115**, 707-747.

Calvet, Laurent E.; John Y. Campbell and Paolo Sodini (2009): "Fight Or Flight? Portfolio Rebalancing by Individual Investors," *Quarterly Journal of*

Economics, **124**, 301-348.

Chamberlain, Gary (1980): "Analysis of Covariance with Qualitative Data," *Review of Economic Studies*, **47**, 225-238.

Choi, James J.; David Laibso and Brigitte C. Madrian (2006): "The Importance of Default Options for Retirement Saving Outcomes; Evidence from the United States," in *Social Security Policy in a Changing Environment*, eds. Jeffrey Brown, Jeffrey Liebman and David A. Wise. University of Chicago Press, Chicago.

Constantinides, George M. (1990): "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy*, **98**, 519-543.

Dabbert, Edith (1992): "Von 280 Mark zu 11830 Mark - 40 Jahre Sparen," in *Im Trabi durch die Zeit - 40 Jahre Leben in der DDR*, ed. Egon Hölder. Metzler-Poeschel: Stuttgart, 121-127.

Deutsche Bundesbank (2008): *Ergebnisse der gesamtwirtschaftlichen Finanzierungsrechnung für Deutschland 1991 bis 2007, Statistische Sonderveröffentlichung 4.* Deutsche Bundesbank: Frankfurt.

Fitzgerald, John; Peter Gottschalk and Robert Moffitt (1998): "An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics," *Journal of Human Resources*, **33**, 251-299.

Fuchs-Schündeln, Nicola (2008): "The Response of Household Saving to the Large Shock of German Reunification," *American Economic Review*, **98**, 1798-1828.

Fuchs-Schündeln, Nicola and Matthias Schündeln (2005): "Precautionary Savings and Self-Selection: Evidence from the German Reunification Experiment," *Quarterly Journal of Economics*, **120**, 1085-1120.

Guiso, Luigi; Tullio Japelli and Daniele Terlizzese (1996): "Income Risk, Borrowing Constraints and Portfolio Choice," *American Economic Review*, **86**, 158-172.

Guiso, Luigi; Michael Haliassos and Tullio Jappelli (2002): *Household Portfolios.* MIT Press: Cambridge.

Guiso, Luigi; Michael Haliassos and Tullio Jappelli (2003): *Stockholding in Europe.* Palgrave: Houndmills.

Guiso, Luigi and Tullio Jappelli (2005): "Awareness and Stock Market Participation," *Review of Finance*, **9**, 537-567.

Guiso, Luigi; Paola Sapienza and Luigi Zingales (2005): "Trusting the Stock Market," *The Journal of Finance*, **63**, 2557-2600.

Haisken-DeNew, John P. and Markus Hahn (2006): "PanelWhiz: A Flexible Modularized Stata Interface for Accessing Large Scale Panel Data Sets," mimeo, (http:// www.PanelWhiz.eu).

Hajivassiliou, V.A. and P.A Ruud (1994): "Classical Estimation Methods for LDV Models using Simulation," in *eds. Robert Engle and Daniel McFadden*, Handbook of Econometrics, Vol. 4. Elsevier, 2383-2441.

Haliassos, Michael and Bertaut, Carol C. (1995): "Why do so few hold Stocks?," *Economic Journal*, **105**, 1110-1129.

Haliassos, Michael and Michaelides,Alexander (2003): "Portfolio Choice and Liquidity Constraints," *International Economic Review*, **44**, 143-178.

Haliassos, Michael (2008): "Household Portfolios," in *The New Palgrave Dictionary of Economics, Second Edition*, eds. Steven N. Durlauf and Lawrence E. Blume. Palgrave Macmillan, .

Heckman, James (1981): "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Dime - Discrete Data Stochastic Process," in *Structural Analysis of Discrete Data and Econometric Applications*, eds. Charles Manski and Daniel McFadden. MIT Press: Cambridge, 179-195.

Heiss, Florian (2007): "Limited Dependent Panel Data Models with AR(1) Error Components in Stata," mimeo, University of Munich.

Heiss, Florian (2008): "Sequential Numerical Integration in Nonlinear State Space Models for Microeconometric Panel Data," *Journal of Applied Econometrics*, **23**, 373-389.

Hong, Harrison; Jeffrey D. Kubik and Jeremy C. Stein (2004): "Social Interaction and Stock Market Participation," *Journal of Finance*, **59**, 137-163.

Honoré, Bo E. and Elie Tamer (2006): "Bounds on Parameters in Panel Discrete Choice Moels," *Econometrica*, **74**, 611-629.

Hsiao, Cheng (1986): *Analysis of Panel Data.* Cambridge University Press: Cambridge.

Hyslop, Dean R. (1999): "State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women," *Econometrica*, **67**, 1255-1294.

KING, MERVYN A. and JONATHAN I. LEAPE (1987): "Asset Accumulation, Information, and the Life Cycle," NBER Working Paper No. 2392.

LEE, LUNG-FEI (1997): "Simulated Maximum Likelihood Estimation of Dynamic Discrete Choice Statistical Models: Some Monte Carlo Results," *Journal of Econometrics*, **82**, 1-35.

MANKIW, N.GREGORY and STEPHEN P. ZELDES (1991): "The Consumption of Stockholders and Non-stockholders," *Journal of Financial Economics*, **29**, 97-112.

MALMENDIER, ULRIKE and STEFAN NAGEL (2007): "Depression Babies: Do Macroeconomic Experiences Affect Risk-Taking?," mimeo, University of California, Berkeley.

MERTON, ROBERT C. (1969): "Lifetime Portfolio Selection Under Uncertainty: The Continuous-Time Case," *Review of Economics and Statistics*, **51**, 247-257.

MINIACI, RAFFAELE AND GUGLIELMO WEBER (2002): "Econometric Issues in the Estimation of Household Portfolios," in *Household Portfolios*, eds. Luigi Guiso, Michael Haliassos, and Tullio Jappelli. MIT Press: Cambridge, 142-178.

MØLLER, STIG VINTHER (2009): "Habit Persistence: Explaining Cross-Sectional Variation in Returns and Time-Varying Expected Returns," *Journal of Empirical Finance*, forthcoming.

MUNDLAK, YAIR (1978): "On the Pooling of Time Series and Cross Section Data," *Econometrica*, **46**, 69-85.

MUÑOZ, SÓNIA (2006): "Habit Formation and Persistence in Individual Asset Portfolio Holdings: The Case of Italy," IMF Working Paper,WP/06/29.

MULLIGAN, CASEY B. and XAVIER SALA-I-MARTIN (2000): "Extensive Margins and the Demand for Money at Low Interest Rates," *Journal of Political Economy*, **100**, 961-91.

PAIELLA, MONICA (2001): "Limited Financial Market Participation: A Transaction Cost-Based Explanation," IFS Working Paper No. 0106.

POTERBA, JAMES M. and ANDREW A. SAMWICK (1995): "Stock Ownership Patterns, Stock Market Fluctuations and Consumption," *Brookings Papers on Economic Activity*, **2**, 295-372.

ROSEN, HARVEY S. and STEPHEN WU (2004): "Portfolio Choice and Health Status," *Journal of Financial Economics*, **72**, 457-484.

Samuelson, Paul A. (1969): "Lifetime Portfolio Selection by Dynamic Stochastic Programming," *Review of Economics and Statistics*, **51**, 39-46.

Samuelson, William and Richard Zeckhauser (1988): "Status Quo Bias in Decision Making," *The Journal of Risk and Uncertainty*, **1**, 7-59.

Schwarzer, Oskar (1999): *Sozialistische Zentralplan-Wirtschaft in der SBZ/DDR. Ergebnisse eines ordnungspolitischen Experiments (1945-1989).* Franz Steiner Verlag: Stuttgart.

Vissing-Jørgensen, Annette (2002a): "Towards an Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income and Participation Cost Structures," NBER Working Paper No. 8884.

Vissing-Jørgensen, Annette (2002b): "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution," *Journal of Political Economy*, **110**, 825-853.

Vissing-Jørgensen, Annette (2003): "Perspectives on Behavioral Finance: Does "Irrationality" Disappear with Wealth? Evidence from Expectations and Actions," in *NBER Macroeconomics Annual*, eds. Mark Gertler and Kenneth Rogoff. NBER: Cambridge, 139-193.

Wooldridge, Jeffrey M. (2002): *Econometric Analysis of Cross Section and Panel Data.* MIT Press: Cambridge.

Wooldridge, Jeffrey M. (2005): "Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Effects," *Journal of Applied Econometrics*, **20**, 39-54.

Figures and Tables

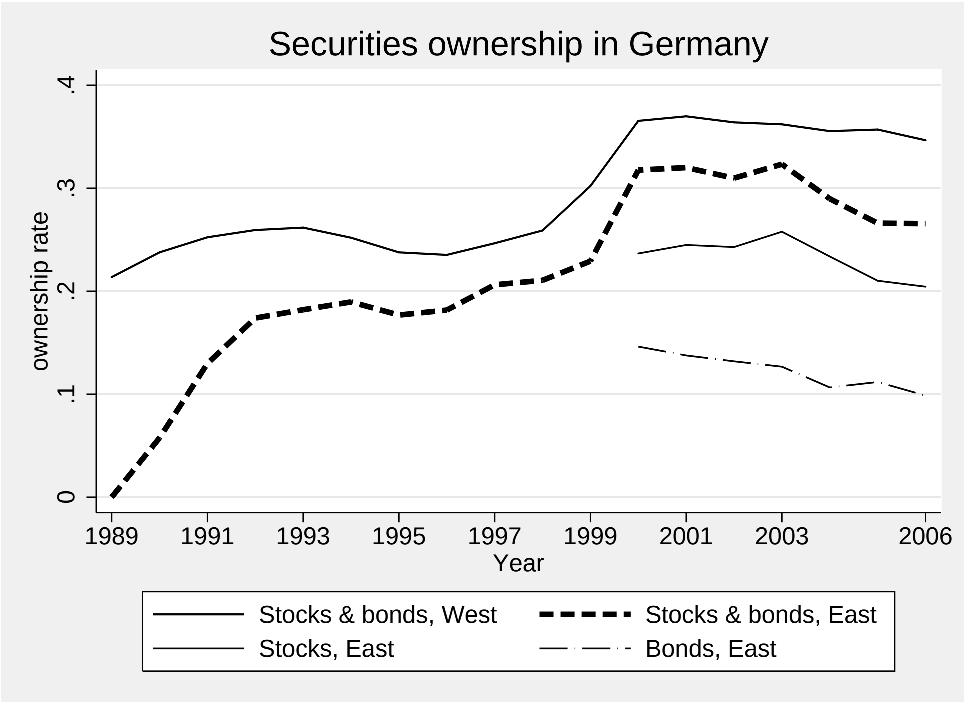

Figure 1.1: The graph depicts securities ownership rates between 1989-2006. Author's calculation based on GSOEP.

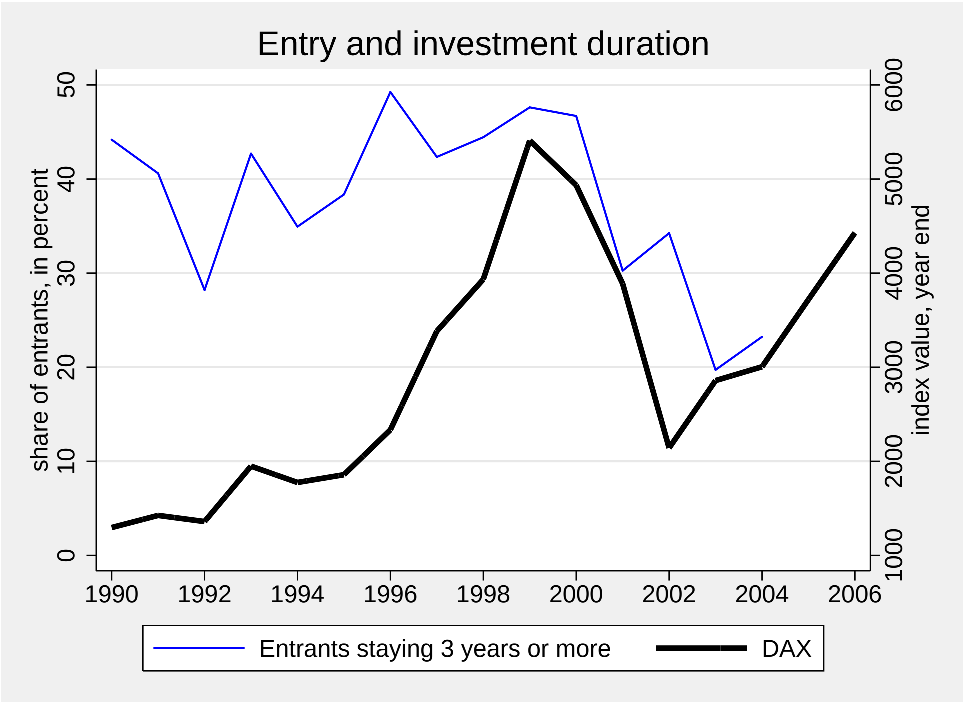

Figure 1.2: The graph depicts the average share of new entrants in a given year who invest into securities for at least the three following years. Author's calculation based on GSOEP.

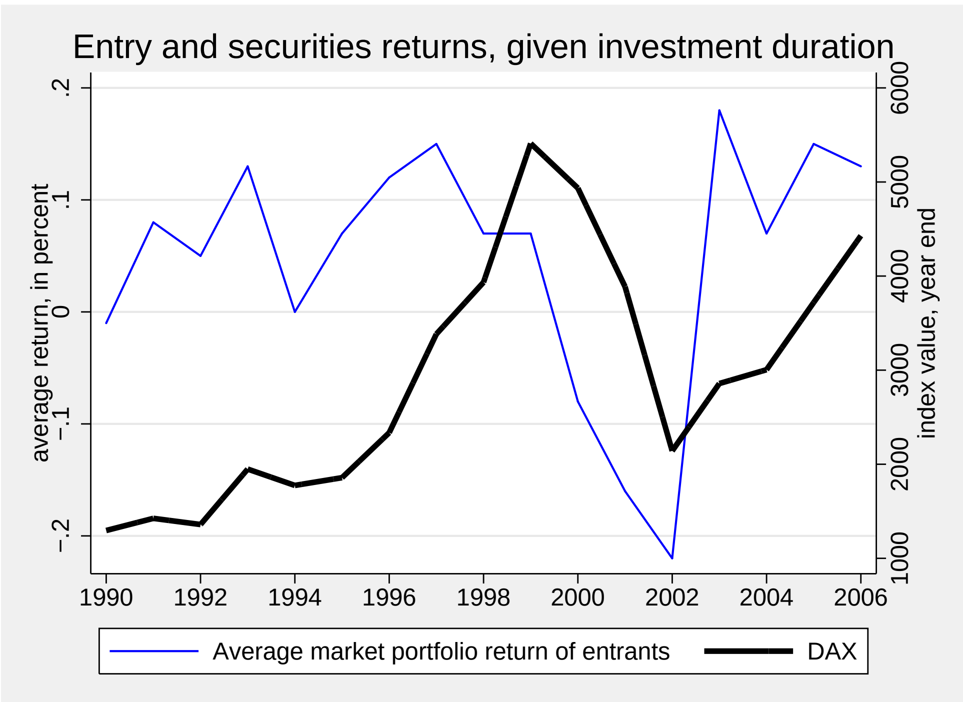

Figure 1.3: The graph depicts the average market return, which entrants in a a given year experienced given the length of their investment. Author's calculation based on GSOEP.

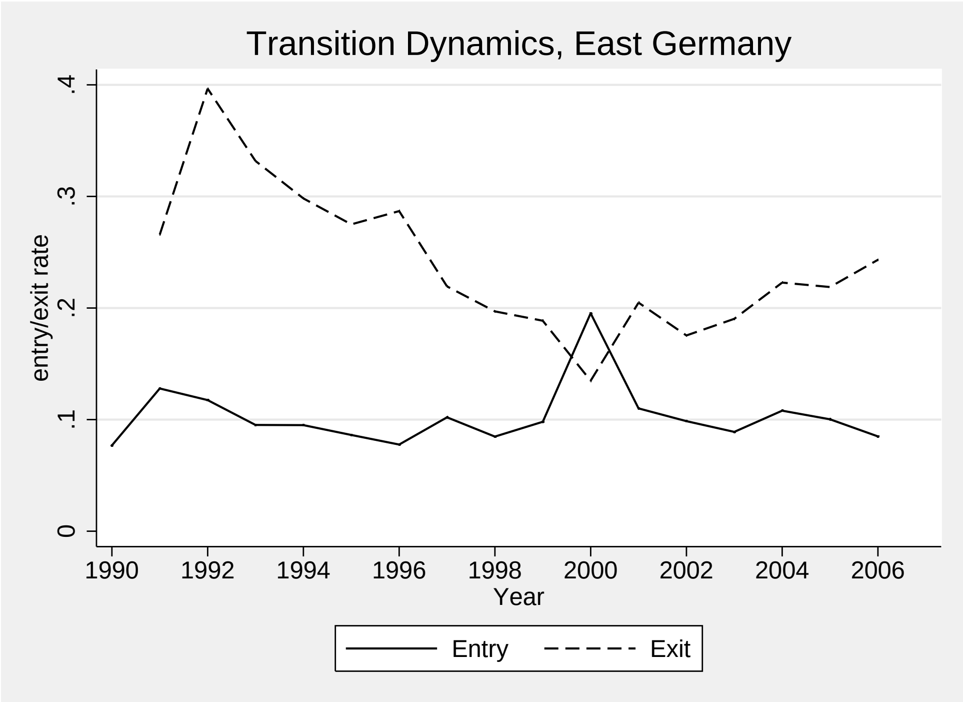

Figure 1.4: The graph depicts annual transition patterns between 1990-2006. Author's calculation based on GSOEP.

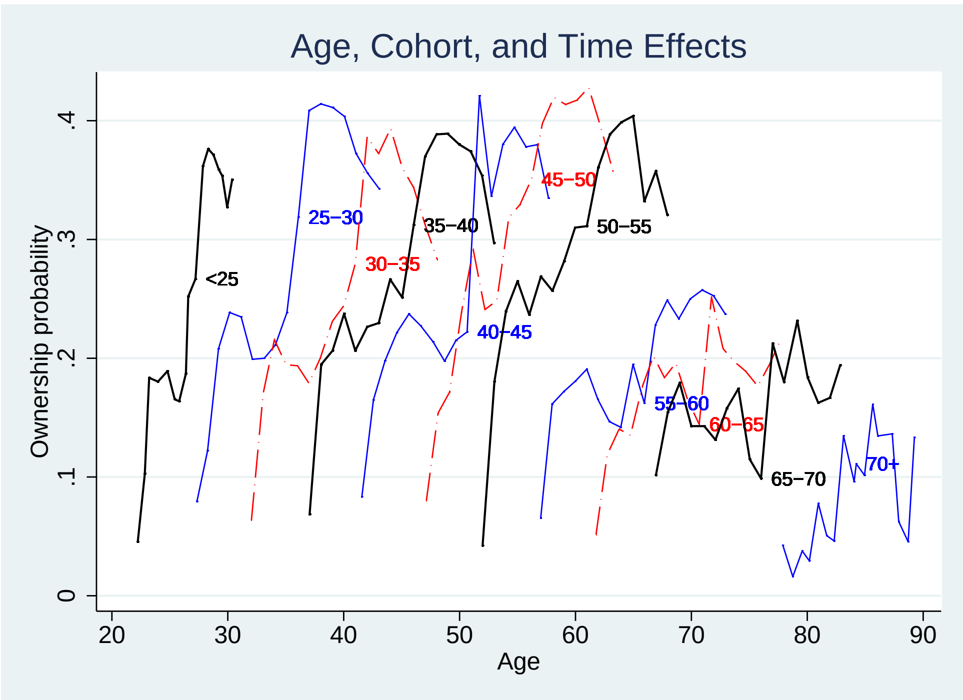

Figure 1.5: The graph depicts age and cohort patterns between 1990-2006. Author's calculation based on GSOEP.

Table 1.1: German household savings

	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Cash	9.4	10.0	10.0	9.8	9.7	9.9	9.6	9.8	9.8	9.7	9.7	12.1	13.1	13.6	14.1	13.9
Time deposits	10.5	11.2	11.4	9.2	7.0	5.6	5.0	4.8	7.0	7.2	7.5	7.4	6.5	6.0	5.7	6.1
Deposits	23.3	22.4	22.3	23.2	23.8	24.1	23.2	22.2	17.6	16.1	15.8	15.9	15.3	15.0	14.2	13.1
Total Cash/Deposits	**43.1**	**43.5**	**43.7**	**42.2**	**40.4**	**39.7**	**37.8**	**36.8**	**34.5**	**33.0**	**32.9**	**35.4**	**34.9**	**34.7**	**33.9**	**33.2**
Savings certificates	5.0	4.7	3.9	3.5	3.6	3.3	3.1	2.8	2.3	2.2	2.1	2.1	1.9	1.8	1.6	1.6
as % of total securities	**17.0**	**16.6**	**13.4**	**12.0**	**11.5**	**10.4**	**9.3**	**8.3**	**6.3**	**6.0**	**5.9**	**6.7**	**5.7**	**5.3**	**4.6**	**4.7**
Commercial paper	0.3	0.4	0.3	0.2	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
as % of total securities	**1.0**	**1.3**	**1.0**	**0.6**	**0.3**	**0.3**	**0.2**	**0.1**	**0.1**	**0.1**	**0.1**	**0.1**	**0.1**	**0.1**	**0.1**	**0.1**
Bonds	8.9	8.1	7.1	7.0	8.4	8.3	7.9	7.2	6.7	6.4	6.8	7.4	7.4	8.1	7.6	8.2
as % of total securities	**30.6**	**28.4**	**24.5**	**23.7**	**27.3**	**26.4**	**23.7**	**20.9**	**18.1**	**17.5**	**18.9**	**23.2**	**22.6**	**24.4**	**22.4**	**23.4**
Derivatives	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
as % of total securities	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Stocks	6.6	5.8	7.4	7.2	7.5	8.2	10.1	11.3	13.9	12.7	10.1	5.7	6.9	6.9	7.6	8.3
as % of total securities	**22.4**	**20.5**	**25.7**	**24.4**	**24.1**	**26.0**	**30.2**	**33.1**	**37.8**	**34.6**	**28.1**	**17.9**	**21.0**	**21.1**	**22.2**	**23.9**
other shareholdings	4.1	4.2	4.3	4.3	4.0	4.0	4.0	3.7	3.3	3.8	4.9	4.6	4.4	4.6	5.1	5.1
as % of total securities	**14.1**	**14.7**	**14.8**	**14.7**	**12.9**	**12.6**	**12.1**	**10.8**	**8.9**	**10.3**	**13.5**	**14.5**	**13.3**	**13.9**	**15.0**	**14.6**
Investment certificates	4.3	5.3	5.9	7.3	7.4	7.6	8.2	9.1	10.5	11.6	12.1	11.9	12.2	11.6	12.2	11.7
as % of total securities	**14.9**	**18.6**	**20.5**	**24.6**	**24.0**	**24.3**	**24.6**	**26.7**	**28.7**	**31.6**	**33.6**	**37.5**	**37.2**	**35.2**	**35.7**	**33.3**
Total Securities	**29.2**	**28.4**	**29.0**	**29.5**	**30.9**	**31.4**	**33.5**	**34.2**	**36.7**	**36.9**	**36.1**	**31.7**	**32.8**	**33.0**	**34.1**	**35.0**
Claims against insurances	19.7	20.1	19.9	20.8	21.2	21.7	21.9	22.2	22.2	23.4	24.1	25.2	24.9	25.0	25.0	25.1
Claims against pension funds	6.8	6.9	6.4	6.4	6.3	6.0	5.7	5.6	5.3	5.5	5.6	6.0	5.9	5.9	5.7	5.7
other claims	1.1	1.1	1.1	1.1	1.1	1.2	1.2	1.2	1.3	1.3	1.3	1.6	1.5	1.5	1.2	1.1
Total Insurance	**26.6**	**27.0**	**26.3**	**27.2**	**27.5**	**27.7**	**27.6**	**27.8**	**27.6**	**28.9**	**29.7**	**31.2**	**30.8**	**30.9**	**30.7**	**30.8**
Total (in billion Euro	**1926.1**	**2059.5**	**2288.2**	**2391.3**	**2563.7**	**2745.7**	**2961.8**	**3168.4**	**3438**	**3509.8**	**3602.1**	**3569.9**	**3802.9**	**3975.9**	**4208.9**	**4411.5**

Note: The table reports the allocation of total private asset holdings for both East an West Germany across different assets classes as a percentage of the total. Total annual private asset holdings are in billion Euro. Source: Deutsche Bundesbank (2008).

Table 1.2: Annual transitions

as % of all observations		
	$S_{it} = 0$	$S_{it} = 1$
$S_{i,t-1} = 0$	89.52	10.48
$S_{i,t-1} = 1$	22.84	77.16
Total	71.5	28.5

Table 1.3: Transitions: 1995 - 2005

	$S_{i,2005} = 0$	$S_{i,2005} = 1$
$S_{i,1995} = 0$	76.68	23.32
$S_{i,1995} = 1$	32.38	67.62
Total	66.35	33.65

Table 1.4: Experience in securities markets

Years of Experience	Frequency	Percent	Cumulative
0	351	32.56	32.56
1	123	11.41	43.97
2	75	6.96	50.93
3	64	5.94	56.86
4	41	3.8	60.67
5	44	4.08	64.75
6	42	3.9	68.65
7	46	4.27	72.91
8	36	3.34	76.25
9	35	3.25	79.5
10	28	2.6	82.1
11	36	3.34	85.44
12	28	2.6	88.03
13	32	2.97	91
14	36	3.34	94.34
15	26	2.41	96.75
16	21	1.95	98.7
17	14	1.3	100
Total	1,078	100	

Note: Experience is constructed as the total number of years during which households owned securities between 1990 and 2006.

Table 1.5: Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
S_t	17179	0.274	0.445	0	1
S_{t-1}	17179	0.255	0.435	0	1
Hhold income 20%-40%	17179	0.197	0.398	0	1
Hhold income 40%-60%	17179	0.222	0.416	0	1
Hhold income 60%-80%	17179	0.216	0.411	0	1
Hhold income 80%+	17179	0.221	0.415	0	1
one child	17179	0.179	0.383	0	1
two children	17179	0.133	0.339	0	1
three children+	17179	0.036	0.186	0	1
married	17179	0.925	0.263	0	1
owns house	17179	0.375	0.484	0	1
unemployed	17179	0.120	0.325	0	1
age/10	17179	5.121	1.359	1.9	9.3
$(\text{age}/10)^2$	17179	28.073	14.457	3.6	86.5
cohort 25-35	17179	0.273	0.445	0	1
cohort 35-45	17179	0.228	0.419	0	1
cohort 45-55	17179	0.249	0.433	0	1
cohort 55-65	17179	0.149	0.356	0	1
cohort 65+	17179	0.0592	0.236	0	1
retiree	17179	0.302	0.459	0	1
10 years schooling	17179	0.451	0.498	0	1
13 years schooling	17179	0.181	0.385	0	1
college	17179	0.198	0.399	0	1
university	17179	0.128	0.334	0	1
dividends_{t-1} 500-2000	16192	0.259	0.438	0	1
dividends_{t-1} 2000-5000	16192	0.079	0.270	0	1
dividends_{t-1} 5000-10000	16192	0.027	0.161	0	1
dividends_{t-1} 10000+	16192	0.009	0.093	0	1
financial worries	17118	0.269	0.443	0	1
job worries	16724	0.160	0.367	0	1

Table 1.6: RE with higher order lags

	(1)	
	β	se
S_{t-1}	1.273***	(0.038)
S_{t-2}	0.578***	(0.041)
S_{t-3}	0.400***	(0.043)
S_{t-4}	0.281***	(0.045)
S_{t-5}	0.179***	(0.045)
Hhold income 20%-40%	0.153*	(0.077)
Hhold income 40%-60%	0.172*	(0.082)
Hhold income 60%-80%	0.233**	(0.086)
Hhold income 80%+	0.281**	(0.091)
one child	0.124	(0.064)
two children	0.093	(0.090)
three children+	0.338*	(0.161)
married	0.013	(0.175)
owns house	-0.009	(0.073)
unemployed	-0.011	(0.061)
age/10	0.003	(0.168)
$(age/10)^2$	-0.003	(0.015)
cohort 25-35	-0.081	(0.101)
cohort 35-45	-0.066	(0.141)
cohort 45-55	0.091	(0.186)
cohort 55-65	0.002	(0.227)
cohort 65+	-0.020	(0.289)
retiree	-0.084	(0.059)
10 years schooling	0.122**	(0.044)
13 years schooling	0.203**	(0.064)
College	0.058	(0.041)
University	0.011	(0.064)
$\overline{Hhold\ income}\ 20\% - 40\%$	0.071	(0.144)
$\overline{Hhold\ income}\ 40\% - 60\%$	0.244	(0.132)
$\overline{Hhold\ income}\ 60\% - 80\%$	0.394**	(0.139)
$\overline{Hhold\ income}\ > 80\%$	0.497***	(0.135)
$\overline{one\ child}$	-0.274*	(0.113)
$\overline{two\ children}$	-0.185	(0.130)
$\overline{three\ children+}$	-0.942***	(0.219)
$\overline{married}$	-0.196	(0.188)
$\overline{owns\ house}$	0.055	(0.085)
$\overline{unemployed}$	-0.196	(0.130)
constant	-1.923***	(0.440)
$\ln(\hat{\sigma}^2)$	-5.967***	(-0.124)
N	12709	
$\chi2$	5006.5	
AIC	8509.5	
BIC	8889.5	

Note: The dependent variable is one if a household owns one or more securities and zero otherwise. Time dummies are suppressed.

Table 1.7: Estimates from standard RE and AR(1)-RE models.

	(2) RE		(3) AR(1)-RE		(4) RE		(5) AR(1)-RE	
	β	se	β	se	β	se	β	se
S_{t-1}	1.191***	(0.036)	0.482***	(0.092)	1.177***	(0.037)	0.342***	(0.103)
exp					0.075***	(0.011)	0.215***	(0.029)
rexp					-0.137	(0.228)	0.161	(0.386)
Hhold income 20%-40%	0.180*	(0.071)	0.240*	(0.102)	0.171*	(0.069)	0.247*	(0.107)
Hhold income 40%-60%	0.230**	(0.075)	0.313**	(0.111)	0.221**	(0.072)	0.326**	(0.117)
Hhold income 60%-80%	0.334***	(0.078)	0.416***	(0.115)	0.324***	(0.075)	0.440***	(0.122)
Hhold income 80+	0.368***	(0.082)	0.489***	(0.124)	0.362***	(0.079)	0.523***	(0.132)
one child	0.079	(0.056)	0.151	(0.092)	0.084	(0.055)	0.179	(0.099)
two children	0.034	(0.077)	0.077	(0.131)	0.053	(0.075)	0.124	(0.140)
three children +	0.079	(0.144)	0.209	(0.240)	0.066	(0.138)	0.208	(0.253)
married	-0.035	(0.150)	-0.010	(0.254)	-0.014	(0.144)	0.029	(0.261)
owns house	0.077	(0.062)	0.059	(0.108)	0.040	(0.061)	0.003	(0.115)
unemployed	0.006	(0.054)	0.002	(0.078)	0.012	(0.053)	0.009	(0.081)
age/10	0.247	(0.173)	0.545	(0.353)	0.219	(0.155)	0.450	(0.332)
$(age/10)^2$	-0.038**	(0.013)	-0.079**	(0.028)	-0.033*	(0.013)	-0.064*	(0.028)
cohort 25-35	0.029	(0.172)	-0.022	(0.314)	-0.008	(0.141)	-0.08	(0.266)
cohort 35-45	0.092	(0.237)	0.114	(0.436)	0.039	(0.194)	0.006	(0.368)
cohort 45-55	0.525	(0.327)	0.840	(0.595)	0.388	(0.266)	0.583	(0.499)
cohort 55-65	0.526	(0.411)	0.891	(0.746)	0.392	(0.333)	0.625	(0.617)
cohort 65+	0.732	(0.512)	1.391	(0.929)	0.575	(0.416)	1.033	(0.767)
retiree	-0.057	(0.063)	-0.071	(0.107)	-0.067	(0.060)	-0.11	(0.111)
10 years schooling	0.180**	(0.068)	0.357**	(0.135)	0.170**	(0.058)	0.336**	(0.119)
13 years schooling	0.296**	(0.094)	0.630**	(0.192)	0.279***	(0.083)	0.588***	(0.176)
college	0.227**	(0.074)	0.419**	(0.145)	0.162**	(0.061)	0.298*	(0.119)
university	0.153	(0.104)	0.249	(0.198)	0.108	(0.089)	0.159	(0.174)
$\overline{Hhold\ income}$ $20\% - 40\%$	0.127	(0.224)	0.302	(0.428)	0.090	(0.181)	0.228	(0.337)
$\overline{Hhold\ income}$ $20\% - 40\%$	0.635**	(0.199)	1.229**	(0.380)	0.473**	(0.163)	0.903**	(0.309)
$\overline{Hhold\ income}$ $20\% - 40\%$	0.686**	(0.210)	1.515***	(0.413)	0.521**	(0.172)	1.181***	(0.334)
$\overline{Hhold\ income}$ $20\% - 40\%$	1.212***	(0.192)	2.332***	(0.396)	0.890***	(0.163)	1.685***	(0.325)
$\overline{one\ child}$	-0.219	(0.181)	-0.443	(0.331)	-0.209	(0.148)	-0.442	(0.275)
$\overline{two\ children}$	-0.096	(0.192)	-0.215	(0.334)	-0.117	(0.159)	-0.281	(0.287)
$\overline{three\ children+}$	-1.340***	(0.311)	-2.505***	(0.622)	-1.074***	(0.264)	-2.058***	(0.527)
$\overline{married}$	-0.422*	(0.199)	-0.863*	(0.361)	-0.349*	(0.177)	-0.722*	(0.331)
$\overline{owns\ house}$	-0.008	(0.096)	0.084	(0.177)	0.017	(0.084)	0.115	(0.159)
$\overline{unemployed}$	-0.205	(0.210)	-0.380	(0.396)	-0.186	(0.169)	-0.348	(0.312)
$\ln(\hat{\sigma})$	-0.421***	(0.082)	0.635***	(0.087)	-1.070***	(0.136)	0.504***	(0.101)
implied $\hat{\sigma}^2$	*0.430*		*3.560*		*0.118*		*2.740*	
$\arctan(\hat{\rho})$			1.703***	(0.043)			1.314***	(0.055)
implied $\hat{\rho}$			*0.936*				*0.865*	
N	17179		17179		17179		17179	
$\chi 2$	2560.3		846.9		2904.0		708.5	
AIC	11908.9		11725.7		11876.8		11667.6	
BIC	12304.2		12128.8		12287.6		12086.2	

Note: The dependent variable is one if a household owns one or more securities and zero otherwise. The constant and time dummies are suppressed.

Table 1.8: Average marginal effects

	(2)		(3)		(4)		(5)	
	ape	se	ape	se	ape	se	ape	se
S_{t-1}	0.286	(0.010)	0.081	(0.017)	0.048	(0.015)	0.052	(0.017)
exp					0.056	(0.008)	0.040	(0.005)
rexp					0.042	(0.100)	0.030	(0.071)
Hhold income 20%-40%	0.033	(0.014)	0.033	(0.015)	0.031	(0.014)	0.032	(0.015)
Hhold income 40%-60%	0.043	(0.015)	0.044	(0.017)	0.042	(0.016)	0.043	(0.017)
Hhold income 60%-80%	0.064	(0.016)	0.061	(0.019)	0.058	(0.017)	0.059	(0.018)
Hhold income 80%+	0.071	(0.017)	0.073	(0.021)	0.069	(0.019)	0.072	(0.020)
10 years schooling	0.034	(0.013)	0.050	(0.021)	0.044	(0.016)	0.045	(0.017)
13 years schooling	0.058	(0.020)	0.095	(0.034)	0.079	(0.026)	0.083	(0.029)
college	0.044	(0.015)	0.059	(0.023)	0.038	(0.016)	0.039	(0.017)
university	0.029	(0.020)	0.033	(0.029)	0.020	(0.023)	0.020	(0.023)
	(5)		(6)		(7)		(8)	
	ape	se	ape	se	ape	se	ape	se
S_{t-1}	0.288	(0.012)	0.056	(0.018)	0.278	(0.012)	0.126	(0.018)
exp	0.038	(0.004)	0.055	(0.007)	0.029	(0.003)	0.034	(0.004)
rexp	-0.030	(0.055)	0.018	(0.078)	-0.030	(0.049)	-0.011	(0.063)
Hhold income 20%-40%	0.030	(0.013)	0.031	(0.014)	0.025	(0.013)	0.026	(0.014)
Hhold income 40%-60%	0.040	(0.014)	0.042	(0.016)	0.034	(0.014)	0.036	(0.015)
Hhold income 60%-80%	0.060	(0.015)	0.059	(0.018)	0.050	(0.015)	0.051	(0.016)
Hhold income 80%+	0.068	(0.016)	0.071	(0.020)	0.056	(0.016)	0.062	(0.018)
10 years schooling	0.030	(0.011)	0.041	(0.016)	0.016	(0.009)	0.020	(0.012)
13 years schooling	0.052	(0.016)	0.077	(0.026)	0.032	(0.013)	0.040	(0.018)
college	0.023	(0.011)	0.033	(0.015)	0.019	(0.009)	0.026	(0.012)
university	0.015	(0.015)	0.017	(0.021)	-0.002	(0.013)	-0.002	(0.016)

Note: Average marginal effects are reported for dummy and continuous variables. Standard errors (calculated with the delta-method) are reported in parenthesis.

Table 1.9: Robustness checks

	(6) RE		(7) AR(1)-RE		(8) RE		(9) AR(1)-RE	
	β	se	β	se	β	se	β	se
S_{t-1}	1.080***	(0.039)	0.355***	(0.103)	1.080***	(0.041)	0.668***	(0.086)
exp	0.229***	(0.021)	0.352***	(0.044)	0.199***	(0.019)	0.245***	(0.026)
exp^2	-0.011***	(0.001)	-0.010***	(0.003)	-0.010***	(0.001)	-0.010***	(0.002)
rexp	-0.316	(0.217)	-0.016	(0.363)	-0.354	(0.216)	-0.231	(0.303)
$rexp^2$	3.296***	(0.960)	1.870	(1.491)	3.657***	(0.978)	2.905*	(1.320)
Hhold income 20%-40%	0.161*	(0.067)	0.232*	(0.102)	0.142*	(0.070)	0.178*	(0.089)
Hhold income 40%-60%	0.212**	(0.071)	0.310**	(0.112)	0.193**	(0.074)	0.242*	(0.095)
Hhold income 60%-80%	0.310***	(0.074)	0.421***	(0.116)	0.275***	(0.078)	0.333***	(0.100)
Hhold income 80%+	0.347***	(0.077)	0.500***	(0.125)	0.306***	(0.081)	0.401***	(0.107)
one child	0.086	(0.054)	0.168	(0.093)	0.078	(0.056)	0.103	(0.076)
two children	0.064	(0.073)	0.129	(0.130)	0.033	(0.076)	0.056	(0.107)
three children+	0.079	(0.135)	0.199	(0.235)	-0.004	(0.139)	0.047	(0.187)
married	0.011	(0.141)	0.038	(0.243)	0.089	(0.154)	0.102	(0.214)
owns house	0.025	(0.060)	-0.003	(0.108)	0.083	(0.063)	0.077	(0.089)
unemployed	0.005	(0.052)	0.006	(0.077)	0.009	(0.057)	0.019	(0.070)
age/10	0.160	(0.144)	0.353	(0.295)	-0.008	(0.140)	0.019	(0.221)
$(age/10)^2$	-0.025*	(0.012)	-0.052*	(0.025)	-0.019	(0.012)	-0.031	(0.019)
retiree	-0.064	(0.058)	-0.107	(0.104)	-0.044	(0.059)	-0.053	(0.084)
10 years schooling	0.155**	(0.053)	0.298**	(0.105)	0.090	(0.049)	0.127	(0.074)
13 years schooling	0.264***	(0.075)	0.524***	(0.156)	0.174*	(0.070)	0.255*	(0.108)
college	0.121*	(0.053)	0.238*	(0.103)	0.103*	(0.047)	0.165*	(0.072)
university	0.077	(0.079)	0.124	(0.151)	-0.014	(0.072)	-0.012	(0.109)
$\overline{Hhold\ income\ 20\% - 40\%}$	0.062	(0.159)	0.177	(0.291)	-0.079	(0.149)	-0.104	(0.215)
$\overline{Hhold\ income\ 40\% - 60\%}$	0.382**	(0.144)	0.746**	(0.271)	-0.089	(0.136)	-0.117	(0.197)
$\overline{Hhold\ income\ 60\% - 80\%}$	0.411**	(0.151)	0.964***	(0.291)	0.020	(0.142)	0.112	(0.206)
$\overline{Hhold\ income\ 80\%+}$	0.714***	(0.147)	1.401***	(0.288)	0.058	(0.137)	0.095	(0.199)
$dividends_{t-1}$ 500-2000					0.011	(0.040)	0.024	(0.050)
$dividends_{t-1}$ 2000-5000					0.003	(0.067)	0.032	(0.084)
$dividends_{t-1}$ 2000-10000					0.117	(0.106)	0.137	(0.132)
$dividends_{t-1}$ 10000+					0.356*	(0.166)	0.426	(0.226)
$\overline{dividends_{t-1}\ 500 - 2000}$					1.071***	(0.102)	1.585***	(0.180)
$\overline{dividends_{t-1}\ 2000 - 5000}$					1.475***	(0.167)	2.192***	(0.281)
$\overline{dividends_{t-1}\ 5000 - 10000}$					1.430***	(0.303)	2.072***	(0.469)
$\overline{dividends_{t-1}\ 10000+}$					1.011*	(0.476)	1.700*	(0.784)
financial worries					-0.067	(0.043)	-0.076	(0.053)
$\overline{financial\ worries}$					-0.210*	(0.107)	-0.327*	(0.161)
job worries					0.064	(0.048)	0.095	(0.060)
$\overline{job\ worries}$					0.031	(0.134)	0.025	(0.199)
$\ln(\hat{\sigma})$	-1.611***	(0.208)	0.365**	(0.111)	-2.306***	(0.246)	-0.100	(0.115)
implied $\hat{\sigma}^2$	*0.040*		*0.481*		*0.010*		*0.819*	
$\arctan(\hat{\rho})$			1.223***	(0.055)			1.185***	(0.048)
implied $\hat{\rho}$			*0.841*				*0.829*	
N	17179		17179		16025		16025	
$\chi 2$	3250.0		775.0		4305.0		1463.5	
AIC	11808.7		11658.0		10711.0		10635.8	
BIC	12235.1		12092.0		11225.7		11158.1	

Note: The dependent variable is one if a household owns one or more securities and zero otherwise. The constant, time dummies, cohort dummies, and selected variables from $\overline{\mathbf{x}}$ are suppressed.

Table 1.10: Testing for attrition bias

	(10) unweighted		(11) weighted		(12) unweighted		(13) weighted	
	β	se	β	se	β	se	β	se
S_{t-1}	1.837***	(0.028)	1.854***	(0.029)	1.196***	(0.037)	1.184***	(0.039)
exp					0.286***	(0.014)	0.298***	(0.015)
exp^2					-0.012***	(0.001)	-0.012***	(0.001)
rexp					-0.304	(0.194)	-0.258	(0.206)
$rexp^2$					3.749***	(0.876)	3.639***	(0.946)
Hhold income 20%-40%	0.142*	(0.060)	0.104	(0.067)	0.140*	(0.062)	0.094	(0.069)
Hhold income 40%-60%	0.176**	(0.064)	0.156*	(0.070)	0.184**	(0.066)	0.157*	(0.072)
Hhold income 60%-80%	0.257***	(0.066)	0.238**	(0.072)	0.271***	(0.068)	0.243**	(0.074)
Hhold income 80%+	0.269***	(0.070)	0.259***	(0.077)	0.302***	(0.071)	0.278***	(0.079)
one child	0.060	(0.050)	0.051	(0.053)	0.085	(0.051)	0.077	(0.054)
two children	0.021	(0.067)	0.015	(0.070)	0.075	(0.068)	0.069	(0.070)
three children+	0.078	(0.120)	0.075	(0.124)	0.061	(0.121)	0.069	(0.124)
married	-0.025	(0.128)	-0.008	(0.143)	0.036	(0.129)	0.041	(0.142)
owns house	0.043	(0.057)	0.030	(0.060)	-0.014	(0.058)	-0.030	(0.061)
unemplreg	0.028	(0.048)	0.013	(0.049)	0.017	(0.048)	-0.001	(0.050)
age/10	0.055	(0.119)	0.087	(0.123)	0.081	(0.121)	0.120	(0.126)
$(age/10)^2$	-0.012	(0.011)	-0.016	(0.011)	-0.013	(0.011)	-0.019	(0.011)
cohort 25-35	-0.004	(0.083)	0.014	(0.085)	-0.072	(0.085)	-0.065	(0.087)
cohort 35-45	0.026	(0.114)	0.046	(0.118)	-0.057	(0.117)	-0.046	(0.120)
cohort 45-55	0.256	(0.150)	0.283	(0.153)	0.139	(0.155)	0.167	(0.158)
cohort 55-65	0.215	(0.183)	0.287	(0.187)	0.114	(0.190)	0.184	(0.194)
cohort 65+	0.309	(0.229)	0.392	(0.234)	0.179	(0.238)	0.275	(0.243)
retiree	-0.039	(0.050)	-0.023	(0.052)	-0.062	(0.052)	-0.047	(0.054)
10 years schooling	0.127***	(0.037)	0.147***	(0.040)	0.116**	(0.038)	0.133**	(0.041)
13 years schooling	0.211***	(0.053)	0.232***	(0.054)	0.198***	(0.055)	0.216***	(0.056)
college	0.110**	(0.035)	0.116**	(0.037)	0.063	(0.035)	0.047	(0.038)
university	0.097	(0.053)	0.067	(0.054)	0.030	(0.055)	-0.002	(0.056)
$\overline{Hhold\ income}\ 20\% - 40\%$	0.086	(0.109)	0.227*	(0.115)	0.038	(0.112)	0.125	(0.119)
$\overline{Hhold\ income}\ 20\% - 40\%$	0.311**	(0.101)	0.385***	(0.108)	0.232*	(0.104)	0.307**	(0.111)
$\overline{Hhold\ income}\ 20\% - 40\%$	0.352***	(0.106)	0.421***	(0.113)	0.257*	(0.108)	0.318**	(0.115)
$\overline{Hhold\ income}\ 20\% - 40\%$	0.646***	(0.101)	0.745***	(0.110)	0.427***	(0.104)	0.511***	(0.113)
$\overline{one\ child}$	-0.177*	(0.089)	-0.175	(0.093)	-0.185*	(0.092)	-0.183	(0.095)
$\overline{two\ children}$	-0.091	(0.099)	-0.087	(0.103)	-0.151	(0.102)	-0.164	(0.106)
$\overline{three\ children+}$	-0.870***	(0.174)	-0.879***	(0.180)	-0.622***	(0.178)	-0.641***	(0.183)
$\overline{married}$	-0.205	(0.139)	-0.272	(0.154)	-0.221	(0.141)	-0.258	(0.154)
$\overline{owns\ house}$	-0.013	(0.066)	0.044	(0.070)	0.053	(0.066)	0.106	(0.070)
$\overline{unemployed}$	-0.135	(0.099)	-0.142	(0.104)	-0.157	(0.101)	-0.157	(0.106)
N	17179		17179		17179		17179	
$\chi 2$	6074.0		5673.9		6166.9		5913.8	
AIC	12616.5		12042.7		11864.7		11274.6	
BIC	13004.0		12430.3		12283.3		11693.2	

Note: Estimates are from pooled probit regressions. Robust standard errors are reported in parentheses. Specifications (11) and (13) are weighted with inverse probability weights from probit regressions of the sampling probability. The constant and time dummies are suppressed.

Chapter 2

Tax Incentives, Bequest Motives, and the Demand for Life Insurance: Evidence from Two Natural Experiments in Germany[1]

2.1 Introduction

Life insurance is one of the most popular financial assets owned by a large number of households in many countries (Guiso, Haliassos, Jappelli, 2002). In its simplest form – term life insurance – it enables the policyholder to pass on bequests to children or other beneficiaries if he or she dies before a certain point in time (the end of the term). However, in many countries, life insurance products are a popular savings vehicle for old age as well. Under whole life insurance contracts, the insurer faces a certain liability over the whole lifetime of the insured, for which the insurer accumulates reserves during the working life of the policyholder. Typically, the policyholder has the right to withdraw the savings component in old age, provided he or she survives. As a result, under whole life insurance term life insurance provisions are coupled with a savings contract. This savings component of whole life insurance often receives tax preferences. Studying the demand for whole life insurance ownership has significant appeal as it allows testing for both the importance of tax incentives and bequest motives in households' savings decisions. This paper explores these two aspects through a study of whole life insurance ownership in Germany.

[1]This chapter is joint work together with Jan Walliser and Joachim Winter.

Interest in providing incentives for retirement savings (including through whole life insurance) intensified during the past few years. Many governments have already reduced the generosity of existing pay-as-you-go pension systems, or are considering doing so as their population ages. Hence, households are increasingly pressed to increase their private savings portfolio in order to sustain the standards of living during retirement, and governments seek to encourage these savings through preferential tax rules. The empirical evidence on the importance of such tax incentives is, however, inconclusive. Scholz (1994) documents little evidence that households modified their portfolios in response to the 1986 US Tax Reform Act. Also, Jappelli and Pistaferri (2003, 2007) do not find significant changes in the demand for life insurance and mortgage debt by those households most affected by incremental tax reforms in Italy. On the other hand, several studies that use cross-sectional data report a positive correlation between marginal tax rates and investments channeled into tax-sheltered assets for Canada (Alan et al, 2008), Denmark (Alan and Leth-Petersen, 2006), the Netherlands (Alessie, Hochguertel, van Soest, 1997), Sweden (Agell and Edlin, 1991), the United Kingdom (Banks and Tanner, 2001), and the US (King and Leape, 1998; Poterba, 2002; Poterba and Samwick, 2003). Yet in cross-sections, it is difficult to disentangle genuine variation in marginal tax rates for given income, from genuine variation in income for given tax rates, because after-tax-yields depend on changing marginal tax rates, which in turn depend on income levels.[2]

Our analysis contributes to the literature on taxation and portfolio choice by exploiting a natural experiment of changes to tax incentives for whole life insurance in Germany. A tax reform in 2000 enables us to identify different investor responses among those affected by the reform, and a control group that remains unaffected without further need for statistical inferences. The changes in tax laws reduced the limit on tax exemptions and created a strong incentive among households that were fully exempt from capital income taxation before the reform to shelter their savings from taxation by investing in (tax-exempt) life insurance contracts afterwards. In contrast to the prior literature, this allows us to review the impact of changing tax incentives at the margin rather than incremental changes in after-tax returns. The results suggest that standard tax revenue estimates, which ignore behavioral changes of portfolio choices, may overestimate potential revenue effects from introducing capital income taxation (Poterba and Verdugo, 2008).

A second aspect studied in this paper relates to the importance of bequest motives as driver for savings behavior. Empirical studies disagree about the strength of bequest

[2]In his seminal contribution, Feldstein (1976) even uses labor income as a proxy for the marginal tax rate.

motives. Estimates of the share of bequests in aggregate private savings range from 17 (Modigliani, 1988) to 46 percent (Kotlikoff and Summers, 1981). Cross-country evidence shows that life insurance demand is higher in countries with a high dependency ratio (Browne and Kim, 1993), high income per capita, low inflation, and a high degree of banking sector development (Beck and Webb, 2003). At the household level, Bernheim (1991) finds that a significant fraction of life insurance demand and consumption can be motivated by the desire to leave bequests to one's children. Kopczuk and Lupton (2007) estimate that households with a bequest motive save about 25 percent more, whereas Hurd (1987, 1989) finds that the marginal utility from bequests in a consumption-savings model is close to zero. Data on direct questions for the intention to leave bequest has been used by Laitner and Juster (1995) and Jürges (2001). Although both find that bequest motives shape savings behavior, altruism toward one's children appears to be of only minor importance. All these studies suffer from the difficulty to distinguish true bequest motives from other savings motives, such as tax, life-cycle, or precautionary motives. In this paper, we are able to identify the importance of bequest motives in the demand for life insurance by exploiting the natural experiment of the division of Germany into two separate states. Owing to the absence of tax incentives and the limited number of consumption and savings possibilities in the former German Democratic Republic, we can isolate the impact of bequest motives on the demand for life insurance, while controlling for the main life-cycle and precautionary savings motives.

Our key empirical findings confirm the predictions from our theoretical model. First, the probability to own tax-exempt whole life insurance contracts increases by 6 percent among households affected by the tax reform in 2000 (i.e., among those households loosing their exemption from capital income taxation). Second, there is also strong indication that households in the former GDR – where life insurance demand was not diluted by tax considerations – purchased life insurance to bequeath wealth to their children, whereas provision for non-working partners seems to play a lesser role.

This paper proceeds by discussing some key theoretical predictions from a formal model of life insurance demand in section 2.2. Section 2.3 describes the data. Section 2.4 analyzes the impact of the German tax reform in 2000 on life insurance demand. In section 2.5, estimates of the strength of bequest motives in GDR life insurance demand are reported. Finally, section 2.6 wraps up the discussion.

2.2 A Life-Cycle Model with Tax incentives and Bequests

A number of papers in the economics literature model the demand for *term* life insurance. Term insurance pays a benefit if the insured dies before a certain date. The first model for term life insurance in a continuous time setting is Yaari (1965). Fischer (1973) develops a life cycle model of term life insurance demand in discrete time and discusses the allocation of insurance purchases over the life cycle. Less common is the modeling of whole life insurance. Whole life insurance requires the build-up of insurance reserves because the insured typically pays premiums only during working life. The premiums must therefore also finance the accumulation of reserves sufficiently to meet expected later obligations. Many whole life insurance contracts enable the insured to take out those reserves (the cash value or surrender value) after a certain age, and therefore resemble a combination of term life insurance with a savings plan. Babbel and Ohtsuka (1989) build a three-period model with uncertainty about future rates of return and health status that allows for simultaneous purchase of term life insurance and whole life insurance, overcoming the problem that whole life insurance is usually dominated by a combination of term life insurance and a savings plan. However, their model is inherently difficult to solve even with sophisticated numerical methods. Moreover, Babbel and Ohtsuka do neither capture the tax preferences of life insurance nor consider the effect of public pension programs on life insurance demand.

Following the standard approach, this paper derives life insurance demand in a model with a "joy-of-giving" bequest motive (one exception is Lewis, 1989). The model has three periods and three types of assets, life insurance, bonds, and public pensions. Life insurance is modeled as a combination of term life insurance and a savings plan. Our specification incorporates the salient features of the German tax and pension system.

In the three-period model, the timing convention used is as follows: consumption streams in the three periods are indexed by 0, 1, and 2, and end-of-period bequests are indexed by 1, 2, 3, respectively. A consumer can use his income to purchase life insurance L or save an amount S of bonds. Bonds earn a rate of return r and the return is subject to a capital income tax of τ^C. Moreover, individuals must contribute to a public pension system with a payroll tax τ^S and they receive pensions in old age. The pension system has an internal rate of return of g.

More formally, consider the following expected utility function in consumption, c, and bequests, b:

$$W(c,b) = \sum_{t=0}^{2} \frac{1}{1-\gamma} \left(\frac{1}{1+\delta}\right)^t [c_t^{1-\gamma} + \eta_{t+1} b_{t+1}^{1-\gamma}(1-\pi_{t+1})] \prod_{s=1}^{t} \pi_s, \tag{2.1}$$

where δ represents the pure rate of time preference, γ is the risk aversion parameter of the constant relative risk aversion utility function, η is the weight on bequests and π_t is the probability to survive at the beginning of period t. Since death at the end of period 2 is certain, $\pi_3 = 0$.

To simplify notation, let $1+r = R$, $1+r(1-\tau^C) = R^C$, and $1+g = G$. The utility maximization is then subject to the following budget constraints in the first two periods ($t = 0, 1$):

$$c_t = w_t(1-\tau^S) - Z_t L_{t+1} - S_{t+1} + S_t R^C + \alpha L_t \tag{2.2}$$

$$b_{t+1} = S_{t+1} R + L_{t+1}. \tag{2.3}$$

Here, w stands for labor earnings. α is the exogenous savings portion of the life insurance contract – if the policy holder survives, a fraction of the insurance sum (the cash value) can be withdrawn. Note also that in case of death the estate receives the full rate of return on bonds, implicitly assuming that there are no estate taxes to be paid.

Consumers retire in their third period of life and receive a public pension. Since life ends with certainty after period 2, there is no role for life insurance in the last period. Consequently, the budget constraints are as follows:

$$c_2 = \tau^S(w_0 G^2 + w_1 G) - S_3 + S_2 R^C + \alpha L_2 \tag{2.4}$$

$$b_3 = S_3 R. \tag{2.5}$$

The first order conditions imply the following relationship between consumption in different periods and consumption and bequest for $t = 1, 2$:

$$\frac{c_t}{c_{t-1}} = \left[\frac{1 - Z_{t-1}R}{\frac{\pi_t}{1+\delta}(R^C - \alpha R)}\right]^{-\frac{1}{\gamma}} \tag{2.6}$$

and

$$\frac{c_t}{b_t} = \left[\frac{(1-\pi_t)\eta_t(1 - Z_{t-1}R)}{\frac{\pi_t}{1+\delta}(R^C Z_{t-1} - \alpha)}\right]^{-\frac{1}{\gamma}} \tag{2.7}$$

Bequests at the end of period 2 are simply:

$$b_3 = c_2 (R\eta_3)^{\frac{1}{\gamma}} \tag{2.8}$$

Using equations (2.6), (2.7) and (2.8), the consumer's maximization problem can be solved recursively. The algebraic solution is fairly complicated and therefore provides few immediate insights (see the Appendix). However, the first-order conditions offer some qualitative predictions for variations in key variables. In general, people buy life insurance for three reasons in our model: first, life insurance enhances bequeathable wealth and is therefore valuable especially at younger ages when savings are still small. Second, life insurance has a tax advantage over other savings. Third, if the consumer considers public pension coverage as too generous he can deannuitize by purchasing life insurance.[3]

Consider first the impact of tax changes on portfolio choices. Suppose two households have the same household income but differ in their tax rate on capital income τ^C. According to equations (2.6) and (2.7), the two households would differ in their consumption, bequest, and portfolio choices. As indicated by equation (2.6), a household facing a lower tax rate (higher R^C) would choose a steeper consumption profile because higher after-tax rates of returns make future consumption "cheaper." As shown in equation (2.7), that household would also choose to bequeath less than the household facing higher tax since lower taxes make future consumption cheaper but do not affect the implicit price for bequests. Equations (2.6) and (2.7) and the budget constraints also imply a different portfolio choice. For reasonable parameter choices, the household with lower tax rates can satisfy (2.6) and (2.7) simultaneously only if it holds less life insurance and more savings than the household with higher tax rates. Lowering life insurance by a dollar and increasing savings by a dollar in period 0 reduces consumption by $Z_0 - 1$ dollars. Under the assumption that insurance is fair, $Z_0 = \frac{1-\pi_1}{R} + \frac{\alpha\pi_1}{R}$, which is less than 1, the reallocation thus reduces resources in the first period. It increases resources in the following period by $R^C - \alpha$ which exceeds 1 for reasonable parameter choices.[4] Moreover, such a reallocation increases bequests by $R - 1$, which is smaller than $R^C - \alpha$ as long as $(R-1)\tau^c + \alpha$ is less than 1, which again is the case for reasonable parameter choices. In summary, reallocating a dollar from life insurance to savings lowers current resources, increases future resources, and increases future resources for

[3] Yaari (1965) discusses why in perfect markets purchasing life insurance is equivalent to purchasing a negative annuity.

[4] For example, assuming interest rates of three percent per year, an α of around 0.2 implies in a three period model that roughly 80 percent of life insurance premiums contribute to the accumulation of reserves in the first period of life.

consumption more than for bequests.

An analogous argument holds for changes in the parameter α that determines the savings content of whole life insurance. Lowering α has the same effect on first-order conditions as lowering the tax rate on capital income. Thus, quite intuitively, equations (2.6) and (2.7) together with the budget constraints also predict that lowering the implicit savings portion of life insurance leads households to shift more resources away from life insurance and towards regular savings.

As equation (2.7) demonstrates, increasing the strength of bequest motives leads to the result that the relative size of bequest to consumption must increase, while the relative size of consumption in different periods remains constant according to equation (2.6). Clearly, the less costly way to increase bequests is to purchase more life insurance. However, unlike the previous results, it depends on specific parameter values whether both saving and life insurance increase or whether life insurance demand increases and savings falls.

Finally, varying the size of the public pension system also matters for both saving and life insurance. As is well-known, public pensions crowd out private savings in a life-cycle model. To the extent that life insurance is a savings instrument, one would therefore expect life insurance demand to fall. However, for people who feel that the public pension is too generous, purchasing more life insurance is a way to increase bequests and reduce the "overannuitization". Thus, the precise effect of public pension coverage on life insurance demand depends on the relative magnitude of the savings and bequest motives.

To summarize, the stylized life-cycle model presented in this section delivers two main testable predictions regarding life-insurance demand. First, controlling for income, people facing lower relative tax rates on other savings should purchase less whole life insurance to accommodate a steeper consumption profile. Second, people with stronger bequest motives, for example married people or households with children, should have stronger incentives to purchase life insurance. The impact of public pension coverage on life insurance demand is ambiguous.

2.3 The Data

The German Socioeconomic Panel (GSOEP) offers a unique opportunity to study the effect of tax reform and bequest motives on the demand for life insurance. It is the only dataset containing annual information about life insurance ownership and portfolio choice of German households that spans from pre- to post-reform years. And, it allows studying portfolios in the territory of the GDR, where a survey containing around 2000 households started in 1990 prior to reunification. The first survey in the West was conducted in 1984. Since then, the sample has been significantly increased in 1998 and 2000. Descriptive evidence for the development of the sample is provided in table 2.1.[5]

Households are asked annually if they owned one or more life insurance policies in the previous year. Thus, we only use observations for households that take part in two successive surveys. If not otherwise stated, socioeconomic characteristics are proxied by the household head. We approximate marginal tax rates by re-calculating each household's taxable income from these (estimated) tax payments, using the official formulas of the federal tax office.[6] A 1 unit change in taxable income is simulated in order to approximate the marginal tax rate.

2.4 Tax Incentives

Life insurance is the second most common asset after savings accounts in Germany. In 2007, 15.6 percent of total private wealth, amounting to 716 billion Euro, was allocated to life insurance (Deutsche Bundesbank, 2008). Overall, 93.9 million life insurance policies existed, of which 7,617,400 had been sold in that year (Gesamtverband der Deutschen Versicherungswirtschaft, 2008). On the flipside of the market, around 49 percent of households own life insurance policies.[7]

One of the main reasons for this unusually high popularity in Germany is the favored fiscal treatment of life insurance policies (and whole life insurance in particular). Firstly, returns on life insurances are tax exempt if the contract lasts for at least 12 years, premiums are paid during at least five years, and the term life insurance component amounts to at least 60 percent of the total benefit paid out at the end of the contract. Secondly, annual contributions to whole life and term life insurance contracts are tax deductible. However, this is typically of little benefit for employees, as they reach the deductibility cap with their obligatory contributions to the social security

[5]The Add-On package PanelWhiz for Stata (http://www.PanelWhiz.eu) has been used for extracting the data. See Haisken-DeNew and Hahn (2006) for details. The PanelWhiz generated DO file to retrieve the data used here is available upon request. Any data or computational errors are our own.

[6]The GSOEP estimates of total tax payments are based on Schwarze's (1995) approach. Schwarze adds up the incomes of all household members and applies standard deductions based on the socioeconomic status of the household.

[7]Authors' calculation based on data from the GSOEP. Typically, life insurance policies have one beneficiary, so that it makes sense for households with several children to invest into several policies.

system. Obligatory contributions are smaller for civil servants and the deductibility cap is higher for the self-employed, who are generally exempt from contributing to the public pension system and must provide for their own retirement income and survivor's benefits (Sommer, 2007). Finally, in the case of bequests, only two-thirds of the cash value of life insurance policies are taxed. It is even possible to avoid estate taxes altogether if, for instance, a husband pays premiums into a life insurance policy owned by his wife, who also is the beneficiary if he dies early.

The 2000 tax reform had a major impact on the treatment of whole life insurance, and it had a visible effect on whole life insurance sales trends. Figure 2.1 plots the development of sales of new life insurance contracts between 1995 and 2003 in Germany. During the entire period, sales of term life insurance policies are relatively constant around 700,000. However, sales of new whole life contracts spike in 1999, indicating an anticipation effect of the tax reform. As taxpayer groups have been affected differently by the reforms, in this paper we can clearly identify the response of households' savings allocation to changes in after-tax yields.

2.4.1 The Tax Reform in 2000

Germany taxes all interest and dividend income exceeding a certain threshold at the households' marginal tax rate. The development of this threshold, the so-called *Sparerfreibetrag* (tax exemption limit), is shown in table 2.1 for the period from 1996 and 2001. In March 1999, a law was passed, cutting the tax exemption limit from DM6,000 (12,000) to DM3,000 (6,0000) for singles (couples) from January 1, 2000 onward. We suspect that households between the old and the new tax exemption limit were disproportionately affected by this reform. As their capital returns were fully exempt from taxation beforehand, the reform created a strong incentive to shelter their savings from taxation by purchasing whole life insurance when the reform was announced. In other words, if these households are responsive to the relative tax treatment, we should see a disproportionate increase of life insurance purchases among the group threatened to have their regular savings income taxed by the reduction of tax exemptions. In what follows, we denote households belonging to this category as the "treatment group".

In order to identify the treatment group, we use survey responses on capital income levels. One quarter of all households report their exact income from interest and dividends in the survey, whereas three quarters indicate on an ordinal scale if their capital income is less than 500, between 500-2,000, 2,000-5,000, 5,000-10,000, or above 10,000 DM. These ordinal thresholds reduce precision of estimating the response to tax changes, biasing the results *against* finding significant differences between groups. We use a difference-in-differences estimator to test if the treatment group is more likely to own one or more (tax-exempt) life insurance policies from 2000 onward.

Our empirical analysis is subject to two additional considerations. First, in June 1999, the government proposed to abolish the tax exemption on life insurance returns by end 1999. Many households were concerned about losing a tax-favored savings opportunity, boosting sales of new contracts by 38.7 percent in 1999. Dolle-Helms (1999a, 199b) provides anecdotal evidence that last minute purchases in 1999 were mainly driven by tax motives. The reform eventually failed in the upper house of parliament (*Bundesrat*) in mid-December and many investors (unsuccessfully) claimed their money back. All households, including those above the new exemption limit were also potentially affected by the proposed and later dropped reform in 1999. To identify the impact of the 2000 reform that affected the relative tax treatment of life insurance and other savings for people below the DM6,000 limit we test if, from 1999 onward, the ownership probability among the treatment group increases relative to wealthier households whose capital income was already above the DM6,000 exemption limit. The implicit underlying identifying assumption is that both the treatment and control group responded equally (in proportional terms) to the announced phasing out of tax advantages for life insurance. Second, the 2000 reform may also have had an impact on those households already paying taxes on capital income since their total tax exempt amounts would fall. However, these households would not be at the margin of having to start paying capital income taxes. If tax incentives matter, households with high capital incomes should already have invested into life insurance before the reform in order to shelter their savings from taxation, and the response would be intra-marginal.

Descriptive evidence in table 2.2 confirms that indeed significant changes only occurred in the treatment group. Life insurance ownership rates remained constant among households below the new tax exemption limit and above the old exemption limit. However, the ownership rate increased strongly from 62.5 to 69.7 percent in the treatment group in 1999. This shows that households affected by the tax reform in 2000 advanced their investments and stocked up on (tax-exempt) life insurance policies before the reform came into effect.

2.4.2 Empirical Results

We estimate a reduced-form model in order to analyze the effect of tax reform on life insurance demand. In particular, a before and after comparison is made between a control group of investors that are unaffected by the reform with a treatment group that is affected by the new tax regime, using a difference-in-differences estimator on repeated cross-sectional data. We denote individual i's binary indicator for the treatment group as G_i. For the reform in 2000, the treated are defined as $G_{it} = \mathbb{1}\{limit^{new} \leq INC_{it}^{CAP} \leq limit^{old}\}$, where INC_{it}^{CAP} denotes total capital income. $T_i = \mathbb{1}\{t \geq 1999\}$ is a time dummy indicating the anticipated reform. To ease the notational burden, we introduce the shorthand $Y_{i\in g,t}$ for $Y_i|G_i = g, T_i = t$. The potential outcomes with and without treatment are Y_i^1 and Y_i^0 respectively. The model for the outcome without intervention is given by

$$Y_i^0 = \alpha T_i + \beta G_i + \epsilon_i,$$

where $\epsilon_i \perp (T_i, G_i)$. The model for the treatment group is

$$Y_i^1 = \alpha + \beta + \tau^{DiD} + \epsilon_i,$$

In the absence of intervention, the average outcome for the treatment group is $E[Y_{i\in 1,1}^0] = E[Y_{i\in 1,0}] + E[Y_{i\in 0,1}] - E[Y_{i\in 0,0}]$. The average treatment effect on the treated is defined as

$$\begin{aligned} \tau^{DiD} &= E[Y_{i\in 1,1}^1] - E[Y_{i\in 1,1}^0] \\ &= E[Y_{i\in 1,1}] - E[Y_{i\in 1,0}] - (E[Y_{i\in 0,1}] - E[Y_{i\in 0,0}]). \end{aligned} \tag{2.9}$$

This estimator requires three identifying assumptions. First, we assume that the tax reform is exogenous to the ownership decision. Investors were hit by surprise, when the tax reforms were announced in 1999, since the reforms were not mentioned in election campaigns or the coalition program of the incoming government. We can also safely exclude the possibility of policy endogeneity, because the reform was not introduced to change the demand for life insurance by different taxpayer groups. It was part of a major tax reform package with the aim of broadening the tax base. Second, we assume that there are no group specific trends in life insurance ownership. This assumption guarantees that the counterfactual of the treated can be inferred from the time trend of the control group. As discussed above, this assumption certainly holds for households above the new exemption limit. Third, we assume that the tax reform is exogenous with respect to sample composition. Essentially, this requires that household income as well as interest and dividend income did not change as a result of the tax reform itself. One caveat could be that interest and dividend income falls when a household buys life insurance. However, households typically pay annual premiums of less than DM2,000 (Sommer, 2007) which would only marginally affect total capital income at reasonable interest rates. Also, the presence of such wealth effects would rather bias the results against our findings.

The upper panel of table 2.3 reports the average effects of the tax reform, using a sample from three years before and three years after the announcement of the reform. While ownership rates of life insurance fell among households in the control group from 1999 onwards, an increase by 4.5 percent can be observed for the treatment group. The difference-in-differences estimate according to equation 2.9 is 5.7 percent for the full sample and 5.9 percent for households above the new exemption limit.[8]

These estimates may be biased for two reasons. First, the estimated probabilities of investing into life insurance do not necessarily lie in the $[0, 1]$ interval. Second, the effects could be blurred because other determinants account for different behavior across groups. Thus, we translate the difference-in-differences approach into a probit regression that imposes bounds on the estimated probabilities and accounts for other covariates.

The model for the outcome without intervention is given by

$$P(Y_i^0 = 1|G_i, T_i, \mathbf{x_i}) = \Phi(\alpha T_i + \beta G_i + \mathbf{x_i}\delta).$$

The model for the treatment group is

$$P(Y_{i\in 1,1}^1 = 1|\mathbf{x_i}) = \Phi(\alpha + \beta + \gamma + \mathbf{x_i}\delta),$$

Puhani (2008) shows that in a nonlinear model, such as probit, the treatment effect on the treated should not be confused with the cross-derivative of the interaction term (Ai and Norton, 2003). Based on the standard probit difference-in-differences model

$$P(Y_i = 1|G_i, T_i, \mathbf{x_i}) = \Phi(\alpha T_i + \beta G_i + \gamma T_i G_i + \mathbf{x_i}\delta),$$

a consistent estimator of the treatment effect is

$$\begin{aligned} \hat{\tau}^{DiD} &= E[Y_{i\in 1,1}^1|\mathbf{x_i}] - E[Y_{i\in 1,1}^0|\mathbf{x_i}] \\ &= \frac{1}{N}\sum_{i=1}^{N}(\Phi(\hat{\alpha} + \hat{\beta} + \hat{\gamma} + \mathbf{x_i}\hat{\delta}) - \Phi(\hat{\alpha} + \hat{\beta} + \mathbf{x_i}\hat{\delta})). \end{aligned} \tag{2.10}$$

Hence, the treatment effect is zero if and only if the coefficient γ is zero. We apply the delta-method to infer statistical significance of the average treatment effect in small samples. Different from the linear model, identification is not provided by the assumption that the cross difference γ is zero for the expected potential outcome Y_i^0,

[8]Estimates are similar, if all observations for the transition year 1999 are dropped.

because group and time differences in the conditional expectation of the potential outcome Y_i^0 are not constant in the nonlinear probit model. However, a nonlinear parametric restriction on that cross-difference guarantees that all expected outcomes (factual or counterfactual) are bounded as required (Athey and Imbens, 2006).

Table 2.4 reports summary statistics of the additional covariates included in the regression. In particular, we include the marginal tax rate to control for differences in after-tax returns. We proxy for the household's net labor income via binary indicators for deciles of the income distribution. Dummies for house ownership as well as interest and dividend returns control for household wealth. Furthermore, martial status and a binary indicator for households with one or more children, as in Hurd (1987, 1989), capture bequest motives. Dummies for emplyoment status, civil servants and the self-employed reflect specific characteristics of the German tax and public pension system. Finally, the model includes gender, education, and linear and non-linear terms of the age of the household head. We use data for three years before and after the reform. The full sample consists of 44,540 observations and 2,419 if we constrain the analysis to households above the new exemption limit.

Table 2.5 reports average marginal effects for continuous and dummy variables. The interaction effect $\hat{\tau}^{DiD}$ is statistically significant at the 5 percent level in both equations. According to equation 2.10, the tax reform increased ownership among households affected by the reform by 5.2 percent. The estimate is 8.9 percent for the restricted sample in column (2). Furthermore, the estimates of the marginal effects in columns (1) and (2) show a highly significant positive correlation between marginal tax rates and investment into life insurance. The model in column (1) suggests that an increase of the marginal tax rate by 10 percentage points increases the ownership probability by 3.3 percentage points. Also, there is evidence that the self-employed, who have larger tax incentives and lower pensions, are more likely to own life insurance policies. We find no evidence that life insurance ownership is higher among civil servants, who typically receive relatively generous survivor benefits.

However, the picture is less clear with regard to bequest motives. Although there is strong evidence that married couples invest into life insurance, we cannot confirm that households with children are more likely to own life insurance. These mixed results are much in line with the previous literature that finds evidence in both directions. One reason for these ambiguous results stems from the impossibility of past studies to neatly single out bequest motives from other forms of savings motives. For instance, we would be unable to identify bequest motives in the presence of strong tax incentives if households with children were liquidity constrained in the short run.

2.5 Bequest Motives

Our identification strategy for bequest motives in life insurance demand relies crucially on the assumption that the specific institutional environment in the GDR allows us to control for the main alternative savings motives discussed in the literature such as tax, life-cycle, and precautionary savings motives (Browning and Lusardi, 1996). The following section describes the peculiarities of the market for life insurance in the GDR. Throughout the section, bequest motives are defined as a willful desire to hand on one's wealth to close relatives or friends during lifetime or posthumously.

2.5.1 Savings Environment in the GDR

Whole life insurance played an important role in household portfolios in the GDR. Before reunification, East Germans could only chose between investing into savings accounts or life insurance. Therefore, life insurance was typically considered as a long-term savings contract (with an additional term life insurance option). While savings accounts offered a unitary interest rate of 3.25 percent (Schwarzer, 1999), calculatory returns on life insurance were about 3.5 percent plus a 15 percent terminal bonus. The only provider of insurance was the *Staatliche Versicherung der DDR*, of which Allianz acquired the private client business after reunification.[9]

Tax incentives. A major advantage of studying life insurance demand in the GDR is that returns on savings accounts and insurances were fully exempt from taxation. Also, the beneficiary of a life insurance policy was exempt from death taxes (Schulze, 1970). Different from all existing studies on bequest motives, our analysis will therefore not be diluted by tax considerations.

Life-cycle savings motives. Another key feature of savings decisions in the GDR is that consumption possibilities were very limited, enabling us to control for ownership of all goods and services for which GDR citizens needed to accumulated large deposits. In particular, we can control for the five main life-cycle and down-payment motives: First, we condition for life-cycle saving effects through linear and nonlinear terms of age as well as an indicator for the retirement status of the household head. Around 40 percent of all retirees also participated in an additional retirement pension supplement plan. However, average pensions were about 450 (550 with the supplement) Mark in

[9] We thank Dr. Michael Lehner from Allianz for providing detailed information about the life insurance market in the GDR.

1986, compared to an average labor income of 960 Mark (Dabbert, 1992).[10] Hence, the elderly had to rely on their savings for a sufficient retirement income. Second, households had little incentive to use life insurance to accumulate deposits for buying an own apartment or house. The communist government restricted ownership of private property and largely subsidized construction of rental housing. Also, it was very cheap to live in a rented apartment, since rents were fixed by the central government and too low to recover maintenance cost (Manzel, 1992). Third, only few durables required large downpayments. There is anecdotal evidence that life insurance contracts were used to buy cars. This was a sensible thing to do, because the average duration of life insurance policies, 11.6 years, matched the average delivery time for a car, 13.5 years (Wolle, 1999).[11] The only other durables for which large deposits were necessary are motorcycles and weekend houses (*Datschas*). The data allow us to control for these three durables when estimating the strength of bequest motives. Fourth, we rule out the possibility that citizens used life insurance as a means to save for travel. The duration of life insurance policies does typically not match the decision to travel. Moreover, travel restrictions were not lifted before the 1970s and even then GDR citizens could only travel to four foreign countries without a visa (Czechoslovakia, Hungary, Poland, Bulgaria). Often, travelling was further complicated by scarce foreign exchange (Saretzki and Kohn, 1992). Other leisure activities did typically not require larger amounts of money. Yet some social activities may affect mortality risk (sports) or provide an information network that increases the awareness for life insurance products. The regression includes indicators for households that go at least once a month to the cinema or theater, a cultural event, the church, do active sports, visit friends, or help neighbors. Fifth, private spending on education did not exist under the communist regime in the GDR. The government fully funded primary and higher education as well as vocational training (Marggraf, 1992). Access to higher education required membership in the GDR youth organization (*FDJ*), and favored entry for children from working class backgrounds.

Precautionary savings motives. Economists disagree sharply as to why people bequeath wealth. In contrast to the view that bequests are intentional, Hurd (1987) suggests that bequest are only an accidental remnant of precautionary savings. Yet the social system in the GDR gave very few reasons to accumulate wealth as a buffer for uncertain times. We believe that our regression captures the remaining precaution-

[10]The Deutsche Mark (DM) should not be confused with the Mark which was the official currency of the GDR. Mark (East) were exchanged 1:1 for Deutsche Mark (DM) in 1990. However, the cash value of life insurance policies and savings above 4,000 Mark were exchanged 2:1.

[11]A fashionable nickname for life insurance used to be *Trabi-Sparvertrag* (savings contract for a *Trabant* ("Trabi") car).

ary motives and therefore yields unbiased estimates of bequest motives. First, East Germans did not have to hedge against income fluctuations, because full employment was constitutionally guaranteed. Fuchs-Schündeln and Schündeln (2005) also argue that income differences and volatility were very low. Second, we can control for the self-assessed health status of each individual, which could affect precautionary savings, although health services were fully provided by the central government.[12] Third, people have been asked if they are dissatisfied or very dissatisfied with the social benefits available in the GDR. Dissatisfaction could denote a larger demand for precautionary savings. Finally, the data provide a proxy for individual risk preferences. People indicate on a 0-10 scale if they consider it to be desirable for one to be security-conscious (0-10 scale).

Bequest motives. If bequests are intentional, they may either reflect altruism (Tomes, 1981), self-interested exchange with one's heirs (Bernheim, Shleifer, Summers, 1985), or the outcome of an intra-household reallocation of incomes (Gandolfi and Miners, 1996). Gandolfi and Miners argue that families insure the labor income of the main bread-earner through life insurance. We proxy for potential reallocation motives by the wife's labor force participation status and the income differential between husband and wife. Like Hurd (1987, 1989) and Jürges (2001), we use a dummy indicating if a household has one or more children in order to proxy for altruistic and strategic motives. The questionnaire also asks the household head if his family is very important to his sense of well-being and personal satisfaction. However, it difficult to differentiate altruistic from strategic motives, since the survey does not ask for the intention of households' bequests.

2.5.2 Empirical Results

Age profiles for life insurance ownership rates in the GDR are depicted in figure 2.2. Note that age and cohort effects cannot be separately identified, as we only use a single cross-section of data in this section. Ownership rates display a hump shape that is broadly consistent with life-cycle insurance demand as derived from the model in section 2.2. Life insurance ownership peaks between ages 20 to 40, while in older ages households cash out their insurance policies.

Descriptive evidence for the presence of bequest motives in GDR life insurance demand is presented in table 2.6. Insurance ownership is clearly higher among married couples, households with children, households with higher labor incomes, civil servants,

[12]The questionnaire reads: "How satisfied are you with your health?" (0-10 scale).

and house owners. Education and wealth seem to play a minor role. The descriptive evidence is hard to reconcile with the notion of intra-household reallocation motives. While ownership rates are higher for households with lager income differences, the contrary holds if the partner is not participating in the labor force.

We estimate probit models for the ownership probability in 1990. Table 2.7 reports average marginal effects for continuous and dummy variables. We do not report separate average partial effects for the hump shaped effects of age and income on the ownership decision. All specifications provide robust evidence for the presence of bequest motives among households with children. Column (4) reports that the probability to own one or more life insurance policies is independent of the number of children within a household. On average, households with children are 7 percentage points more likely to own life insurance. Column (5) also controls for the attitudes of investors. Only investors for whom family is very important show a significantly higher participation probability of around seven percentage points. No significant correlation can be identified between insurance demand and attitudes such as security-consciousness, importance of social security, or self-assessed health status. We also control for different indicators of leisure activities in column (6), which reflect social interaction effects (Hong, Kubik, Stein, 2004). However, these indicators are neither individually nor jointly significant.

Because many households cash out their whole life insurance policies at retirement, we test whether that fact has a significant impact on regression results. Column (7) presents estimates for a subsample of households with a head of less than 65 years of age. In this subsample, age effects are insignificant indicating that age profiles are similar among the working population. The previous findings are broadly confirmed in this smaller sample. Estimates of potential bequest motives are statistically significant and of similar size, as in the full sample. There is no indication that investors insure the labor income of the main earner.

2.6 Conclusion

Whole life insurance plays an important role in household saving. In a stylized model both tax incentives and bequest motives drive whole life insurance demand. While a bequest motive could be satisfied by term life insurance, sheltering savings from capital income taxation is only possible with whole life policies. The empirical evidence presented is consistent with the theoretical predictions. In particular, we study two natural experiments in Germany. A tax reform in 2000 halved the tax exemption limit for capital income. We find that the demand for life insurance increases particularly among that group of households, which did not pay taxes on capital returns prior to the reform.

Our results contrast to Jappelli and Pistaferri (2003, 2007) who do not find that tax incentives matter for life insurance demand in Italy. However, anecdotal evidence tells that sales agents exploited the favorable market situation in Germany, whereas Italian insurers lacked the vital initiative to point out the effects of the tax reform among potential investors. Moreover, the tax incentive resulting from incremental changes in after-tax yields in Italy might be too small to induce significant changes in investment behavior if inertia are present. The specific features of the German reform which establish a natural experiment and the richness of our data provide a truly unique opportunity to show that increases in capital income taxation induce a shift of the portfolio towards tax-exempt assets. The results suggest that standard tax revenue estimates, which assume that current investors would stick to their asset choices if capital taxation were introduced, may be misleading. Governments need to account for changes in investment behavior due to tax reforms (Poterba and Verdugo, 2008).

With regard to bequest motives, we analyze the demand for life insurance in the experimental setting of the GDR, where our estimates are not diluted by tax considerations or life-cycle and precautionary savings motives. We find a significantly higher ownership probability among households with children and a high regard for the family. Life insurance demand does not seem to depend on intra-household allocation motives. As a note of caution, we admit that our results are based on are very peculiar institutional setting. Yet in contrast to our expectations and in favor of a broader applicability of our findings, GDR life insurance demand demonstrates the importance of bequest motives despite the omnipresence of a paternalist communist state.

References

Ai, Chunrong and Edward Norton (2003): "Interaction Terms in Logit and Probit Models," *Economics Letters*, **80**, 123-129.

Alan, Sule and Søren Leth-Petersen (2006): "Tax Incentives and Household Portfolios: A Panel Data Analysis," CAM Working Papers 2006-13, University of Copenhagen.

Alan, Sule; Kadir Atalaya; Thomas F. Crossley and Sung-Hee Jeon (2008): "New Evidence on Taxes and Portfolio Choice," Unpublished manuscript, University of Cambridge.

Alessie, Rob; Stefan Hochguertel and Arthur van Soest (1997): "Saving Accounts versus Stocks and Bonds in Household Portfolio Allocation," *Scandinavian Journal of Economics*, **99**, 81-97.

Agell, Jonas and Per-Anders Edin (1990): "Marginal Taxes and the Asset Portfolios of Swedish Households," *Scandinavian Journal of Economics*, **9**, 47-64.

Babbel, David F. and Eisaku Ohtsuka (1989): "Aspects of Optimal Multiperiod Life Insurance," *Journal of Risk and Insurance*, **56**, 460-481.

Banks, James and Sarah Tanner (2002): "Household Portfolios in the UK," in *Household Portfolios*, (eds.) Luigi Guiso, Michael Haliassos, and Tullio Jappelli. MIT Press: Cambridge, 219-250.

Beck, Thorsten and Ian Webb (2003): "Economic, Demographic, and Institutional Determinants of Life Insurance Demand across Countries," *The World Bank Economic Review*, **17**, 51-88.

Bernheim, Douglas; Andrei Shleifer and Lawrence H. Summers (1985): "The Strategic Bequest Motive," *Journal of Political Economy*, **93**, 1045-1076.

Bernheim, Douglas (1991): "How Strong Are Bequest Motives? Evidence Based on Estimates of the Demand for Life Insurance and Annuities," *Journal of Political Economy*, **99**, 899-927.

Browne, Mark J. and Kihong Kim (1993): "An International Analysis of Life Insurance Demand," *Journal of Risk and Insurance*, **60**, 460-481.

Browning, Martin and Annamaria Lusardi (1996): "Household Saving: Micro Theories and Micro Facts," *Journal of Economic Literature*, **34**, 1797-1855.

Dabbert, Edith (1992): "Von 280 Mark zu 11830 Mark - 40 Jahre Sparen," in *Im*

Trabi durch die Zeit - 40 Jahre Leben in der DDR, (ed.) Egon Hölder. Metzler-Poeschel: Stuttgart, 121-127.

DEUTSCHE BUNDESBANK (2008): *Ergebnisse der gesamtwirtschaftlichen Finanzierungsrechnung für Deutschland 1991 bis 2007, Statistische Sonderveröffentlichung 4*. Deutsche Bundesbank: Frankfurt.

DOLLE-HELMS, ELKE (1999a): "Geschäft des Lebens," *Die Zeit*, **36**, 28.

DOLLE-HELMS, ELKE (1999b): "Voreilig abgeschlossen?," *Die Zeit*, **52**, 32.

FISCHER, STANLEY (1973): "A Life Cycle Model of Life Insurance Purchases," *International Economic Review*, **14**, 132-152.

FUCHS-SCHÜNDELN, NICOLA and MATTHIAS SCHÜNDELN (2005): "Precautionary Savings and Self-Selection: Evidence from the German Reunification Experiment," *Quarterly Journal of Economics*, **120**, 1085-1120.

GANDOLFI, ANNA SACHKO and LAURENCE MINERS (1996): "Gender-Based Differences in Life Insurance Ownership," *The Journal of Risk and Insurance*, **63**, 683-693.

GESAMTVERBAND DER DEUTSCHEN VERSICHERUNGSWIRTSCHAFT (2008): *Geschäftsentwicklung 2007. Die deutsche Lebensversicherung in Zahlen.* GDV: Berlin.

HAISKEN-DENEW, JOHN P. and MARKUS HAHN (2006): "PanelWhiz: A Flexible Modularized Stata Interface for Accessing Large Scale Panel Data Sets," mimeo, (http:// www.PanelWhiz.eu).

HONG, HARRISON; JEFFREY D. KUBIK and JEREMY C. STEIN (2004): "Social Interaction and Stock-Market Participation," *Journal of Finance*, **59**, 137-163.

HURD, MICHAEL D. (1987): "Savings of the Elderly and Desired Bequests," *American Economic Review*, **77**, 298-312.

HURD, MICHAEL D. (1989): "Mortality Risk and Bequests," *Econometrica*, **57**, 779-813.

JAPPELLI, TULLIO and LUIGI PISTAFERRI (2003): "Tax Incentives and the Demand for Life Insurance," *Journal of Public Economics*, **87**, 1779-1799.

JAPPELLI, TULLIO and LUIGI PISTAFERRI (2007): "Do People Respond to Tax Incentives? An Analysis of the Italian Reform of the Deductibility of Home Mortgage Interest ," *European Economic Review*, **51**, 247-273.

JÜRGES, HENDRIK (2001): "Do Germans Save to Leave an Estate? An Examination of the Bequest Motive," *The Scandinavian Journal of Economics*, **103**, 391-414.

King, Mervyn A. and Jonathan I. Leape (1998): "Wealth and Portfolio Composition: Theory and Evidence," *Journal of Public Economics*, **69**, 155-193.

Kopczuk, Wojciech and Joseph P. Lupton (2007): "To Leave or Not to Leave: The Distribution of Bequest Motives," *Review of Economic Studies*, **74**, 207-235.

Laitner, John and Thomas Juster (1996): "New Evidence on Altruism: A Study of TIAA-CREF Retirees," *American Economic Review*, **86**, 893-908.

Lewis, Frank D. (1989): "Dependents and the Demand for Life Insurance," *American Economic Review*, **79**, 452-467.

Manzel, Karl-Heinz (1992): "Von der Wohnlaube zum Wohnblock - Ziel der 'registrierten Antragstellung'," in *Im Trabi durch die Zeit - 40 Jahre Leben in der DDR*, (ed.) Egon Hölder. Metzler-Poeschel: Stuttgart, 251-264.

Marggraf, Hans-Jörg (1992): "Von der 'Demokratisierung der deutschen Schule' zum sozialistischen Bildungssystem - der Weg vom Kindergarten zum Beruf," in *Im Trabi durch die Zeit - 40 Jahre Leben in der DDR*, (ed.) Egon Hölder. Metzler-Poeschel: Stuttgart, 49-66.

Poterba, James M. (2002): "Taxation and Portfolio Structure: Issues and Implications," in *Household Portfolios*, (eds.) Luigi Guiso, Michael Haliassos, and Tullio Jappelli. MIT Press: Cambridge, 103-142.

Poterba, James M. and Andrew A. Samwick (2003): "Taxation and Household Portfolio Composition: U.S. Evidence from the 1980s and 1990s," *Journal of Public Economics*, **87**, 5-38.

Poterba, James and Ramirez Verdugo (2008): "Portfolio Substitution and the Revenue Cost of Exempting State and Local Government Interest Payments from Federal Income Tax," Working Paper, MIT.

Puhani, Patrick A. (2008): "The Treatment Effect, the Cross Difference, and the Interaction Term in Nonlinear "Difference in Differences" Models," IZA Discussion Paper No. 3478.

Saretzki, Hans-Ulrich and Ursula Kohn (1992): "Vom gewerkschaftlich organisierten Urlaub zum begrenzten Tourismus - Reisen als Beitrag zur Lebensqualität," in *Im Trabi durch die Zeit - 40 Jahre Leben in der DDR*, (ed.) Egon Hölder. Metzler-Poeschel: Stuttgart, 329-341.

Scholz, John K. (1994): "Portfolio Choice and Tax Progressivity: Evidence from the Surveys of Consumer Finances," in *Tax Progressivity and Income Inequality*, (ed.) Joel Slemrod. New York: Cambridge University Press, 219-267.

SCHULZE, SIEGFRIED (1970): *Versicherungsrecht.* Staatsverlag der DDR: Berlin.

SCHWARZE, JOHANNES (1995): "Simulating German Income and Social Security Tax Payments Using the GSOEP," Cross-National Studies in Aging Program Project Paper 19.

SCHWARZER, OSKAR (1999): *Sozialistische Zentralplan-Wirtschaft in der SBZ/DDR. Ergebnisse eines ordnugspolitischen Experiments (1945-1989).* Franz Steiner Verlag: Stuttgart.

SOMMER, MATTHIAS (2007): "Savings Motives and the Effectiveness of Tax Incentives - an Analysis based on the Demand for Life-Insurance in Germany," MEA Discussion Paper 99-28, University of Mannheim.

TOMES, NIGEL (1981): "The Family, Inheritance, and the Intergenerational Transmission of Inequality," *Journal of Political Economy*, **89**, 928-958.

WOLLE, STEFAN (1999): *Die heile Welt der Diktatur. Alltag und Herrschaft in der DDR 1971-1989.* Bundeszentrale für politische Bildung: Bonn.

YAARI, MENAHEM E. (1965): "Uncertain Lifetime, Life Insurance and the Theory of the Consumer," *Review of Economic Studies*, **32**, 137-150.

Technical Appendix

The solution for first period consumption c_0 can be derived as follows:

$$
\begin{aligned}
c_0 = & \\
& \left[\left(1 + \frac{1}{R}(R\eta_3)^{\frac{1}{\gamma}}\right)\left[\frac{(R^C - R\alpha)^2 \frac{\pi_1\pi_2}{1+\delta}}{(1-Z_0R)(1-Z_1R)}\right]^{\frac{1}{\gamma}} + \right. \\
& \left(\frac{R^C - \alpha R}{1 - Z_1 R} Z_1 - \alpha\right)\left(\frac{(1-\pi_2)\eta_2 \frac{\pi_1}{1+\delta}(R^C - \alpha R)^2}{(Z_1 R^C - \alpha)(1 - Z_0 R)}\right)^{\frac{1}{\gamma}} + \\
& \left(\frac{R^C - \alpha R}{1 - Z_1 R}\right)\left(\left(\frac{(R^C - \alpha R)\frac{\pi_1}{1+\delta}}{1 - Z_0 R}\right)^{\frac{1}{\gamma}} + \right. \\
& \left.\left(\frac{R^C - \alpha R}{1 - Z_1 R} Z_1 - \alpha\right)\left(\frac{(1-\pi_1)\eta_1 (R^C - \alpha R)}{Z_0 R^C - \alpha}\right)^{\frac{1}{\gamma}}\right) + \\
& \left.\frac{(R^C - \alpha R)^2}{(1 - Z_0 R)(1 - Z_1 R)}\right]^{-1} \times \\
& \left[\frac{(R^C - \alpha R)^2}{(1 - Z_0 R)(1 - Z_1 R)} w_0 (1 - \tau^S) + \right. \\
& \left.\frac{R^C - \alpha R}{1 - Z_1 R} w_1 (1 - \tau^S) + \tau^S (w_0 G^2 + w_1 G)\right]
\end{aligned}
$$

The solution for c_0 in combination with equations (2.6), (2.7) and (2.8) immediately implies values for c_1, c_2, b_1, b_2, b_3 and thus, by applying the budget constraints, also for L_1 and L_2.

Figures and Tables

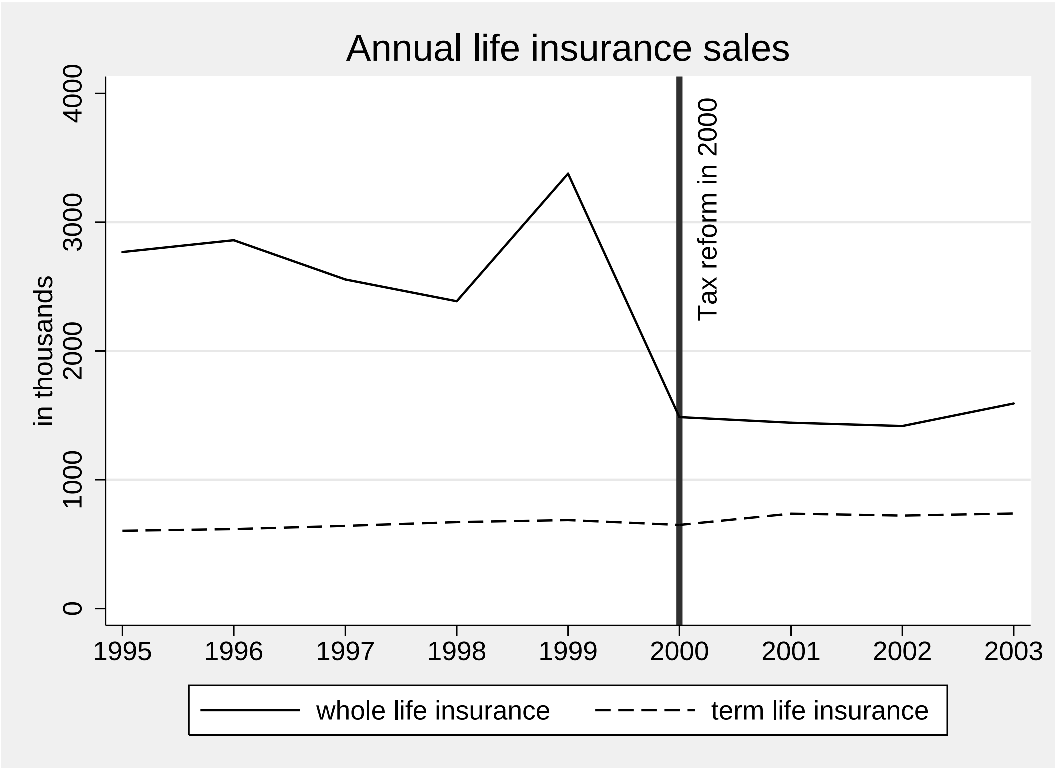

Figure 2.1: The graph depicts sales of new life insurance contracts in Germany, 1996-2001. Source: Gesamtverband der Deutschen Versicherungswirtschaft (2008).

Table 2.1: Tax exemption limits on capital income

Period	Exemption limit on capital income	Treatment group [exact]	Treatment group [categorical]
		$limit^{new}$-$limit^{old}$	$\widetilde{limit}^{new}$-$\widetilde{limit}^{old}$
1996-1999			
single	6000 DM	3,000-6,000 DM	2,000-5,000 DM
(couple)	(12,000)	(6,000-12,000)	(5,000-10,000)
2000-2001			
single	3000 DM	3,000-6,000 DM	2,000-5,000 DM
(couple)	(12,000)	(6,000-12,000)	(5,000-10,000)

Note: The table reports the development of tax exemption limits on capital income in Germany for singles (married couples). The thresholds for the old and new exemption limits, $limit^{old}$ and $limit^{new}$, define the upper and lower bounds of the treatment group and are either assigned by exact or categorical (indicated by tilde) interest and dividend returns.

Table 2.2: Tax incentives - average ownership rates 1996-2001

as a % of all observations in the subpopulation	1996	1997	1998	1999	2000	2001
Full sample	54.72	55.79	54.56	55.29	54.67	52.38
N	*6,594*	*6,383*	*7,159*	*6,980*	*11,662*	*11,193*
$INC^{CAP} < limit^{new}$	54.28	55.52	54.20	54.48	54.09	51.85
N	*6,278*	*6,092*	*6,816*	*6,533*	*10,959*	*10,703*
$INC^{CAP} > limit^{old}$	64.84	67.71	60.36	62.58	63.11	62.29
N	*91*	*96*	*111*	*163*	*225*	*175*
$limit^{new} < INC^{CAP} < limit^{old}$	62.67	58.46	62.50	69.72	64.02	65.08
N	*225*	*195*	*232*	*284*	*478*	*315*

Note: The table reports average ownership rates of life insurance policies for different subpopulations. INC^{CAP} denotes total capital income.

Table 2.3: Tax incentives - difference-in-differences

	treated	non-treated	Difference between groups	N
Effect of the tax reform.				
N	1,729	47,961		49,690
After the reform (1999-2001)	0.658	0.536	0.123	29,554
	(0.014)	(0.003)	(0.015)	
Before the reform (1996-1999)	0.613	0.548	0.066	20,136
	(0.019)	(0.004)	(0.020)	
Difference within groups	0.045	-0.012	**0.057**	
	(0.024)	(0.005)	**(0.000)**	
Effect of the tax reform, $INC^{CAP} > limit^{new}$.				
N	1,729	861		2,590
After the reform (1999-2001)	0.658	0.627	0.031	1,640
	(0.014)	(0.020)	(0.025)	
Before the reform (1996-1999)	0.613	0.641	-0.028	
	(0.028)	(0.019)	(0.034)	950
Difference within groups	0.045	-0.014	**0.059**	
	(0.024)	(0.035)	**(0.001)**	

Note: The upper panel reports average ownership rates of life insurance policies for the years 1996-2001. The bottom panel reports averages for all households with a capital income $INC^{CAP} > limit^{new}$. The difference-in-difference estimate is reported in bold face in the bottom right cell of each panel. Standard errors are reported in parentheses.

Figure 2.2: The graph depicts the (smoothed) average life insurance ownership rate in East Germany, 1990. Source: GSOEP.

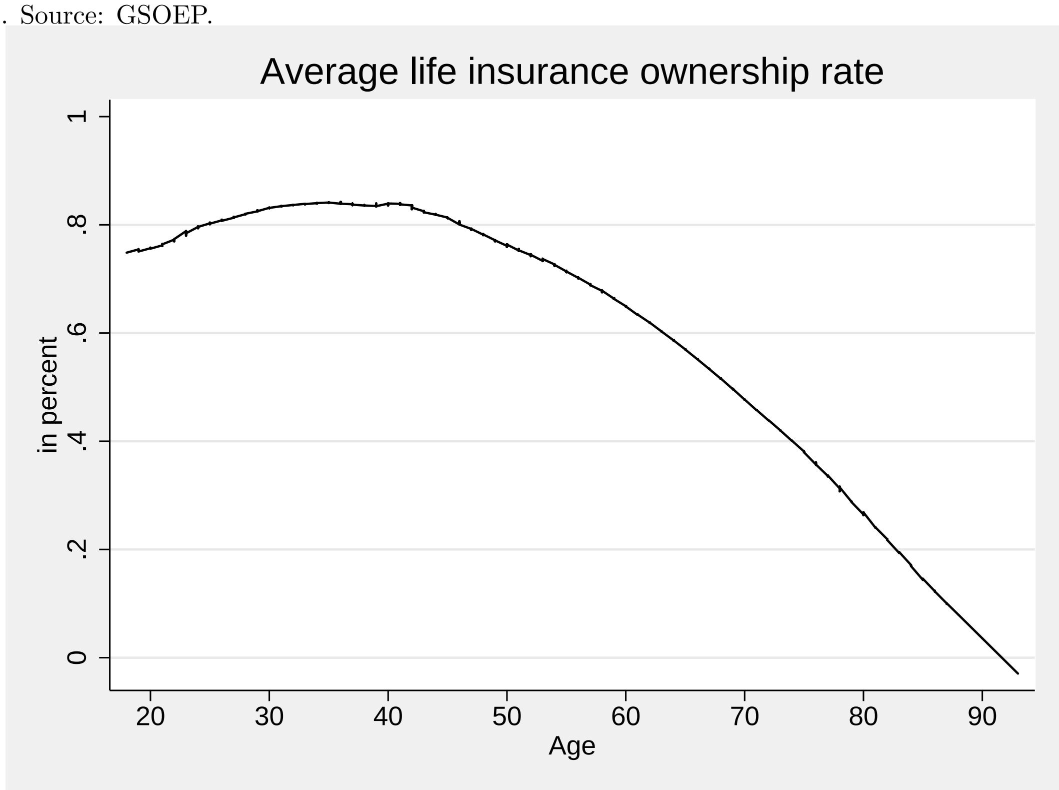

Table 2.4: Tax incentives - summary statistics

		full sample	$INC^{CAP} > limit^{new}$
marginal tax rate		0.249	0.338
woman	D	0.376	0.277
age		48.87	52.39
children	D	0.340	0.218
married	D	0.829	0.730
10 years schooling	D	0.280	0.250
13 years schooling	D	0.200	0.424
college	D	0.086	0.132
university	D	0.100	0.240
self-employed	D	0.057	0.136
civil servant	D	0.045	0.070
retired	D	0.283	0.320
unemployed	D	0.078	0.035
Hhold income decile 1	D	0.099	0.026
Hhold income decile 2	D	0.103	0.039
Hhold income decile 3	D	0.097	0.054
Hhold income decile 4	D	0.101	0.070
Hhold income decile 6	D	0.111	0.100
Hhold income decile 7	D	0.097	0.078
Hhold income decile 8	D	0.097	0.111
Hhold income decile 9	D	0.099	0.174
Hhold income decile 10	D	0.094	0.284
owns house	D	0.405	0.632
returns < 2,000 DM	D	0.231	
returns < 5,000 DM	D	0.084	
returns < 10,000 DM	D	0.033	
returns > 10,000 DM	D	0.016	
N		44,540	2,419
Sample years		1996-2001	1996-2001

Note: The samples are from the GSOEP. Demographic variables refer to the household head. Dummy variables are marked by D.

Table 2.5: Tax incentives - average marginal effects

		(1) full sample		(2) $INC^{CAP} > limit^{new}$	
		estimate	st.error	estimate	st.error
$\hat{\tau}^{DiD}$	D	0.052**	(0.024)	0.089**	(0.046)
T	D	0.010	(0.006)	-0.040	(0.037)
G	D	-0.006	(0.021)	-0.057	(0.036)
marginal tax rate		0.329***	(0.026)	0.250***	(0.089)
woman	D	0.003	(0.008)	-0.075**	(0.031)
age/10		0.153***	(0.017)	0.102*	(0.056)
$(\text{age}/10)^2$		-0.021***	(0.002)	-0.017***	(0.005)
children	D	-0.000	(0.009)	0.033	(0.035)
married	D	0.062***	(0.012)	0.099***	(0.037)
10 years schooling	D	0.059***	(0.009)	-0.037	(0.035)
13 years schooling	D	0.003	(0.012)	-0.096***	(0.036)
college	D	0.005	(0.014)	0.048	(0.039)
university	D	-0.024	(0.015)	-0.002	(0.040)
self-employed	D	0.044***	(0.014)	0.020	(0.041)
civil servant	D	0.025	(0.018)	-0.002	(0.048)
retired	D	-0.008	(0.013)	-0.078*	(0.044)
unemployed	D	-0.022**	(0.011)	0.031	(0.057)
Hhold income decile 1	D	-0.173***	(0.014)	-0.035	(0.084)
Hhold income decile 2	D	-0.082***	(0.013)	0.060	(0.063)
Hhold income decile 3	D	-0.068***	(0.013)	-0.018	(0.061)
Hhold income decile 4	D	-0.040***	(0.011)	0.029	(0.051)
Hhold income decile 6	D	0.023**	(0.011)	0.047	(0.047)
Hhold income decile 7	D	0.030**	(0.012)	0.054	(0.050)
Hhold income decile 8	D	0.039***	(0.012)	0.045	(0.048)
Hhold income decile 9	D	0.044***	(0.013)	0.075	(0.046)
Hhold income decile 10	D	0.068***	(0.014)	0.156***	(0.041)
owns house	D	0.071***	(0.008)	0.057*	(0.029)
returns < 2,000 DM	D	0.080***	(0.007)		
returns < 5,000 DM	D	0.034***	(0.011)		
returns < 10,000 DM	D	0.030*	(0.016)		
returns > 10,000 DM	D	0.039	(0.026)		
Suppressed: year dummies, constant.					
N		44,540		2,419	
Pseudo-R2		0.171		0.209	
χ^2 (*prob.*)		3,249.2 (0.000)		275.9 (0.000)	
Sample years		1996-2001		1996-2001	

D indicates dummy variables. Average marginal effects are reported. Robust and clustered standard errors are reported in parentheses. ***, **, * indicate significance at the 0.01, 0.05, 0.1 level.

Table 2.6: Bequest motives - summary statistics

	subsamples		All observations			
	owner	non-owner				
	Mean	Mean	Mean	Std. Dev.	Min	Max
age	43.2	53.5	46.1	15.9	17	93
woman	0.50	0.55	0.51	0.50	0	1
married	0.77	0.53	0.70	0.46	0	1
10 years schooling	0.48	0.34	0.44	0.50	0	1
13 years schooling	0.15	0.13	0.14	0.35	0	1
master craftsman	0.08	0.07	0.08	0.27	0	1
college	0.20	0.17	0.19	0.39	0	1
university	0.10	0.10	0.10	0.30	0	1
returns < 200 Mark	0.24	0.21	0.23	0.42	0	1
returns < 500 Mark	0.24	0.23	0.24	0.43	0	1
returns < 1,000 Mark	0.13	0.09	0.12	0.33	0	1
returns > 1,000 Mark	0.06	0.08	0.06	0.24	0	1
Hhold income/10,000	0.18	0.13	0.17	0.07	0.03	0.51
partner no job	0.09	0.16	0.11	0.32	0	1
partner income diff./1,000	0.39	0.26	0.35	0.43	0	3.63
retired	0.10	0.38	0.17	0.38	0	1
self-employed	0.03	0.02	0.02	0.15	0	1
civil servant	0.31	0.19	0.27	0.45	0	1
owns house	0.30	0.24	0.29	0.45	0	1
owns weekend house	0.17	0.11	0.15	0.36	0	1
no car	0.36	0.56	0.42	0.49	0	1
motorbike	0.42	0.23	0.37	0.48	0	1
children	0.54	0.28	0.47	0.50	0	1
one child	0.25	0.14	0.22	0.42	0	1
two children	0.24	0.12	0.20	0.40	0	1
three children +	0.05	0.02	0.04	0.20	0	1
family very important	0.89	0.74	0.85	0.36	0	1
unsatisfied social benefits	0.56	0.54	0.55	0.50	0	1
security consciousness	8.83	8.75	8.81	1.83	0	10
Health satisfaction	6.80	6.25	6.65	2.64	0	10
classical concerts, theatre	0.13	0.11	0.12	0.32	0	1
pop concerts, movies, discos	0.13	0.16	0.15	0.36	0	1
active sports	0.14	0.15	0.14	0.35	0	1
meet friends, neighbors	0.58	0.66	0.64	0.48	0	1
help friends, neighbors	0.40	0.49	0.46	0.50	0	1
attend church services	0.11	0.07	0.08	0.28	0	1
N	1487	562	2049			

Note: The sample is the 1990 GSOEP for East Germany. Demographic variables refer to the household head.

Table 2.7: Bequest motives - average marginal effects

		(3)	(4)	(5)	(6)	(7) only <65
age/10		0.165***	0.165***	0.168***	0.167***	0.048
		(0.041)	(0.041)	(0.042)	(0.042)	(0.069)
$(\text{age}/10)^2$		-0.019***	-0.019***	-0.019***	-0.019***	-0.004
		(0.004)	(0.004)	(0.004)	(0.004)	(0.008)
woman	D	0.041**	0.041**	0.037*	0.045**	0.031
		(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
married	D	0.024	0.025	0.013	0.025	0.010
		(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
10 years schooling	D	-0.006	-0.006	-0.001	-0.006	-0.003
		(0.026)	(0.026)	(0.027)	(0.026)	(0.028)
13 years schooling	D	-0.000	-0.000	0.002	0.007	-0.007
		(0.043)	(0.043)	(0.043)	(0.043)	(0.044)
master craftsman	D	0.034	0.034	0.027	0.036	-0.003
		(0.036)	(0.036)	(0.037)	(0.036)	(0.040)
college	D	-0.046*	-0.046*	-0.049*	-0.042	-0.041
		(0.027)	(0.027)	(0.027)	(0.027)	(0.028)
university	D	-0.101**	-0.100**	-0.095*	-0.096*	-0.088*
		(0.049)	(0.049)	(0.049)	(0.049)	(0.051)
returns < 200 Mark	D	0.021	0.021	0.023	0.021	0.012
		(0.024)	(0.024)	(0.024)	(0.024)	(0.025)
returns < 500 Mark	D	0.016	0.016	0.019	0.019	0.017
		(0.024)	(0.024)	(0.024)	(0.024)	(0.025)
returns < 500 Mark	D	0.070**	0.069**	0.071**	0.072**	0.063**
		(0.029)	(0.029)	(0.029)	(0.029)	(0.028)
returns > 1,000 Mark	D	-0.075	-0.075*	-0.078*	-0.069	-0.099**
		(0.046)	(0.046)	(0.046)	(0.046)	(0.050)
Hhold income/10,000	D	2.102***	2.106***	1.776***	2.115***	2.054***
		(0.532)	(0.533)	(0.539)	(0.530)	(0.564)
$(\text{Hhold income}/10{,}000)^2$		-3.919***	-3.927***	-3.119**	-3.938***	-3.613***
		(1.249)	(1.250)	(1.268)	(1.241)	(1.299)
partner no job		-0.006	-0.007	-0.006	-0.005	0.005
		(0.031)	(0.031)	(0.032)	(0.031)	(0.039)
partner income diff./1000	D	0.022	0.022	0.022	0.022	0.021
		(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
retired		-0.045	-0.044	-0.053	-0.053	0.036
		(0.046)	(0.046)	(0.047)	(0.046)	(0.056)
self-employed	D	0.023	0.023	0.018	0.028	0.026
		(0.059)	(0.059)	(0.059)	(0.058)	(0.056)
civil servant	D	0.027	0.026	0.025	0.027	0.017
		(0.022)	(0.022)	(0.022)	(0.022)	(0.021)

. . . Table 2.7 continued . . .

		(3)	(4)	(5)	(6)	(7) only <65
owns house	D	0.016	0.016	0.015	0.013	0.024
		(0.022)	(0.022)	(0.022)	(0.022)	(0.023)
owns weekend house	D	0.023	0.023	0.023	0.021	0.037
		(0.026)	(0.026)	(0.026)	(0.026)	(0.025)
no car	D	0.012	0.011	0.009	0.011	0.007
		(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
motorbike	D	0.058***	0.058***	0.059***	0.057***	0.056***
		(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
children	D	0.070***		0.063**	0.070***	0.071***
		(0.025)		(0.026)	(0.026)	(0.026)
one child	D		0.080***			
			(0.029)			
two children	D		0.068**			
			(0.032)			
three children +	D		0.071			
			(0.054)			
family very important	D			0.070**		0.077**
				(0.028)		(0.033)
unsatisfied social benefits	D			-0.006		0.006
				(0.018)		(0.019)
security conscious				0.001		-0.001
				(0.005)		(0.005)
health satisfaction				-0.000		0.001
				(0.004)		(0.004)
classical concerts, theatre	D				-0.038	
					(0.031)	
pop concerts, movies, discos	D				0.017	
					(0.028)	
active sports	D				-0.036	
					(0.028)	
meet friends, neighbors	D				0.025	
					(0.020)	
help friends, neighbors	D				0.018	
					(0.019)	
attend church services	D				-0.023	
					(0.035)	
N		2,049	2,049	2,024	2,049	1,715
Pseudo-R2		0.145	0.145	0.145	0.148	0.064
χ^2		306.4	306.3	300.1	307.7	107.0
p		0.000	0.000	0.000	0.000	0.000
AIC		2,108.8	2,112.7	2,082.5	2,114.2	1,689.6
BIC		2,249.5	2,264.6	2,245.2	2,288.6	1,847.6

D indicates dummy variables. Average marginal effects are reported. Standard errors are reported in parentheses. ***, **, * indicate significance at the 0.01, 0.05, 0.1 level.

Chapter 3

Do Investors Respond to Tax Reform? Evidence from a Natural Experiment in Germany[1]

3.1 Introduction

While theoretical models of portfolio decisions imply that households take into account the after-tax return of each asset, empirical studies of the importance of tax incentives provide ambiguous results (Poterba, 2002). Studies using data from cross-sections, as for instance Poterba and Samwick (2003), typically face the difficulty of disentangling genuine variation in income, for given tax rates, from genuine variation in after-tax yields, for given income, because marginal tax rates are inherently linked to labor income. Even studies that analyze the impact of tax reforms on the demand for life insurance, which is in many developed countries one of the most tax-advantaged assets, cannot provide conclusive evidence for the importance of tax incentives (Jappelli and Pistaferri, 2003).

We revisit the link between taxation and portfolio choice by analyzing households' responses to a tax reform in Germany which revoked the tax exemption of life insurance returns for all policies bought after January 1, 2005. Using a difference-in-differences estimator on repeated cross-sectional data, we test if a treatment group of investors that are affected by the new tax regime is more likely to own life insurance after the reform than a control group of investors that are unaffected by the reform. We find conclusive evidence that the reform was anticipated and that demand increased among

[1]This chapter is joint work together with Joachim Winter.

households in the top tax quartile in the year before the tax exemption was abolished.

We proceed by discussing some key features of the reform in section 3.2. We describe the data in section 3.3. The empirical analysis is presented in section 3.4 before we conclude in section 3.5.

3.2 The Tax Reform of 2005

Life insurance is the second most popular financial asset in German households' portfolios, after savings accounts. In 2007, 15.6 percent of total private wealth, amounting to 716 billion Euro, was allocated to life insurance (Deutsche Bundesbank, 2008). Around 49 percent of households own life insurance policies.[2] One of the main reasons for this unusually high popularity is the favored fiscal treatment of life insurance returns. Returns have historically been fully exempt from taxation if the contract lasts for at least 12 years, premiums are paid during at least five years, and the term life insurance component amounts to at least 60 percent of the total benefit paid out at the end of the contract. However, due to a tax reform that was announced in mid-2004, returns of all policies bought after January 1, 2005 are taxed at half the individual marginal tax rate (under the above conditions). This reform was unanticipated and not intended to offset group-specific trends in life insurance ownership. To illustrate the effect of the announcement of the reform on life insurance demand, the left panel of figure 3.1 depicts an index of internet search volumes for the term *Lebensversicherung* (life insurance) in Germany, relative to the average search volume for this term between 2004 and 2008. Search volumes increased substantially during the months preceding the tax reform. At the end of 2004, searches were three times larger than the average search volume over the 2004–2008 period. The right panel of figure 1 shows life-insurance sales in Germany over the 2000–2007 period. Sales of (tax-exempt) whole-life insurance policies, which combine a term-life insurance contract with a savings plan, spike in 2004, whereas sales of pure term-life insurance policies remain relatively constant over the entire period. These graphs show that after the announcement of the reform, households compared the conditions of different insurers (higher internet search volume) and acquired life insurance before the new tax regime came into effect (higher sales volume only for whole-life insurance).

[2]Authors' calculation based on the data from the GSOEP, described in section 3.3.

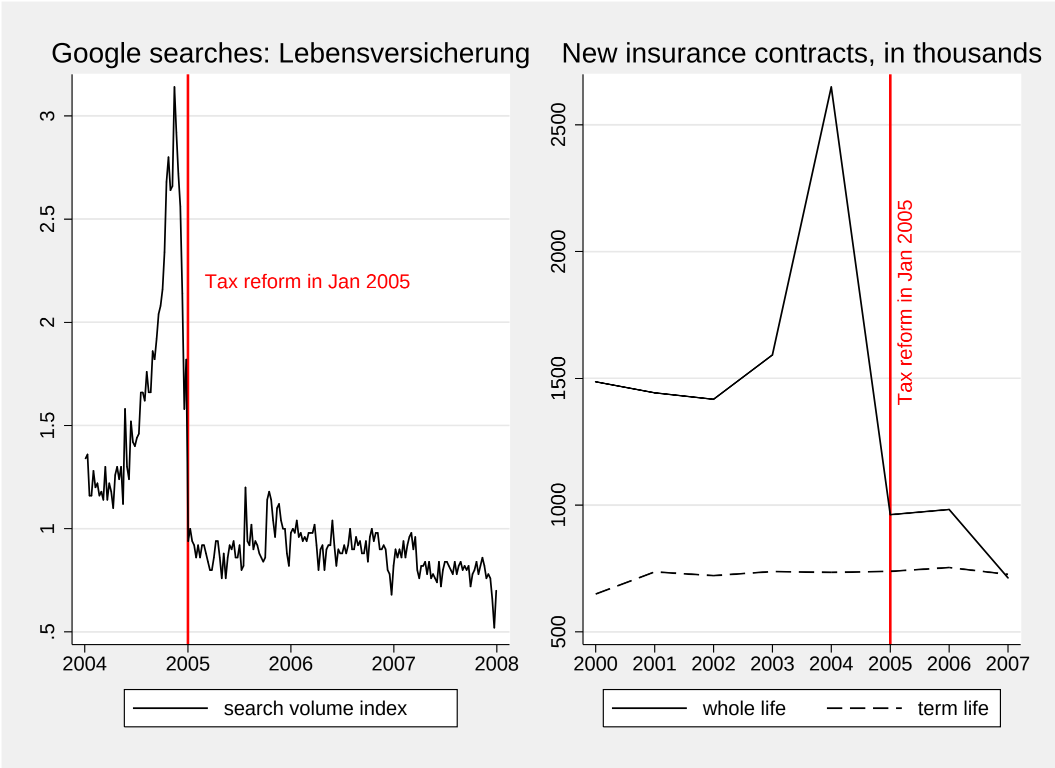

Figure 3.1: Internet search volumes and sales of new contracts around the 2005 tax reform. *Sources*: Google Trends (http://www.google.de/trends), left panel; Gesamtverband der Deutschen Versicherungswirtschaft (2008), right panel.

3.3 The Data

Our data are taken from the German Socioeconomic Panel (GSOEP), the only dataset containing annual information on life insurance ownership of German households that covers pre- and post-reform years. The add-on package PanelWhiz for Stata (http://www.PanelWhiz.eu) has been used for extracting our data from the GSOEP files; see Haisken-DeNew and Hahn (2006) for details.[3] Our data span a before-reform period (2000–2003) and the year when the reform was announced (2004). The dependent variable in our subsequent analysis is life insurance holdings, defined as a binary variable. Each year households are asked if they owned one or more life insurance policies in the previous year. The independent variable of interest is the marginal tax rate which is unobserved in the GSOEP data. We approximate marginal tax rates by re-calculating each household's taxable income from (estimated) tax payments, using the official formulas of the federal tax office.[4] A one unit change in taxable income is

[3]The Stata program generated by PanelWhiz to retrieve the data is available upon request. Any data or computational errors are our own.

[4]The estimates of total tax payments provided by the GSOEP are based on Schwarze's (1995) approach. Schwarze adds up the incomes of all household members and applies standard deductions

simulated in order to approximate the marginal tax rate. The socioeconomic characteristics that are used as additional independent variables in our multivariate analysis (see table 3.1 below) are naturally defined and refer to the household head.

3.4 Difference-in-Differences Estimates

As the post-reform tax regime links after-tax yields to marginal tax rates, households in higher tax brackets are more likely to avoid taxation by preponing life insurance purchases. Table 3.1 reports that the ownership rate was already very high (around 75.3 percent) in 2000–2003 among households in the top quartile of the marginal tax distribution. Thus, we are most likely to observe changes at the intensive rather than at the extensive margin. However, the GSOEP data do not provide information on the amounts invested.

Table 3.1: Anticipation effect of the tax reform

	treated	non-treated	Difference between groups	N
N	14,997	42,184		57,181
Anticipating the reform (2004)	0.754	0.440	0.314	11,086
	(0.008)	(0.005)	(0.010)	
Before the reform (2000-2003)	0.753	0.460	0.293	46,095
	(0.004)	(0.003)	(0.005)	
Difference within groups	0.001	-0.020	**0.021**	
	(0.009)	(0.006)	**(0.000)**	

Note: The table reports average ownership rates of life insurance policies for the years 2000–2004. The unconditional difference-in-difference estimate is reported in bold face in the bottom right cell of the panel. Standard errors are reported in parentheses.

We use the data in table 3.1 to compute an unconditional difference-in-differences estimate of the effect on the taxpayer group that is most affected by the reform. The ownership probability increased by 2.1 percent for households in the top quartile of the marginal tax distribution. Next, we make sure that the estimate is bounded in the $[0;1]$ interval and control for additional covariates that could account for different behavior across groups. We denote individual i's binary indicator for the treatment group as $G_i = \mathbb{1}\{MTR_{it} \geq MTR_t^{75}\}$, where MTR_t^{75} denotes the 75^{th} percentile of the marginal tax distribution in year t. We assume that group membership is exogenously determined, i.e., households did not change marginal tax brackets as a result of the reform itself. $T_i = \mathbb{1}\{t \geq 2004\}$ defines a time dummy for the reform year. To ease the notational burden, we introduce the shorthand $Y_{i \in g,t}$ for $Y_i | G_i = g, T_i = t$. The

based on the socioeconomic status of the household.

potential outcomes with and without treatment are Y_i^1 and Y_i^0 respectively. Based on the standard probit difference-in-differences model

$$P(Y_i = 1|G_i, T_i, \boldsymbol{x}_i) = \Phi(\alpha T_i + \beta G_i + \gamma T_i G_i + \boldsymbol{x}_i \boldsymbol{\delta}), \tag{3.1}$$

where $\boldsymbol{x}_i$ is a vector of additional independent variables with coefficients $\boldsymbol{\delta}$, Puhani (2008) shows that a consistent estimator of the average treatment effect on the treated is

$$\begin{aligned} \widehat{\tau}^{DiD} &= E[Y^1_{i\in 1,1}|\boldsymbol{x}_i] - E[Y^0_{i\in 1,1}|\boldsymbol{x}_i] \\ &= \frac{1}{N}\sum_{i=1}^{N}\left[\Phi\left(\widehat{\alpha}+\widehat{\beta}+\widehat{\gamma}+\boldsymbol{x}_i\widehat{\boldsymbol{\delta}}\right) - \Phi\left(\widehat{\alpha}+\widehat{\beta}+\boldsymbol{x}_i\widehat{\boldsymbol{\delta}}\right)\right]. \end{aligned} \tag{3.2}$$

We apply the delta method to infer statistical significance of the average treatment effect in small samples. Table 3.2 reports summary statistics for the additional covariates. Also, average marginal effects for the probability to hold one or more life insurance policies are reported for continuous and dummy variables. Our estimate of the average treatment effect on the treated, $\widehat{\tau}^{DiD}$, remains similar to the unconditional estimate when controlling for other potential determinants of life insurance demand and is statistically significant at the 5 percent level.

3.5 Conclusion

The abolishment of tax favors for life insurance bought in Germany after January 1, 2005 triggered last-minute purchases in anticipation of the reform. At the extensive margin, life insurance demand increased by 2 percent among households in the top tax quartile. However, the increase of internet search volumes in late 2004 suggests that changes at the intensive margin, for which coherent data is not available, are likely to be larger. Our results contrast with Jappelli and Pistaferri (2003) who cannot find that tax incentives matter for life insurance demand in Italy. Anecdotal evidence suggests that insurance sales agents exploited the favorable market situation in Germany in 2004, whereas Italian insurers lacked the vital initiative to point out the effects of the tax reform among potential investors.

Table 3.2: Summary statistics and probit estimates

		Summary statistics		Average marginal effects	
		owners	non-owners	estimate	st.error
$\hat{\tau}^{DiD}$	D	0.07	0.02	0.020**	(0.010)
T	D	0.19	0.20	-0.020***	(0.006)
G	D	0.36	0.12	0.049***	(0.009)
woman	D	0.33	0.42	-0.006	(0.007)
age/10		4.66	5.52	0.159***	(0.016)
$(\text{age}/10)^2$		23.4	33.7	-0.022***	(0.002)
children	D	0.39	0.21	-0.021***	(0.008)
married	D	0.84	0.81	0.048***	(0.011)
10 years schooling	D	0.33	0.23	0.041***	(0.008)
13 years schooling	D	0.29	0.20	-0.009	(0.011)
college	D	0.11	0.07	0.020	(0.013)
university	D	0.15	0.09	-0.003	(0.013)
self-employed	D	0.09	0.04	0.028**	(0.013)
civil servant	D	0.07	0.03	-0.003	(0.016)
retired	D	0.16	0.46	-0.070***	(0.013)
unemployed	D	0.05	0.09	-0.080***	(0.011)
Hhold income decile 1	D	0.04	0.17	-0.235***	(0.012)
Hhold income decile 2	D	0.06	0.15	-0.123***	(0.012)
Hhold income decile 3	D	0.07	0.13	-0.099***	(0.012)
Hhold income decile 4	D	0.09	0.12	-0.053***	(0.011)
Hhold income decile 6	D	0.13	0.10	0.032***	(0.010)
Hhold income decile 7	D	0.11	0.07	0.043***	(0.011)
Hhold income decile 8	D	0.13	0.06	0.061***	(0.012)
Hhold income decile 9	D	0.13	0.05	0.071***	(0.012)
Hhold income decile 10	D	0.14	0.04	0.120***	(0.013)
owns house	D	0.27	0.19	0.069***	(0.007)
returns < 2,000 DM	D	0.11	0.08	0.073***	(0.007)
returns < 5,000 DM	D	0.05	0.03	0.043***	(0.010)
returns < 10,000 DM	D	0.03	0.02	0.048***	(0.014)
returns > 10,000 DM	D	0.53	0.38	0.022	(0.017)
Suppressed: year dummies, constant.					
N		27,289	23,385	50,674	
Pseudo-R2				0.180	
χ^2 (*prob.*)				3,733.6 (0.000)	
Sample years				2000-2004	

Note: ***, **, * indicate significance at the 0.01, 0.05, 0.1 levels. The dependent variable is 1 if the households owns one or more life insurance policies and zero otherwise. D indicates dummy variables. Average marginal effects are reported. Robust and clustered standard errors are reported in parentheses.

References

Deutsche Bundesbank (2008): *Ergebnisse der gesamtwirtschaftlichen Finanzierungsrechnung für Deutschland 1991 bis 2007, Statistische Sonderveröffentlichung 4.* Deutsche Bundesbank: Frankfurt.

Haisken-DeNew, John.P. and Michael Hahn (2006): "PanelWhiz: A Flexible Modularized Stata Interface for Accessing Large Scale Panel Data Sets," mimeo, (http:// www.PanelWhiz.eu).

Gesamtverband der Deutschen Versicherungswirtschaft (2008): *Geschäftsentwicklung 2007: Die deutsche Lebensversicherung in Zahlen.* GDV: Berlin.

Jappelli, Tullio and Luigi Pistaferri (2003): "Tax Incentives and the Demand for Life Insurance," *Journal of Public Economics*, **87**, 1779-1799.

Poterba, James M. (2002): "Taxation and Portfolio Structure: Issues and Implications," in *in: Luigi Guiso, Michael Haliassos, Tullio Jappelli (eds.)*, Household Portfolios. MIT Press: Cambridge, 103-142.

Poterba, James M. and Andrew Samwick (2003): "Taxation and Household Portfolio Composition: U.S. Evidence from the 1980s and 1990s," *Journal of Public Economics*, **87**, 5-38.

Puhani, Paul A. (2008): "The Treatment Effect, the Cross Difference, and the Interaction Term in Nonlinear "Difference in Differences" Models," IZA Discussion Paper No. 3478.

Schwarze, Johannes (1995): "Simulating German Income and Social Security Tax Payments using the GSOEP," Cross-National Studies in Aging Program Project Paper 19.

Chapter 4

Adopting to New Financial Products: Evidence from the Demand for Building Society Contracts in East Germany

4.1 Introduction

The rising importance of financial services in modern economies and a high pace of financial innovation foster the interest in the determinants of financial innovation. While several studies analyze supply-side factors driving the introduction of new financial tools, empirical studies of the determinants of households' adoption to new financial products are relatively rare. The only two examples of studies assessing investor behavior, which are listed in the recent overview paper by Frame and White (2004), are Mantel (2000) and Mantel and McHugh (2001). The authors study who first uses electronic bill payment and debit card services. They find that usage is positively related to age, income, and gender (female). This dearth of empirical research on adoption processes to new financial products is mainly due to the lack of accessible data that allows analyzing investor behavior.

This paper contributes to the literature by studying a unique natural experiment in the German financial market. For the first time in 1990, East Germans had the opportunity to save into building society contracts (BSCs) after 40 years of communism. As both home ownership and wealth were low in the former East Germany, BSCs were an attractive new financial product for people of all age groups to fulfill their dream of an own home by starting to save for the downpayment of a mortgage. This paper analyzes (i) who uses BSCs to save for a house purchase, and (ii) how long it takes after reunification until future BSC investors start to save into BSCs.

BSCs are a subsidized and tax-favored savings product that facilitates saving for downpayments of mortgage credits. They offer a savings plan with a predetermined value and interest rate. If the investor accumulated half of the contract's value, the building society typically allocates credit for the remaining value.

Given the importance of BSCs in the portfolios of German households, studying the adoption to BSCs in East Germany can improve our understanding as to why some households postpone investing into financial products they will eventually own. We find that households with close ties to their families in the West (which were aware of the features of BSCs) invest earlier, pointing to the presence of information asymmetries and the importance of social networks in the adoption process. This is in line with Hong, Kubik, Stein (2004), who show that social investors have lower entry costs into stock markets. There is also evidence that households trade-off long-term savings goals for short-term consumption, because households owning a car already in 1990 are more likely to invest into BSCs and to do so earlier. Life insurance appears to be a substitute for BSCs, which is not surprising, given that both provide a tax-favored means to accumulate savings over the long-term.

The remainder of this paper is structured as follows. In section 4.2, the data used in this study is presented. The empirical methodology applies a split-population survival model, which is discussed in section 4.3. Results from the estimation are reported in section 4.4. Section 4.5 concludes.

4.2 Descriptive Analysis

Owner-occupied housing typically accounts for a major share of households' wealth holdings and is the largest financial transaction conducted by most households during their lifetime. For instance, around 80% of US households are house owners, investing 47 percent of their wealth into real estate (Lusardi and Mitchell, 2007). The current financial crisis, which began in the US housing market, illustrates the crucial impact of people's house financing decisions on the economy as a whole. Before the financial crisis, cheap credit encouraged banks to reduce downpayment requirements, so that potential home owners could raise mortgages that were solely backed by the face value of the house they acquired. Such times are over now, so that the eligibility for mortgage credit will again depend on a household's ability to first pay a downpayment out of its savings. BSCs combine a savings plan for the downpayment with a traditional mortgage contract.

Figure 4.1 depicts average ownership rates of owner-occupied housing and BSCs for East and West Germany. Between 1990 and 2006, ownership rates were relatively stable in West Germany with a constant wedge between house and BSC ownership. In the East, only around 26 percent of the population owned a house or apartment in 1990. After reunification, house ownership picked up only slowly. However, the adaption to BSCs happened quickly, so that in 1994 nearly as many East as West Germans invested into BSCs.

The data used in this study come from the German Socioeconomic Panel (GSOEP), which is a longitudinal survey of private households in Germany.[1] The subsample covering the territory of the former GDR started in 1990. The GSOEP contains data on the ownership of BSCs in each year. As BSCs were not available before 1990, all households from the initial sample can be observed as either saving or not saving into BSCs at some point during the 17 years following reunification. The dependent variable in all regressions is the number of years since 1990 until first BSC ownership. For households not owning BSCs throughout the entire sampling period, the dependent variable is set to 17. Several households also drop out of the sample before the sampling period ends. Most of these early drop-outs occurred during the first three sampling years. In total, the sample consists of 1774 households.

Figure 4.2 provides non-parametric Kaplan-Meier estimates of the hazard and survivor functions. The non-parametric survival function in figure 4.2 reaches a limit at around 0.35, corresponding to the proportion of the sample that had never invested into BSCs until the last year of observation. The non-parametric estimate of the hazard function resembles closely to hazards from a continuous-time parametric model. In what follows, we thus limit ourselves to modeling a continuous-time split-population duration model.[2]

We explain ownership probabilities and duration based on the initial characteristics of the household at reunification. Summary statistics of the explanatory variables are reported in table 4.1. We use dummies for age, gender, martial status, education, income, and wealth. Figure 4.3 depicts age and cohort effects. There is strong indication that households that were older than 60 years in 1990 did not save for buying a house anymore. In contrast, differences in age seem to be negligible for the younger

[1] The Add-On package PanelWhiz for Stata (http://www.PanelWhiz.eu) has been used for extracting the data. See Haisken-DeNew and Hahn (2006) for details. The PanelWhiz generated DO file to retrieve the data used here is available upon request.

[2] Discrete-time split-population models, as proposed by Stephen Jenkins (*spsurv*), show very poor convergence properties.

generations.

Determinants of adoption behavior that are of particular interest are information asymmetries and consumption-savings trade-offs. We study if households adopt quicker to BSCs if they owned a house, car, or life insurance in 1990, or received gifts from their relatives. Figure 4.4 shows that house owners were more likely to invest into BSCs. Given the bad condition of many estates, BSCs provided a subsidized and tax-favored tool to save for the renovation of owner-occupied homes. The difference in BSC ownership is most pronounced among households with a car. Initial car owners have a much higher BSC ownership rate, indicating that consumption was more important to most households than long-term savings goals. Similarly, households receiving gifts from relatives in the West are slightly more likely to invest into BSCs. However, life insurance ownership is strongly negatively correlated with savings into BSCs due to the substitutability of both financial products.

4.3 Methodology

For the BSC-investors in the sample, we use the reported age of first ownership, so that the duration can be interpreted as a complete spell. However, 49 percent of the observations in the sample are not buying BSCs in any of the 17 years we observe. In a parametric duration model, these observations would be interpreted as incomplete, right-censored spells, assuming that these individuals will eventually fail and buy BSCs. Therefore, we use the split-population model proposed by Foster and Jones (2000), which applies the duration process only to those individuals that are predicted to eventually invest into BSCs. Defining $b = 1$ for a household that will eventually invest into BSCs and modeling eventual failure by using a probit specification yields:

$$\begin{aligned} &\textit{P(eventually invest into BSCs)}= P(b = 1|\mathbf{z_i}) = \Phi(\alpha'\mathbf{z_i}) \\ &\textit{P(never invest into BSCs)} = P(b = 0|\mathbf{z_i}) = 1 - \Phi(\alpha'\mathbf{z_i}), \end{aligned}$$

where $\mathbf{z_i}$ is a vector of covariates for household i. The probability of investing into BSCs at a given time t is then defined conditionally on an eventual investment. Based on standard model selection criteria (table 4.2) for the unconditional case as well as Cox-Snell residuals (figure 4.5), we choose a log-logistic distribution to model duration. We use the plots of the cumulative Cox-Snell residuals to assess the general fit of the models. A correctly fitted model should yield cumulative Cox-Snell residuals which resemble a sample from a standard exponential distribution. A plot of the non-parametric estimate of the cumulative hazard function for these data should therefore lie on a 45^0 degree line through the origin. We find the best fit for a log-logistic function, for which the Cox-Snell residuals are closest to the diagonal.

In order to check if also the proportional hazard assumption holds for the variables of interest, Jenkins (2005) suggests to take recourse to the log-odds survival interpretation of the log-logistic function. For a non-parametric (Kaplan-Meier product-limit) estimate of the survivor function, $S(.)$, the log-odds property suggests that a plot of $\ln(\frac{S(t,\mathbf{x_i})}{1-S(t,\mathbf{x_i})})$ against $\ln(t)$, should be a straight line if the log-logistic model is appropriate. If plotted separately for two different groups classified by combinations of $\mathbf{x_i}$, the lines should move parallel. Figure 4.6 shows that the log-odds of survival plots are straight, which confirms the choice of the log-logistic model of duration. In addition, the lines are parallel in all four cases, suggesting that the variables of interest affect the hazard proportionally.

The probability density function $f(.)$ and the survival function $S(.)$ of the log-logistic distribution for those households eventually starting to invest are respectively

$$f(t|b=1,\mathbf{x_i}) = \frac{\lambda^{\frac{1}{\gamma}} t^{\frac{1}{\gamma}-1}}{\gamma[1+(\lambda t^{\frac{1}{\gamma}})]^2} \tag{4.1}$$

$$S(t|b=1,\mathbf{x_i}) = \frac{1}{1+\lambda t^{\frac{1}{\gamma}}}, \tag{4.2}$$

where $\lambda = exp(-\beta'\mathbf{x_i})$, $\mathbf{x_i}$ is a vector of time-invariant covariates and γ is a scale parameter. For identification, $\mathbf{z_i}$ includes in addition to the variables included in vector $\mathbf{x_i}$ also a variable that is 1 if a the interview in 1990 lasted for more than 15 minutes and zero otherwise. We expect that the 41% of households taking less than 15 minutes to answer the survey are most likely to drop out during the first waves. This provides identification of the probit regression as long as disinterest in participating in the survey is orthogonal to the decision to save into BSCs. The log-likelihood contributions for the split population model then become:

$$L = c_i \ln[f(t|b=1,\mathbf{x_i})] + (1-c_i)\ln[1-\Phi(\alpha'\mathbf{z_i}) + \Phi(\alpha'\mathbf{z_i})S(t|b=1,\mathbf{x_i})]. \tag{4.3}$$

For those who are observed as BSC-investors, $c_i = 1$, the contribution is simply the probability density function of investing at some point, $\Phi(\alpha'\mathbf{z_i})$, multiplied by the probability density function of the observed starting date of the contract, $f(.)$. For those who are observed as not starting (including right-censored observations), $c_i = 0$, the contribution is the logarithm of the probability of never saving into BSCs, $1-\Phi(\alpha'\mathbf{z_i})$, plus the probability of investing after the last observed survey date, $\Phi(\alpha'\mathbf{z_i})S(.)$.

4.4 Results

The results of the estimation are presented in table 4.3. There is indeed evidence that age has a significantly negative impact on BSC ownership among older investors. Households older than 65 years at reunification are 22 percentage points less likely to invest into BSCs. We find little indication that gender, martial status, or education effect ownership of BSCs.

While initial financial wealth does not have a significant impact on BSC ownership, households with higher labor income are more likely to invest into BSCs. Moreover, initial house owners are more likely to save in BSCs, whereas life insurance ownership in 1990 has a significantly negative impact on investment into BSC. This supports the idea that life insurance and BSCs are partial substitutes. Finally, households taking more than 15 minutes to answer the questionnaire are less likely to drop out early, confirming the validity of our instrument.

Regarding the timing of the investment decision, households with an upper medium income invest earlier. Furthermore, house and car ownership in 1990 reduces the time until a household starts to save via BSCs. Households receiving gifts from relatives are also more likely to invest relatively early, although the estimate is only significant at the 10 percent level. Figure 4.7 shows graphically that the estimated hazard probabilities for initial house owners, car owners and gift receivers have a higher-than-average hazard probability, while life insurance owners have a lower-than-average hazard probability.

4.5 Conclusion

Despite the importance of financial innovation for modern economies, determinants of the adoption process of investors to new financial products have so far received little attention by researchers. In this paper, we argue that German reunification provides a natural experiment to study ownership and the timing of investments into BSCs. After 1990, East German households could decide to use BSCs in order to save for the downpayment of a mortgage and when to do so. Overall, households adopt quickly to this new savings product and ownership rates catched up to West German levels.

The evidence presented points towards the importance of information networks and consumption-savings trade-offs in savings decisions. We find that households with close ties to their families in West Germany take less time until they start saving into BSCs, suggesting that information and financial help from relatives in the West facilitate BSC investments. We also find strong indication that long-term investments are traded-off for short-term consumption goals, as initial car owners are more likely to own BSCs and invest earlier. Life insurance, however, appears to be a substitute for BSCs. Contrary to what the literature on financial literacy would suggest, however, we find no differences in ownership and duration with regard to education characteristics of the households.

References

Forster, Martin and Andrew Jones (2000): "The Role of Tobacco Taxes in Starting and Quitting Smoking: Duration Analysis of British Data," *Journal of the Royal Statistical Society*, **164**, 517-547.

Frame, Scott W. and Lawrence J. White (2004): "Empirical Studies of Financial Innovation: Lots of Talk, Little Action?," *Journal of Economic Literature*, **42**, 116-144.

Haisken-DeNew, John P. and Markus Hahn (2006): "PanelWhiz: A Flexible Modularized Stata Interface for Accessing Large Scale Panel Data Sets," mimeo, (http:// www.PanelWhiz.eu).

Hong, Harrison; Jeffrey D. Kubik and Jeremy C. Stein (2004): "Social Interaction and Stock Market Participation," *Journal of Finance*, **59**, 137-163.

Jenkins, Stephen (2005): *Survival Analysis.* Manuscript, University of Essex, http://www.iser.essex.ac.uk/teaching/degree/stephenj/ec968/index.php.

Lusardi, Annamaria and Olivia Mitchell (2007): "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth," *Journal of Monetary Economics*, **54**, 205-224.

Mantel, Brian (2000): "Why Do Consumers Pay Bills Electronically? An Empirical Analysis," *Economic Perspectives, Federal Reserve Bank of Chicago*, **4**, 32-47.

Mantel, Brian and Timothy McHugh (2001): "Competition and Innovation in the Consumer e-Payments Market? Considering Demand, Supply, and Public Policy Issues," Federal Reserve Bank of Chicago, Emerging Payments Occasional Working Paper Series, EPS-2001-4.

Figures and Tables

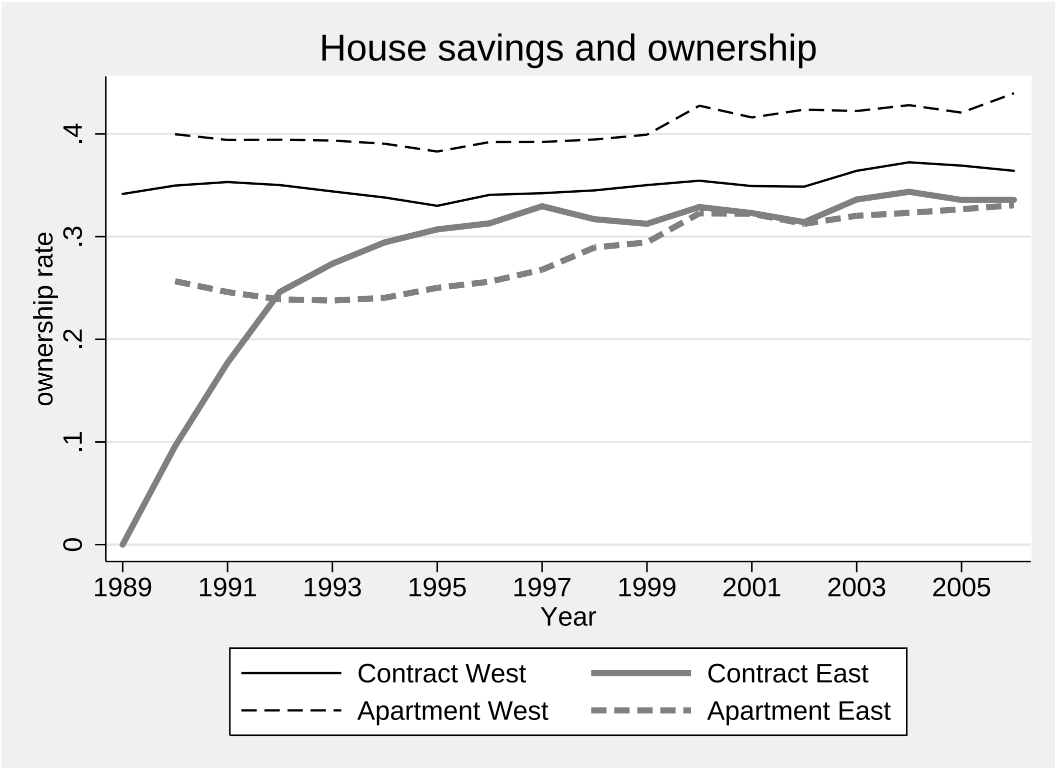

Figure 4.1: Average ownership rates of building society contracts and owner-occupied homes. Author's calculations based on data from the GSOEP.

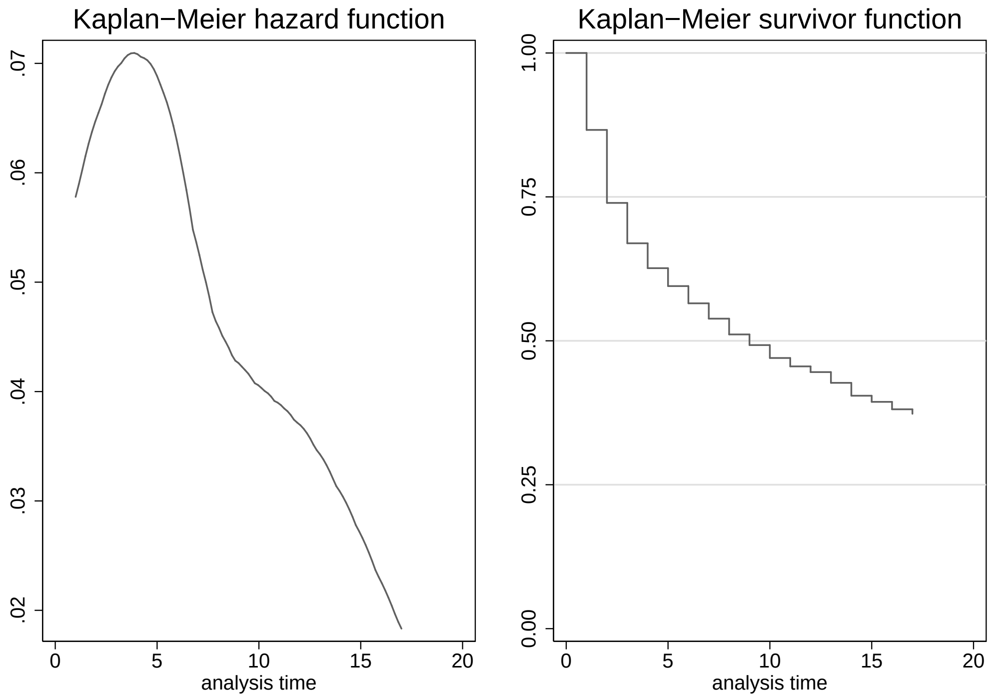

Figure 4.2: Non-parametric estimates of the hazard and survival functions. Author's calculations.

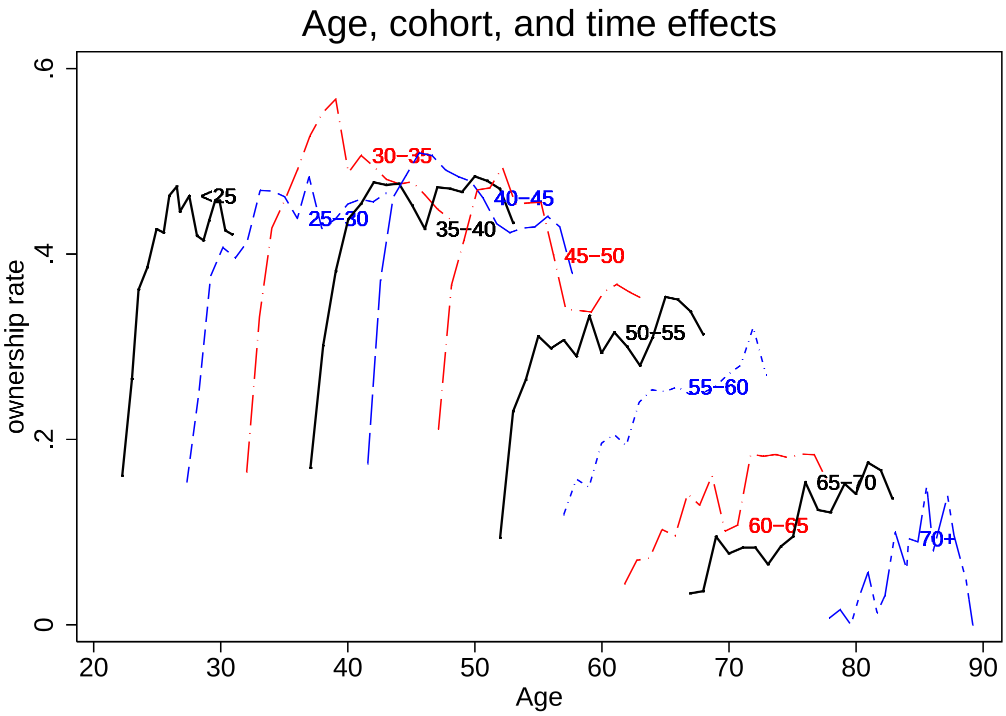

Figure 4.3: Average ownership rate of building society contracts across cohorts. Author's calculations.

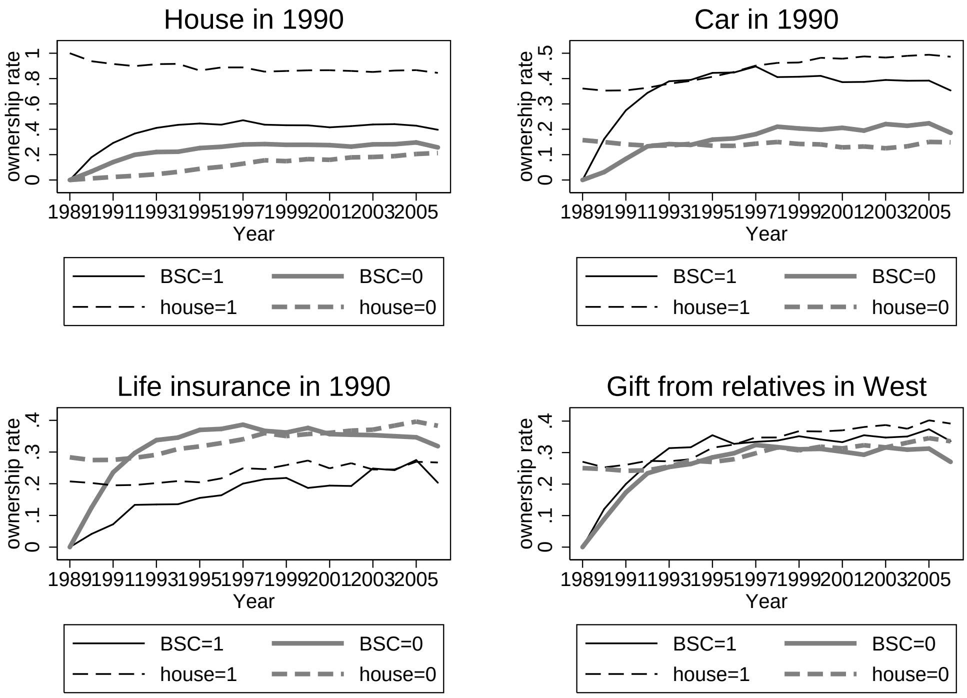

Figure 4.4: Average ownership rates of building society contracts (BSCs) and owner-occupied homes by different initial conditions. Author's calculations based on data from the GSOEP.

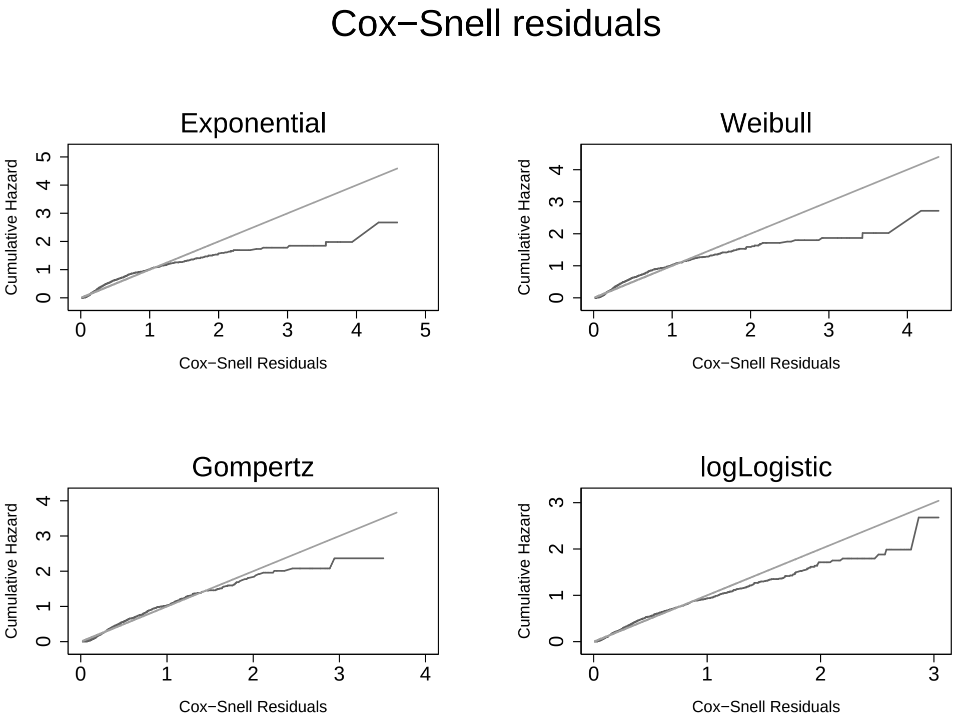

Figure 4.5: Estimates of the Cox-Snell residuals from regressions that include the full set of explanatory variables. Author's calculations.

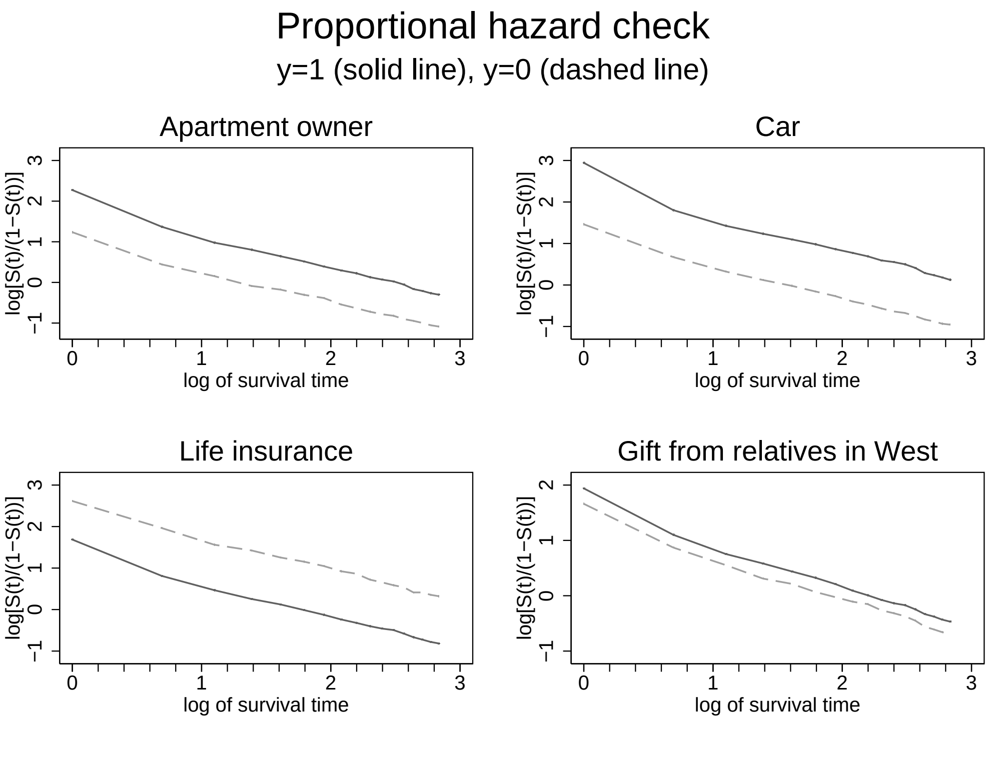

Figure 4.6: Check of the proportional hazard specification. Author's calculations.

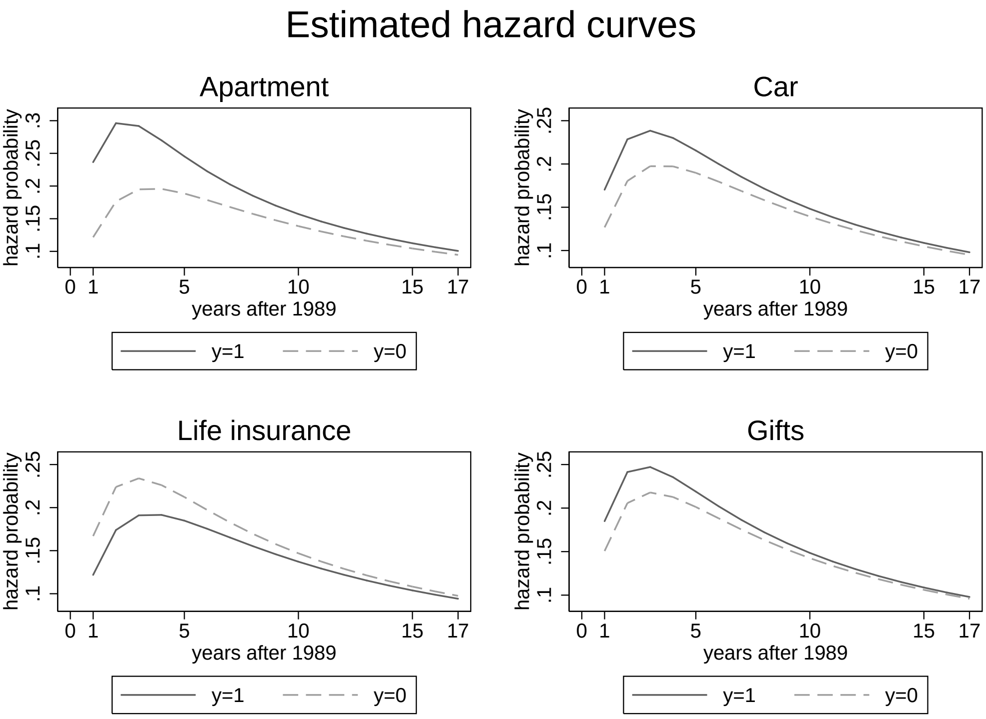

Figure 4.7: Estimated hazard curves of model 3. Author's calculations.

Table 4.1: Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
cohort 35-44	1774	0.21	0.41	0	1
cohort 45-54	1774	0.21	0.40	0	1
cohort 55-64	1774	0.14	0.34	0	1
cohort 65+	1774	0.14	0.35	0	1
woman	1774	0.51	0.50	0	1
married	1774	0.91	0.28	0	1
10 years schooling	1774	0.45	0.50	0	1
13 years schooling	1774	0.15	0.35	0	1
college	1774	0.28	0.45	0	1
income quintile 2	1774	0.20	0.40	0	1
income quintile 3	1774	0.20	0.40	0	1
income quintile 4	1774	0.20	0.40	0	1
income quintile 5	1774	0.19	0.39	0	1
wealth < 6,000	1774	0.19	0.40	0	1
wealth < 15,000	1774	0.17	0.38	0	1
wealth < 30,000	1774	0.07	0.26	0	1
wealth 30,000+	1774	0.05	0.21	0	1
kids	1774	0.46	0.50	0	1
house in 1990	1774	0.31	0.46	0	1
car in 1990	1774	0.61	0.49	0	1
life insurance in 1990	1774	0.26	0.44	0	1
gifts from west	1774	0.24	0.43	0	1
interview > 15 min.	1774	0.59	0.49	0	1

Table 4.2: Model choice

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
Exponential	1774	-2303.8	-2057.4	23	4160.8	4286.8
Weibull	1774	-2291.8	-2056.1	24	4160.3	4291.8
Gompertz	1774	-2239.0	-2023.0	24	4093.9	4225.4
log-logistic	1774	-2235.7	-1979.0	24	4005.9	4137.5

Table 4.3: Split-population regression

	Hazard		Probit			
	estimate	st.err.	estimate	st.err	marginal effect	st.err
cohort 35-44	-0.023	(0.099)	-0.161	(0.167)	-0.026	(0.030)
cohort 45-54	-0.104	(0.128)	-0.493***	(0.187)	-0.089	(0.043)
cohort 55-64	0.300*	(0.177)	-0.675***	(0.213)	-0.140	(0.057)
cohort 65+	0.398*	(0.241)	-0.963***	(0.252)	-0.224	(0.077)
woman	0.087	(0.072)	0.102	(0.104)	0.016	(0.015)
married	-0.389*	(0.200)	-0.145	(0.285)	-0.021	(0.045)
10 years schooling	-0.135	(0.107)	-0.100	(0.143)	-0.016	(0.024)
13 years schooling	-0.140	(0.139)	-0.353*	(0.198)	-0.063	(0.042)
college	-0.110	(0.088)	0.134	(0.138)	0.020	(0.019)
income quintile 2	-0.257	(0.170)	0.245	(0.191)	0.035	(0.023)
income quintile 3	-0.172	(0.166)	0.435**	(0.202)	0.059	(0.020)
income quintile 4	-0.368**	(0.168)	0.271	(0.198)	0.039	(0.024)
income quintile 5	-0.215	(0.173)	0.635***	(0.223)	0.082	(0.019)
wealth < 6,000	0.096	(0.093)	0.054	(0.134)	0.008	(0.020)
wealth < 15,000	0.095	(0.096)	0.202	(0.145)	0.029	(0.018)
wealth < 30,000	0.096	(0.136)	0.171	(0.200)	0.024	(0.026)
wealth 30,000+	0.353	(0.215)	0.255	(0.292)	0.034	(0.033)
kids	-0.151	(0.099)	0.249*	(0.148)	0.039	(0.020)
house in 1990	-0.417***	(0.074)	0.401***	(0.116)	0.059	(0.013)
car in 1990	-0.183**	(0.091)	0.223*	(0.121)	0.036	(0.017)
life insurance in 1990	0.195*	(0.105)	-0.365***	(0.127)	-0.065	(0.027)
gifts from west	-0.130*	(0.078)	0.128	(0.117)	0.019	(0.016)
interview > 15 min.			0.223**	(0.095)	0.036	(0.013)
constant	2.247***	(0.218)	0.184	(0.295)		
shape	0.547***	(0.019)				
N	1774					
ll	-2851.98					
χ^2	116.61					
AIC	5799.95					
BIC	6063.04					

Note: Significant at the * 0.10, ** 0.05, *** 0.01 level.

Chapter 5

Talking Trade: Language Barriers in Intra-Canadian Commerce

5.1 Introduction

In the light of falling tariffs and transport cost, the importance of institutional barriers to trade has captured much attention in recent research.[1] The existence of a language barrier in trade has been documented in numerous empirical studies. Rose (2000) finds that countries sharing a common language trade 1.5 times more with each other. Anderson and van Windcoop (2004) estimate that the tax equivalent of the language barrier amounts to seven percent.

While gravity models of aggregate trade flows find robust evidence for the language barrier, these models remain silent on the question of the channel through which language affects trade. It is even questionable if language should affect international trade at all, given that international trade flows consist mainly of manufactures. Yet in order to trade two manufacturing goods between for example the US and China only one translator is required, whose services are unlikely to affect total trading cost. Also the fact that with China and Japan two countries with relatively few fluent English speakers are among the top five trading nations contradicts the importance of language for manufacturing trade. Services trade, on the other hand, often requires the ability of both the service provider and his customer to communicate directly with each other.

A second shortcoming of the studies mentioned above is their opaque measurement of the language barrier, which is typically represented as either a binary indicator for countries that share a common official language (e.g. Frankel and Rose 2002), or as

[1] Recent examples are Rauch (2002), Nunn (2007), and Levchenko (2007).

the probability that two randomly chosen people from two countries share a common mother tongue (e.g. Melitz, 2008). Alternatively, Hutchinson (2002) and Ku and Zussman (2008) suggest to use the fluency in English - the lingua franca of international trade - as a proxy for the ability of natives from two countries to communicate in a common third language. Yet what is really required for trade is that there is a sufficient number of people in both countries who are proficient in at least one of the other's language(s), irrespective of whether they speak a lingua franca, the same official, native, or second language. Also, the common proxies might take up all other kinds of bilateral institutional similarities, thereby imposing an upward bias on the estimate for the language barrier in gravity models.

This paper provides one way to resolve the missing motivation of the language barrier and to reduce measurement bias of the effect of language on trade. In particular, I test if communication-intensive industries trade more between Canadian provinces with a good knowledge of the other's language(s) compared to those industries that require less communication with the trading partner. Such a finding could justify the alleged role of language as a trade barrier. Though it is less general than conventional gravity models, this simple approach has two advantages: First, it tests for one specific mechanism through which language affects trade. Second, it corrects for other institutional factors that could bias the estimates via fixed-bilateral effects between Canadian provinces.

Previous work that comes closest to this paper is from Fink et al. (2005), who show that trade is significantly lower between countries with high bilateral international calling prices. They find that the price effect is larger for trade in differentiated products compared to goods that are traded over organized exchanges, which corroborates the hypothesis that trade in communication-intensive goods is more sensitive to deficiencies in direct communication. However, they estimate that halving the importer's calling prices would boost aggregate trade by 42.5%, which seems unreasonably high. Melitz (2008) estimates the effect of sharing a common mother tongue on international trade flows. In contrast to my paper, Melitz's variables on language commonality do not measure the knowledge of second languages. My measure incorporates the two-sided knowledge of English, French, and Chinese as first or second languages between Canadian provinces, which is a better proxy for the language-trade link, as the empirical evidence in section 5.3 shows. So far, only Fidrmuc and Fidrmuc (2008) use data on the actual knowledge of foreign languages in the Europe. Yet their estimates are based on aggregate trade flows, so that they cannot attribute the effects to a specific channel through which language erects a trade barrier.

My results suggest that commerce in industries that require direct communication for trade increases with the probability that people in another Canadian province speak the same language. I cannot find evidence for an impact of indirect communication via mail or advertising on intra-Canadian trade flows. The estimates imply that Canada's minority language regions face a potential burden from expensive services exports and imports.

The remainder of the paper is structured as follows: In section 5.2 the estimation equation is motivated. Section 5.3 describes the data. In section 5.4, the baseline results are discussed. In section 5.5, I control for potential endogeneity. Section 5.6 consists of robustness checks. Finally, section 5.7 wraps up the discussion. A detailed description of the variable labels is provided in table 5.8.

5.2 Empirical Model

While there is strong support for the language barrier in empirical research, hardly any theoretical work has analyzed this issue, probably because it seems self-evident that people can only trade if they are able to communicate with each other. Yet the case for language in trade is not clear-cut: For instance, while rice or oil can be bought at the merchandise exchange without the need to learn any Asian or Arabic languages, a buyer of a laptop will require explanations, software, and support services in a language he speaks. To see exactly how language can affect trade patterns, imagine the following scenario: There are two regions, whose populations speak different languages. Translation is costly. If some products require more communication between buyer and seller for trade to proceed, translation cost will more adversely affect trade in those products. If more people learn the other's language, total translation cost will fall. Hence, I propose the following hypothesis: ceteris paribus, a high language commonality between two regions should disproportionately help communication-intensive industries to trade.

With respect to the type of communication used, I distinguish between direct or spoken communication and indirect or written communication. Direct communication is expected to have a larger effect on the volume of trade than indirect communication, because total translation costs are higher for direct communication, which cannot be replicated but has to occur simultaneously. Hence, direct communication-intensive industries are more likely to be affected by the language barrier.

This study focuses on Canada, which is the only OECD country with more than one official language for which detailed inter-regional data on trade flows is available. While this choice limits the scope of the study and the number of potential sources of language variation, it offers at least three advantages. Firstly, the relative uniformity of Canada's legal and social system alleviates institutional bias that is possibly present in studies of international trade. As communication-intensive industries are often contract-intensive as well, estimates from cross-country regressions would be likely to incorporate effects of comparative advantage in regions with sound legal institutions (Nunn 2007). Secondly, it downweights the possibility that my language estimates capture some home bias (or border) effects that are well-known to the international trade literature.[2] Hummels and Hillberry (2003) showed that intra-US trade is unlikely to suffer from intra-national border effects.[3] Similarly, Combes et al. (2003) estimate that in France more than 60% of the potential intra-national home bias can be explained by internal migration and cultural networks. Such network effects between Canadian provinces and territories are likely to be primarily determined by linguistic differences, since Helliwell (1997) already pointed out that internal migration has little trade creating effect within Canada. Finally, the arguments presented above for the existence of a language-trade channel should be mainly relevant for service-intensive industries. Therefore I refrain from studying intra-European trade (which otherwise would make a perfect case for the language-trade link), because services are not sufficiently liberalized across EU members (e.g. Kox and Lejour 2005; Kox and Lejour 2006).

This paper introduces a new way to thinking about the gravity model of trade, which rests on the work by Rajan and Zingales (1998), Romalis (2004), and Nunn (2007). These papers use industry- and cross-country-variation to identify sources of a country's comparative advantage across industries. I adapt their approach for a single country setting, where I exploit trade variation across industries and bilateral province pairings. Specifically, I eliminate any variation, which is not needed to test the main hypothesis: trade in communication-intensive industries is higher between provinces with a higher language commonality. The model I estimate is then:

$$\ln trade_{ijk} = \delta_{ij} + \delta_k + \beta_1 \ln(trans_k dist_{ij}) + \beta_2 prod_{ijk} + \beta_3 c_k lang_{ij} + \epsilon_{ijk}, \qquad (5.1)$$

where $\ln trade_{ijk}$ is the natural logarithm of the bilateral trade flow from province i to province j in industry k. The fixed-bilateral effects δ_{ij} pick up all trade variation for each country pairing that is constant across industries. Similarly, δ_k are industry fixed-effects that are constant across bilateral trade flows. Compared to the traditional gravity model, the bilateral fixed-effects do not allow to use variables that are constant across country pairings. Thus, the impact of distance on trade is

[2] See McCullum (1995), Helliwell (1996).

[3] Wolf's (2000) dataset does not properly account for intra-US trade distances and wholesale trade flows.

proxied with the log of the interaction between the transport-intensity of a sector, $trans_k$, and distance, $dist_{ij}$. In order to capture differences in comparative advantage, $prod_{ijk} = \frac{production_{ik}}{GDP_i} - \frac{production_{jk}}{GDP_j}$ reflects differences in the structure of production between two provinces. The main variable of interest is the interaction of c_k and $lang_{ij}$, in which c_k reflects the need for communication and $lang_{ij}$ stands for the language commonality between two provinces. ϵ_{ijk} is a random error. As common in the literature (e.g. Anderson and van Windcoop 2003; Melitz 2008), I assume that imports and exports are affected symmetrically by the interaction effects.

The approach here is conceptually different from industry-level gravity models that estimate the semi-elasticity of the language commonality with respect to trade (e.g. Deardorff 1998; Hummels 2001). The bilateral fixed-effects capture the direct effect of the language commonality on the volume of trade in my estimation equation. Hence, the coefficient of interest β_3 only captures the effect that language commonality has on the pattern of trade and provides no direct interpretation as a semi-elasticity of the language barrier.

The estimates of (5.1) should not be regarded as conclusive evidence for the language-trade channel. First, there may be determinants of trade that are omitted from (5.1). As a matter of fact Canada's English speaking provinces tend to be richer and domicile more Protestants than Catholics compared to their French speaking counterparts. Therefore, a primary concern is that $c_k lang_{ij}$ may be simply capturing the fact that wealth and religion shape intra-Canadian trade patterns. I carefully control for these alternative determinants of the language-trade channel. Second, the direction of causality implied by equation (5.1) may be wrong. If trade fosters the adoption of the other's language, causality might run from trade to language. In consequence, estimates of β_3 may be biased. In section 5.5, I instrument for language variation that is unaffected by this feedback effect. Finally, this paper concentrates on the analysis of positive exports and imports. Thus the interpretation of the estimates is conditional on a province trading in an industry, thereby disregarding the effect of language on the decision to enter an industry. I check for the sensitivity of my results to the inclusion of zero trade in section 5.6.

5.3 The Data

The most disaggregated inter-provincial trade data available for Canada are at the 2-digit industry level. The data comprise all recorded (non-zero) inter-regional trade flows of Canada's ten provinces and three territories for the year 2001.[4] The final data classify in 38 industries that comprise agriculture, manufacturing and service industries. For numerous reasons, I study trade across all sectors, which is different from other studies that solely focus on manufacturing trade (e.g. Nunn 2007; Romalis 2004; Hummels 2001). The first reason is that Canada's internal trade differs from international trade, where services trade is negligible compared to manufactures. In 2001, service trade accounted for 56.7% of total intra-Canadian trade flows. The second reason is that the language channel should be present across all sectors of the economy. Particularly, the service sector is likely to be more language-sensitive than manufactures. So leaving out one of the sectors would narrow the scope of this study.

Provincial gross domestic products in current prices as well as population estimates have been retrieved from the Statistics Canada home-page. The distance variable is from Feenstra (2004), who provides distances between the capitals of Canadian provinces. I added distances for each pairing that involves trade with the three territories, using the respective longitudes and latitudes.

5.3.1 Language Variables

In contrast to the language proxies used in previous studies, this paper measures language commonality as the probability that any two people from different provinces picked at random will be able to communicate with each other.[5] This variable is more in line with theory, because trade only requires a sufficient knowledge of the trading partner's language in order to reduce translation costs.

The measure of language commonality between provinces is constructed from the Census survey. The survey asks for mother tongue, knowledge of official languages, and use of languages at work. Table 5.1 depicts the percentage of speakers of English and French as a mother tongue in Canadian provinces. While English is the dominant mother tongue (59.5%), 22.7% of the Canadian population are native French speakers. French mother tongue speakers, are mainly concentrated in Quebec (81.2% French mother tongue speakers) and New Brunswick (32.9%).

The statistics show clearly that the language barrier cannot be represented by the distribution of mother tongues within the population: 17.6% of all Canadians have a mother tongue different from the two official languages. Yet only 1.5% of all Canadians

[4]Although industry-level trade data is available from 1997 to 2004, Census data is only available for the years 1996 and 2001. For a discussion of the derivation of inter-provincial trade flows from IO-tables, see Généreux and Langen (2002).

[5]For an excellent overview of language measures used in previous research, see Melitz (2008).

are unable to speak at least one of the two official languages. Across provinces, only the Inuit population in the Nunavut territories constitutes an exception with 13.1% of the population knowing neither English nor French. While in most provinces more than 97% of the population speaks English as first or second language, Quebec (45.4%) and New Brunswick (90.7%) are the two exceptions with relatively few English speakers.[6] In addition, I control for potential Chinese networks within Canada that have been shown to affect international trade flows (Rauch 2002). With 2.7%, Canada's Chinese minority supplies the third largest language group of Canada's working population. Different from other minorities, 42% of speakers with a 'Chinese' mother tongue also use non-official languages often or sometimes at work. As a proxy for the knowledge of a Chinese language, I use the population share with Chinese origin.

For each bilateral pairing ij, the variable for language commonality is constructed as follows :

$$lang_{ij} = \sum_{l=1}^{L} (knowledge^l)_i (knowledge^l)_j, \tag{5.2}$$

where $l = \{English, French, 'Chinese'\}$. $English$ is the sum of people knowing English and people knowing English and French, divided by the total population of the province. The $French$ and $'Chinese'$ measures are constructed similarly. Due to data limitations and the fact that indigenous languages are not used outside the three territories, I do not consider native languages separately. $lang_{ij}$ is not bounded at one, since people may be fluent in several languages. However, I restrict the probability that two randomly chosen people are able to communicate with each other to one in cases where I calculate values slightly larger than one. Based on equation (5.2), I also construct a measure for religious (denominational) commonality, where l={*Anglicans, Baptists, Buddhists, Catholics, Hindus, Muslims, Jews, United church*}.

Table 5.1 depicts the resulting language commonality for all bilateral country pairings. The pairings range from Quebec and Nunavut, where the probability that two randomly chosen people understand each other is 43.1%, to Saskatchewan and Prince Edward Island, where everybody speaks the same language. Virtually all variation in language knowledge comes from the two French speaking provinces and the territories, whereas the pairing Ontario/British Columbia exhibits the lowest language commonality (0.962) among the English speaking provinces.

[6]Unfortunately, the data do not allow to draw explicit conclusions on the fluency of language knowledge. Yet Hutchinson (2002) cannot find a statistically significant difference between speakers of English as a mother tongue or second language, when analyzing the volume of US exports and imports.

5.3.2 Communication-intensities of Industries

Identification hinges on a careful choice of c_k, the proxy for differences in the industry-specific need of direct and indirect communication between importer and exporter. Rauch (1999) classifies manufacturing goods on whether they are traded on an exchange, reference priced or neither of both. However, I refrain from using his classification for two reasons. Firstly, it only captures manufacturing goods, yet the services sector is an important pillar of intra-Canadian trade that accounts for much of the language-related variation, as will be seen below. Secondly, it is not possible to extend his classification to services, because services are typically neither reference priced nor traded on exchanges.[7]

I construct a new measure for the communication-intensity of industries that takes advantage of detailed input output (IO) tables. IO data are available for the manufacturing as well as the service sector and allow me to rank all industries according to their need for communication between trading partners. Thereby I implicitly assume that the input structure of communication services proxies the need for direct and indirect communication between exporter and importer. Given the relatively high level of aggregation of the trade data, all this assumption postulates is that if the printing industry needs a larger share of communication inputs than the paper industry relative to its total inputs, trading printing products also requires more communication for trade. I measure the direct communication-intensity by the share of telecommunications services in total inputs for each industry. For the industry-specific need to communicate indirectly via written language, I employ two measures: the input share of post services and the input share of promotion services (i.e. advertising and entertainment inputs). Since the IO tables at M-level aggregation (2-digit level) exist only for Canada as a whole, I assume that the average Canadian input structure persists across provinces. This strong assumption is less problematic in the Canadian case, where production structures are relatively similar, than in cross-country studies (e.g. Rajan and Zingales, 1998; Nunn, 2007).

The resulting ranking seems reasonable by common sense, as can be seen in table 5.3. The share of telecommunication inputs ranges from 0.05% (Fishery) to 5.3% (Professional services). As expected, table 5.3 shows that service industries are more communication-intensive than manufacturing industries. Among manufacturing industries, more complex products are generally ranked higher, which is consistent with

[7]Experiments with the Rauch data proved inconclusive. I manually matched classifications and calculated the percentage of goods that is neither reference priced nor traded on public exchanges for each industry. However, the estimated effects are only significant if trade flows in the (language-insensitive) petroleum and coal industry are included into the sample.

the Rauch (1999) classification, where more complex manufactures such as electronic equipment are rarely reference-priced or traded on an exchange. Also with respect to post service inputs, services are more communication-intensive than manufactures. The input share of promotion services differs from the other two measures.

Finally, industries are ranked according to their relative cost of transportation, presupposing that distance has a larger trade diverting impact on industries with higher transport cost. I calculate the share of transport margins in total inputs, which is defined as the charges paid to a third party in order to deliver a product from the producer to the (intermediate or final) purchaser. The ranking of the transport variable in table 5.3 shows that services generally have lower transportation cost than manufactures. Particularly heavy industries rank high, e.g. metal, mineral products, chemical, and motor vehicle industries. If not stated differently, I drop the fuel as well as the petroleum and coal industries from the sample. As these industries are unlikely to be sensitive to language, the high trade volumes in both industries would bias the estimates downward (which can be seen in table 5.6).

5.4 Empirical Results

Columns (1), (4), and (7) of table 5.4 report simple correlations between the interactions of telecommunication, post, and promotion services with the log of the bilateral trade volume. All three interactions exhibit a significant positive correlation, which provides preliminary evidence in favor of the proposed language-trade channel. Estimates of equation (5.1) are reported in columns (3), (6), and (9). As expected, the transport-distance interaction has a significant negative impact on inter-provincial trade. The intuition for this estimate is that trade with distant provinces is particularly low for transport-intensive industries. Moreover, I find that specialization affects trade positively, which is in accordance with standard trade theory. The estimate implies an average impact of $prod_{ijk}$ on trade of 39%, given the standard deviation of 0.045 of the production differential within an industry.[8] The interaction between language commonality and telecommunication-intensity is statistically significant at the 5% level. This implies an increase of trade volume by 2.64% for an average communication-intensive industry, when shifting from the 25^{th} to the 75^{th} percentile of the distribution of $lang_{ij}$. For a service-intensive industry such as health, this effect would correspond

[8] I calculated the effect as $\%\Delta trade_{ijk} = 100 * \beta_2 * 0.045 = 100 * 8.647 * 0.045 = 38.91\%$.

to an increase of trade by 6.91%.[9] There is, however, less evidence for the presence of an indirect communication channel. The $post_k lang_{ij}$ variable in column (6) is statistically and economically insignificant. Similarly, the interaction between promotion services and language commonality is statistically insignificant in column (9), once I control for bilateral- and industry-fixed-effects. This indicates that industries that rely on direct interaction in order to export their products trade more between regions with a high language commonality, whereas indirect communication seems to play a lesser role.

I also test if the estimated effects are only due to variation in the industry dimension or result from joint variation of the interaction effect across industries and bilateral pairings. Columns (2) and (3) show that the coefficient of the interaction term is of similar size and also statistically significant at the 10% level if I include both $telecom_k lang_{ij}$ and $telecom_k$. With respect to post and promotion services, however, most variation seems to result from the communication-intensities $post_k$ and $promo_k$.

The estimates of β_3 might be biased if determinants of trade have been omitted from (5.1) that are correlated with the explanatory variables. Warren (2003) argues that the economic development of the French speaking provinces was retarded. Within Quebec most businesses were in the hands of an English speaking minority before strong French-promoting legislation was passed in the 1970s. If Canada's English speaking population were more affluent, all I capture with the language interaction would be a wealth effect. Hence, I control for the interaction between telecommunication and joint provincial GDP per capita in column (1) of table 5.5. The insignificance of the estimate and the fact that the estimated β_3 remains practically unchanged indicate that I am really capturing language effects.

Another reason for bias of β_3 could be that other institutional variables that are correlated with language have been omitted from equation (5.1). It could be that the foremost Catholic population in French-speaking Canada distrusts Protestant business partners or exhibits different demand patterns. If this were the case, the alleged language effect would really capture religious affiliation. Although Lipset (1990) argues that religion has a smaller role in Canadian everyday life than in the US, religious commonality has been shown to affect international trade patterns (e.g. Lewer and van den Berg, 2007; Helble, 2007). Hence, I control for the probability that two randomly chosen people from two states have the same denomination. The religion measure is highly correlated with language commonality (0.72). Yet the estimate in column (2) is

[9] These numbers have been calculated for the pairings NL-NU (25^{th} percentile) and ON-PE (75^{th} percentile), where the effect for an average industry is $\%\Delta trade_{ijk} = 100 * \beta_3 * \overline{telecom}_{ijk} * (lang_{ij}^{75} - lang_{ij}^{25}) = 100 * 20.23 * 0.0107 * (0.989 - 0.868)$.

insignificant, while β_3 remains a significant explanatory variable of trade. The fact that the estimate still is of similar magnitude is evidence in favor of the language channel.[10]

5.5 IV Results

Although the approach taken here reduces several potential sources of bias that are present in standard gravity models, the estimate of β_3 could still be subject to endogeneity. I deal with this issue using legal language status as an instrumental variable (IV).

While Canada's Official Language Act of 1969 guarantees equal legal status of both English and French with respect to federal administrative services, federal courts, and in Parliament, some provinces enacted additional language laws. Particularly, Quebec and New Brunswick passed own Official Language Acts during the 1970s that promote the use of French at the work place, in educational institutions, and for administrative procedures. The *Official Languages of New Brunswick Act* was first enacted in 1973 and later on revised. Likewise, Amendment 16.1 of the Canadian constitution, which was enacted in 1993, reinforces the equal status of the French language in New Brunswick. Quebec passed the *Official Language Act* (Bill 22) in 1974 and the *Charter of the French Language* (Bill 101) in 1977. Warren (2003) argues that these laws triggered a revival of the French language in everyday life and also in business, where English was to become the primary language in the 1970s. Moreover, Lazear (1999) shows that the protection of minority interests by the government reduces incentives to learn the majority language, implying lower knowledge of English in those regions that guarantee specific language rights.

Therefore I use the legal language status across provinces as an instrument for the probability that two people from two provinces speak the same language. In particular, I use the interaction $c_k legal_{ij}$ as an instrument for $c_k lang_{ij}$, where c_k is assumed to be exogenous. As legal language status is predetermined and unaffected by the trade flow in 2001, it is a suitable instrument to isolate exogenous variation in language commonality. The variable $legal_{ij}$ is a dummy, which is one if Quebec or New Brunswick are a trading partner in a bilateral pairing, two for trade flows between these two provinces, and zero otherwise.

[10] Another variable that could be correlated with language commonality is ethnic origin. However, the data do not allow to disentangle ethnic origin and language ties for French Canadians. Similarly, constructing an aggregate measure along the lines of equation (5.2) will not yield a consistent proxy for ethnicity, because large ethnic groups within Canada have ethnic origins that are unlikely to affect trading patterns, e.g. English, Irish, Scottish or Welsh.

The IV estimates are reported in table 5.5. I only report second stage estimates. The statistics from the first stage regressions indicate that the IV estimator can be used. Columns (3) to (6) report large F statistics and high partial R^2s of the first stage regressions. Also, the instrument $c_k legal_{ij}$ is significantly partially correlated with $c_k lang_{ij}$ in the first stage regressions. The IV estimate of $lang_{ij} telecom_k$ in (3) is positive and statistically significant, supporting the hypothesis of a language-trade channel. However, the estimate is larger than the OLS estimate in table 5.4, not smaller, as the potential reverse causality suggests. This could indicate a weak instrument, yet neither the partial correlations, nor the t-statistic suggest presence of a weak correlation. Also the Cragg and Donald (1993) test for weak instruments rejects the hypothesis that the equation is only weakly identified. To test for local average treatment effects, I drop the three territories from the sample, since they are partly inhabited by natives. This reduces much of the language variation that cannot be attributed to laws affirming the use of French in business. The estimate in (4) decreases to a value 23.91 which is close to the OLS estimates. This corresponds to an increase of trade in the health industry by 0.26% if $lang_{ij}$ increases by 1%.[11]

As with OLS, the IV estimate of the post service interaction is not significant. The coefficient for $promo_{ij} telecom_k$ is positive and significant but too high compared to the OLS estimate.

5.6 Robustness

The following section tests the sensitivity of my results to the choice of the sample and potential bias of the estimates due to the focus on positive trade flows. Table 5.6 reports the sensitivity of the estimated coefficients to the removal of influential observations and the choice of the language variable. The estimates of β_3 are obtained from separate regressions of equation (5.1), using one interaction term in each run. The upper third of table 5.6 reports coefficients for the interactions between direct communication-intensities and language commonalities. The middle third reports coefficients for interactions with the input share of post services. The bottom of the table reports interactions with the promotion services-intensity. Direct communication interactions are statistically significant in all models. Also the 1998 sample confirms the previous results using Census data on language and ethnicity variables from 1996. The estimates for the indirect communication channel are in line with the preceding

[11]The calculation of the partial derivative with respect to $lang_{ij}$ yields: $\frac{\partial \ln trade_{ijk}}{\partial lang_{ij}} = \beta_3 * telecom_k = 23.91 * 0.0107 = 0.26$.

results: A higher language commonality has neither a significant impact on trade in post service-intensive industries nor in promotion-service-intensive industries.

As a final test, I check if the above analysis is sensitive to the exclusion of zero trade flows from the sample. In order to account for zero trade flows, I set all observations for which trade flows are not reported to zero. However, a log-transformation of zero values is not possible. Several methods have been suggested to deal with this issue. Sample selection procedures would probably be the most elegant way to adjust the estimates for zero observations. However, the estimation of sample selection models requires that at least one independent variable explains the selection process but is not partially correlated with the dependent variable (in order not to rely on distributional assumptions). It is hardly possible to find such a variable for the trade data used here.

To get around the selection problem, Silva and Tenreyro (2006) suggest the use of the Poisson model for gravity equations. The Poisson estimator uses all positive and zero observations in a way that allows to interpret the coefficients similar to gravity estimates. Although it is typically used for count data, the Poisson estimator is consistent as long as the mean function is correctly specified. Helpman et al. (2008) find that a Poisson regression yields estimates that are comparable to a sample selection procedure.

Column (1) of table 5.7 reports regression results for all combinations of provinces and territories across all industries. The estimates of the interaction terms are of similar magnitude compared to the fixed-effects estimates. In columns (2)-(4), I drop the fuel as well as the petroleum and coal industries from the sample for comparative purposes. Now the magnitude of the language interaction resembles the IV estimate from table 5.5.

While most estimates of the interaction terms are of similar size as with the fixed-effects estimator, the inclusion of zero trade flows has a strong effect on the distance estimate. The estimated trade barrier of distance is more than twice as large. This indicates that the predominant reason not to enter a trade relationship with another province is transport cost.

5.7 Conclusion

In this paper, I find robust evidence for one mechanism that could justify the empirical evidence for the language barrier to trade in gravity models. Industries that require more communication with the business partner in order to trade their products, trade more between Canadian provinces with a high proportion of same-language speakers.

This language channel appears to depend on direct (spoken) communication rather than indirect communication via mail or advertising. This is in line with Fink et al. (2005), who demonstrate the importance of international calling prices for the volume of bilateral trade. Finally, the significant negative relationship between the volume of trade and the distance-transport cost interaction holds potential for future applications of this methodology.

The results indicate that the French-speaking parts of Canada do not only have a comparative disadvantage in providing communication-intensive services to the rest of the country, but also face a higher burden on imports of such services. This might expose the entire economy of minority language regions to higher cost of acquiring up-to-date services from English-speaking providers. Hence, language might turn out to be a source of comparative advantage or disadvantage that allows regions with a higher share of majority language speakers to specialize in more advanced goods and services. Future research might study in how far this language-trade channel also applies to international trade. It is likely that language will prove to be an impediment to trade in services and complex goods that require direct communication with the foreign importer. Such a finding would suggest that developing countries with few English-speakers will find it hard to develop competitive services industries in the future.

References

Anderson, James E. and Eric van Wincoop (2003): "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, **93:1**, 170-192.

Anderson, James E. and Eric van Wincoop (2004): "Trade Costs," *Journal of Economic Literature*, **42:3**, 691-751.

Combes, Pierre-Philippe; Miren Lafourcade and Thierry Mayer (2003): "Can Business and Social Networks Explain the Border Effect Puzzle?," CEPR Discussion Paper, 3750.

Cragg, John G. and Stephen G. Donald (1993): "Testing Identifiability and Specification in Instrumental Variables Models," *Econometric Theory*, **9:2**, 222-240.

Deardorff, Alan V. (1998): "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?," in *The Regionalization of the World Economy*, Jeffrey A. Frankel. University of Chicago Press: Chicago, 7-32.

Fidrmuc, Jan and Jarko Fidrmuc (2008): "Foreign Languages and Trade," CEPR Discussion Paper 7228.

Fink, Carsten, Aaditya Mattoo and Ileana Cristina Neagu (2005): "Assessing the Impact of Communication Cost on International Trade," *Journal of International Economics*, **67**, 428-445.

Frankel, Jeffrey A. and Andrew Rose (2002): "An Estimate of the Effect of Common Currencies on Trade and Income," *Quarterly Journal of Economics*, **117:2**, 437-466.

Généreux, Pierre A. and Brent Langen (2002): "The Derivation of Provincial (Inter-regional) Trade Flows: The Canadian Experience," Statistics Canada Working Paper, 98.

Helble, Matthias (2007): "Is God Good for Trade?," *Kyklos*, **60:3**, 385-413.

Helliwell, John F. (1996): "Do National Borders Matter for Quebec's Trade?," *Canadian Journal of Economics*, **29:3**, 507-522.

Helliwell, John F. (1997): "National Borders, Trade and Migration," *Pacific Economic Review*, **2:3**, 165-185.

Helpman, Elhanan; Marc Melitz and Yona Rubinstein (2008): "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of*

Economics, **123:3**, 441-487.

Hummels, David (2001): "Towards a Geography of Trade Costs," mimeo, University of Chicago.

Hummels, David and Russell Hillberry (2003): "Intranational Home Bias: Some Explanations," *The Review of Economics and Statistics*, **85:4**, 1089-1092.

Hutchinson, William K. (2002): "Does Ease of Communication Increase Trade? Commonality of Language and Bilateral Trade," *Scottish Journal of Political Economy*, **49:5**, 544-556.

Lipset, Seymour Martin (1990): *Continental Divide. The Values and Institutions of the United States and Canada*. Routledge: London.

McCallum, John (1995): "National Borders Matter: Canada-U.S. Regional Trade Patterns," *American Economic Review*, **83:3**, 615-623.

Kox, Henk and Arjan Lejour (2005): "Regulatory Heterogeneity as Obstacle for International Services Trade," CPB Discussion Paper, 49.

Kox, Henk and Arjan Lejour (2006): "The Effects of the Services Directive on Intra-EU Trade and FDI," *Revue Economique*, **57:4**, 747-769.

Ku, Hyejin and Asaf Zussman (2008): "Lingua Franca: The Role of English in International Trade," mimeo, Cornell University.

Lazear, Edward (1999): "Culture and Language," *Journal of Political Economy*, **107:6**, 95-125.

Levchenko, Andrei (2007): "Institutional Quality and International Trade," *Review of Economic Studies*, **74:3**, 791-819.

Lewer, Joshua J. and Hendrik van den Berg (2007): "Estimating the Institutional and Network Effects of Religious Cultures on International Trade," *Kyklos*, **60:2**, 255-277.

Melitz, Jacques (2008): "Language and Foreign Trade," *European Economic Review*, **52:4**, 667-699.

Nunn, Nathan (2007): "Relationship-Specific, Incomplete Contracts, and the Pattern of Trade," *Quarterly Journal of Economics*, **122:2**, 569-600.

Rajan, Raghuram G. and Luigi Zingales (1998): "Financial Dependence and Growth," *American Economic Review*, **88:4**, 559-586.

Rauch, James E. (1999): "Networks versus Markets in International Trade," *Journal* *of International Economics*, **48:1**, 7-35.

RAUCH, JAMES E. and VITOR TRINDADE (2002): "Ethnic Chinese Networks in International Trade," *The Review of Economics and Statistics*, **84:1**, 116-130.

ROSE, ANDREW K. (2001): "One Money, One Market: The Effect of Common Currencies on Trade," *Economic Policy*, **30:1**, 7-46.

SILVA, J. M. C. SANTOS and SILVANA TENREYRO (2006): "The Log of Gravity," *The Review of Economics and Statistics*, **88:4**, 641-658.

WAGNER, DON; KEITH HEAD and JOHN RIES (2002): "Immigration and the Trade of Provinces," *Scottish Journal of Political Economy*, **49:5**, 507-525.

WARREN, JEAN-PHILIPPE (2003): "The History of Quebec in the Perspective of the French Language," in *Linguistic Conflict and Language Laws. Understanding the Quebec Question*, Larrivée, Pierre. Palgrave Macmillan: Houndmills, 57-86.

WOLF, HOLGER. C. (2000): "Intranational Home bias in Trade," *The Review of Economics and Statistics*, **82:4**, 555-563.

Tables

Table 5.1: Mother tongues and knowledge of official languages

	Mother Tongues				Knowledge of Official Languages				Total Population
	English	French	English and French	Other	English only	French only	English and French	None	
Canada	59.3%	22.7%	0.4%	17.6%	67.50%	13.30%	17.70%	1.50%	30,007,094
Newfoundland and Labrador (NL)	98.4%	0.4%	0.1%	1.1%	95.70%	0.00%	4.10%	0.10%	512,930
Prince Edward Island (PE)	93.9%	4.3%	0.3%	1.5%	87.90%	0.10%	12.00%	0.00%	135,294
Nova Scotia (NS)	93.0%	3.8%	0.3%	3.0%	89.70%	0.10%	10.10%	0.10%	908,007
New Brunswick (NB)	64.7%	32.9%	0.7%	1.7%	56.50%	9.20%	34.20%	0.10%	729,498
Quebec (QC)	8.0%	81.2%	0.8%	10.0%	4.60%	53.80%	40.80%	0.80%	7,237,479
Ontario (ON)	71.6%	4.4%	0.4%	23.7%	85.90%	0.40%	11.70%	2.10%	11,410,046
Manitoba (MB)	75.8%	4.1%	0.3%	19.9%	89.70%	0.10%	9.30%	0.80%	1,119,583
Saskatchewan (SK)	85.7%	1.9%	0.2%	12.2%	94.50%	0.00%	5.10%	0.30%	978,933
Alberta (AB)	81.8%	2.0%	0.2%	16.0%	92.00%	0.10%	6.90%	1.10%	2,974,807
British Columbia (BC)	74.1%	1.5%	0.2%	24.3%	90.30%	0.00%	7.00%	2.70%	3,907,738
Yukon Territory (YT)	87.1%	3.1%	0.3%	9.5%	89.40%	0.20%	10.10%	0.30%	28,674
Northwest Territories (NT)	78.1%	2.6%	0.2%	19.0%	90.40%	0.10%	8.40%	1.00%	37,360
Nunavut (NU)	27.6%	1.5%	0.1%	70.8%	83.00%	0.10%	3.80%	13.10%	26,745

Table 5.2: Bilateral language commonality, 2001

$lang_{ij}$	Exporter	Importer	Mean Trade	Mean Exports	Mean Imports
0.431	QC	NU	3.56	4.25	0.95
0.492	QC	NL	24.57	30.49	18.31
0.501	SK	QC	29.64	21.05	37.96
0.509	QC	BC	91.80	111.13	71.30
0.515	QC	AB	112.26	122.93	101.59
0.530	QC	NT	5.64	6.79	2.09
0.539	QC	MB	42.85	40.31	45.72
0.549	QC	NS	41.98	51.03	32.09
0.550	YT	QC	1.20	0.60	1.41
0.557	QC	ON	838.45	774.08	901.12
0.568	QC	PE	7.26	8.69	5.45
0.804	NU	NB	0.49	0.10	0.62
0.822	QC	NB	64.18	64.05	64.32
0.847	NU	BC	1.35	0.68	1.55
0.851	ON	NU	4.13	4.86	1.70
0.861	NU	AB	3.21	0.95	3.94
0.861	NU	NT	2.99	1.03	4.95
0.863	NU	MB	0.81	0.50	0.92
0.867	SK	NU	0.33	0.34	0.24
0.868	YT	NU	0.95	1.78	0.13
0.868	NU	NL	0.34	0.25	0.37
0.870	NU	NS	1.01	0.34	1.28
0.872	PE	NU	0.30	0.20	0.60
0.913	NB	BC	5.38	5.17	5.59
0.924	NL	NB	17.05	20.65	13.79
0.926	SK	NB	1.83	1.65	2.03
0.927	NB	AB	5.83	4.92	6.73
0.934	NT	NB	1.23	0.20	1.43
0.938	ON	NB	61.90	86.32	35.27
0.939	NB	MB	3.11	3.00	3.22
0.948	YT	NB	0.58	0.15	0.65
0.949	NS	NB	32.15	31.90	32.38
0.958	PE	NB	8.84	6.42	11.05
0.962	ON	BC	276.05	373.43	175.81
0.968	NT	BC	8.71	15.03	5.65
0.970	BC	AB	230.30	226.31	234.39
0.971	MB	BC	30.89	34.11	27.77
0.974	SK	BC	32.15	28.39	35.91
0.974	NL	BC	3.73	2.47	4.55
0.975	ON	AB	410.28	463.42	355.57
0.975	ON	NT	18.92	13.35	34.68
0.977	YT	BC	6.02	7.51	5.07
0.978	NS	BC	7.86	6.94	8.77
0.978	ON	MB	120.77	135.62	104.65
0.979	SK	ON	108.92	92.42	126.39
0.979	ON	NL	52.53	68.39	35.14
0.980	PE	BC	1.13	0.62	1.46
0.984	NT	AB	9.57	2.61	13.49
0.984	YT	ON	4.02	2.40	4.60
0.986	ON	NS	81.28	109.05	50.91
0.986	MB	AB	78.38	67.17	88.96
0.987	NT	MB	1.50	0.82	1.76
0.989	SK	AB	110.14	85.05	135.23
0.989	PE	ON	14.67	9.64	18.17
0.989	SK	NT	0.83	0.85	0.78
0.990	NL	AB	8.42	2.70	12.30
0.991	NT	NL	0.39	0.23	0.44
0.992	SK	MB	36.57	35.92	37.31
0.992	YT	AB	2.90	1.28	3.60
0.993	NL	MB	1.91	1.37	2.20
0.993	YT	NT	1.21	2.25	0.28
0.994	NS	AB	11.27	10.11	12.36
0.995	NS	NT	1.84	2.64	0.34
0.996	YT	MB	0.42	0.41	0.42
0.996	PE	AB	1.29	1.14	1.39
0.997	SK	NL	1.39	1.45	1.27
0.997	YT	SK	0.29	0.33	0.28
0.998	PE	NT	0.23	0.23	0.20
0.998	NS	MB	4.12	3.86	4.38
0.998	YT	NL	0.10	0.05	0.11
1.000	SK	NS	3.07	2.97	3.18
1.000	PE	MB	0.95	1.27	0.74
1.000	PE	NL	1.98	2.58	1.43
1.000	NS	NL	15.29	19.88	10.84
1.000	YT	NS	0.29	0.13	0.36
1.000	PE	NS	7.54	5.20	9.80
1.000	YT	PE	0.20	-	0.20
1.000	SK	PE	0.52	0.63	0.36

Note: Trade is the average bilateral trade across all reported industries in million Canadian \$. $lang_{ij}$ is the probability that two randomly selected people from both regions are able to communicate with each other in English, French, or 'Chinese'. Own calculations.

Table 5.3: Input shares by sector, 2001

Industry	Trade *in million $*	Telecoms *in %*	Post *in %*	Promotion *in %*	Transport *in %*
Fishery	6.78	0.05	-	0.09	0.68
Metal	90.31	0.07	0.02	0.24	2.18
Paper	60.41	0.09	0.05	0.62	3.94
Petroleum and Coal	102.60	0.09	0.01	0.18	0.59
Fuels	680.76	0.10	0.03	0.58	0.11
Lumber and Wood	35.25	0.14	0.04	0.51	2.14
Beverages and Tobacco	17.75	0.15	0.07	5.30	0.76
Residential Construction	74.65	0.15	0.07	0.61	0.76
Leather	29.74	0.19	0.19	1.68	0.94
Textiles	22.58	0.19	0.15	0.96	0.54
Hosiery	29.27	0.22	0.25	2.09	0.26
Fabricated Metal	50.42	0.22	0.11	0.91	1.48
Furniture	27.18	0.23	0.21	1.85	1.09
Mineral products	16.89	0.24	0.11	1.10	2.58
Minerals	8.48	0.24	0.08	1.59	0.70
Ores	77.29	0.24	0.08	1.59	0.70
Machinery	31.58	0.27	0.12	1.50	0.90
Motor vehicles, parts	93.38	0.27	0.14	2.01	1.84
Print	34.58	0.28	0.33	1.08	1.41
Manufactured Products	28.69	0.37	0.39	2.56	0.70
Accommodation and Meals	26.22	0.37	0.07	3.11	0.51
Chemical and Pharmaceutical	94.92	0.38	0.22	4.10	2.33
Forestry	26.02	0.40	0.04	0.79	0.20
Mining services	17.71	0.44	0.06	2.57	0.81
Electronic equipment	48.16	0.57	0.20	3.35	1.27
Grains	27.53	0.63	0.00	0.04	1.05
Fruits	82.46	0.63	0.00	0.04	1.05
Meat	88.13	0.63	0.00	0.04	1.05
other Agriculture	59.06	0.83	0.05	0.94	1.50
Retail	25.63	1.01	1.87	5.11	0.07
Finance and Insurance	143.51	1.07	0.56	3.13	0.02
Utilities	43.39	1.13	0.68	3.70	1.01
Educational services	3.32	1.41	0.71	3.63	0.05
Wholesale	140.05	2.27	1.08	6.14	0.13
Communication services	45.52	2.43	5.17	4.69	0.11
Health	3.70	2.52	0.45	1.58	0.05
other Services	51.25	2.82	2.32	17.44	0.39
Transport and Storage	74.00	4.21	1.22	5.18	0.88
Professional services	152.78	5.30	2.55	13.34	0.69

Note: Trade is the average trade in this industry across all reported bilateral flows. Telecommunication, post, promotion and transportation services inputs shares calculated as the percentage of total inputs in that industry. Own calculations.

Table 5.4: Baseline estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\ln(trans_k dist_{ij})$		0.142***	-0.409***		0.155***	-0.415***		0.105***	-0.410***
		(0.0172)	(0.119)		(0.0181)	(0.123)		(0.0169)	(0.119)
$prod_{ijk}$		8.999***	8.647***		8.620***	7.854***		9.069***	8.639***
		(0.967)	(0.965)		(0.937)	(0.933)		(0.957)	(0.965)
$telecom_k lang_{ij}$	29.25***	16.59*	20.23**						
	(1.826)	(8.983)	(8.943)						
$telecom_k$		15.62*							
		(8.066)							
$post_k lang_{ij}$				22.09***	-2.736	2.654			
				(1.898)	(8.695)	(7.884)			
$post_k$					29.85***				
					(7.731)				
$promo_k lang_{ij}$							7.973***	1.216	2.910
							(0.733)	(3.262)	(2.934)
$promo_k$								7.050**	
								(2.896)	
Industry FE	NO	NO	YES	NO	NO	YES	NO	NO	YES
Bilateral FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	3463	3330	3330	3394	3261	3261	3463	3330	3330
F-statistic	256.4	106.9	60.71	135.5	78.26	63.91	118.4	68.83	61.65
R^2	0.054	0.105	0.412	0.025	0.080	0.411	0.034	0.081	0.411

Note: Note: The estimates are from fixed-effects regressions of equation (5.1). The dependent variable is the natural logarithm of bilateral trade between provinces i and j. All specifications include fixed-bilateral and -industry effects. Robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10, 5, and 1 percent levels.

Table 5.5: Additional regressors and instrumental variables

	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(trans_k dist_{ij})$	-0.409***	-0.409***	2.034**	1.259	-0.0801	3.950**
	(0.119)	(0.119)	(0.993)	(1.020)	(1.024)	(1.759)
$prod_{ijk}$	8.646***	8.647***	8.657***	9.591***	7.858***	8.650***
	(0.967)	(0.965)	(0.895)	(0.896)	(0.865)	(0.893)
$telecom_k lang_{ij}$	20.12**	20.22**	38.73***	23.91**		
	(8.84)	(10.96)	(11.24)	(11.41)		
$telecom_k GDPpc_{ij}$	234.6					
	(3065)					
$telecom_k religion_{ij}$		-0.0379				
		(20.12)				
$post_k lang_{ij}$					13.38	
					(12.62)	
$promo_k lang_{ij}$						10.71***
						(3.675)
Observations	3330	3330	3327	2554	3258	3327
F-stat	59.18	59.14	62.76	80.94	63.18	62.92
R^2	0.412	0.412	0.411	0.492	0.410	0.410
Cragg-Donald test (p-value)			0.000	0.000	0.000	0.000
1^{st}-stage F stat			184.6	155.3	103.4	119.6
1^{st}-stage partial R^2			0.435	0.423	0.434	0.441

Note: The estimates are from fixed-effects and fixed-effects instrumental variables regressions of equation (5.1). The dependent variable is the bilateral trade between provinces i and j. Legal language status is used as an instrument for $lang_{ij}$. All specifications include fixed-bilateral and -industry effects. In column (4), the sample is restricted to Canada's ten provinces. The Craag and Donald (1993) statistic tests the null hypothesis that the model is weakly identified. Robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10, 5, and 1 percent levels.

Table 5.6: Robustness and sensitivity analysis

	Full sample	Only Provinces	1998 sample
Telecommunication services			
$lang_{ij}$	18.17*	19.30**	24.44***
	(9.313)	(9.161)	(9.325)
	3409	2554	3072
	[154]	[90]	[131]
$work_{ij}$	18.33**	17.03**	
	(7.519)	(7.333)	
	3409	2554	
	[154]	[90]	
Postal services			
$lang_{ij}$	0.833	3.039	6.658
	(8.204)	(8.795)	(7.057)
	3340	2492	3021
	[154]	[90]	[131]
$work_{ij}$	2.509	4.205	
	(6.697)	(7.087)	
	3340	2492	
	[154]	[90]	
Promotional Services			
$lang_{ij}$	2.186	1.155	5.854*
	(2.905)	(3.143)	(3.008)
	3409	2554	3072
	[154]	[90]	[131]
$work_{ij}$	3.105	1.545	
	(2.396)	(2.559)	
	3409	2554	
	[154]	[90]	

Note: The regressions are estimates of equation (5.1). The dependent variable is the natural logarithm of bilateral trade between provinces i and j. All specifications include fixed-bilateral and -industry effects. Each entry of the table reports the estimated coefficients for β_3 with robust standard errors reported in parentheses. Below this the number of observations in the regression is reported. The number of bilateral pairings is given in brackets. *, **, and *** indicate significance at the 10, 5, and 1 percent levels. The full sample includes also the following sectors: Fuels, Petroleum and Coal.

Table 5.7: Robustness to zeros: poisson estimates

	(1)	(2)	(3)	(4)
$\ln(trans_k dist_{ij})$	-0.974***	-0.972***	-0.979***	-0.976***
	(0.254)	(0.254)	(0.254)	(0.254)
$prod_{ijk}$	7.029***	6.916***	7.001***	6.920***
	(1.496)	(1.668)	(1.494)	(1.693)
$telecom_k lang_{ij}$	19.30***	23.36***		
	(6.062)	(5.930)		
$post_k lang_{ij}$			11.82	
			(7.356)	
$promo_k lang_{ij}$				5.690**
				(2.001)
Observations	5610	5466	5454	5466
Pseudo-R^2	0.891	0.908	0.892	0.907

Note: The estimates are from poisson regressions of equation (5.1). The dependent variable is the bilateral trade between provinces i and j. Column (1) uses all available observations. All specifications include fixed-bilateral and -industry effects. In columns (2)-(4), the industries Fuels, Petroleum and Coal have been dropped. Robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10, 5, and 1 percent levels.

Table 5.8: Variable labels

Label	Explanation
$trade_{ijk}$	The trade volume between province or territory i and province or territory j in industry k in million Canadian dollars, including zero trade flows.
$\ln trade_{ijk}$	The natural logarithm of $trade_{ijk}$.
$trans_k$	The share of transport margins in total inputs of industry k.
$dist_{ij}$	The bilateral distance between two capital cities of provinces or territories.
$prod_{ijk}$	The industry differential in the total production of two provinces.
$telecom_k$	The share of telecommunication services in total inputs of industry k.
$post_k$	The share of postal services in total inputs of industry k.
$promo_k$	The share of promotional services in total inputs of industry k.
$lang_{ij}$	The probability that two randomly chosen people from province i and j are able to communicate with each other.
$work_{ij}$	The probability that two randomly chosen people from province i and j use the same official language at work.
$GDPpc_{ij}$	The joint GDP per capita of provinces i and j.
$religion_{ij}$	The probability that two randomly chosen people from province i and j have the same religion or denomination.

Curriculum Vitae

10/2006-01/2010	Ph.D. program in Economics Munich Graduate School of Economics Ludwig-Maximilians-University, Munich, Germany
09/2005 - 09/2006	Studies in International Economics, Master of Science Erasmus University Rotterdam, The Netherlands
10/2002 - 09/2005	Studies in Economics, Social Sciences and History, Bachelor of Arts University of Erfurt, Germany
09/2001-09/2002	Military Service Heeresmusikkorps I, Hannover, Germany
06/2001	Abitur Gymnasium am Fredenberg, Salzgitter, Germany
22/11/1981	born in Salzgitter, Germany