Essays on the Distribution of Income, Attention, and Rewards

DISSERTATION
of the University of St. Gallen,
School of Management, Economics,
Law, Social Sciences and International Affairs
to obtain the title of

Doctor of Philosophy in Economics and Finance

submitted by

Stefan Legge

from Germany

Approved on the application of

Prof. Dr. Reto Föllmi

and

Prof. Dr. Winfried Koeniger

Prof. Dr. Josef Zweimüller

Dissertation no. 4537

Difo-Druck GmbH, Bamberg 2016

The University of St. Gallen, School of Management, Economics, Law, Social Sciences and International Affairs hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St. Gallen, May 9, 2016

The President:

Prof. Dr. Thomas Bieger

Acknowledgements

In general, the process of writing a dissertation is a rewarding and instructive endeavor. At times, however, it can and will be a challenging and dispiriting task. A large number of colleagues and friends have supported me throughout this time and it is my greatest pleasure to express my gratitude to them.

First of all, I wish to thank my supervisor Reto Föllmi who provided me with an excellent, intellectually challenging, and cordial environment. In all my academic endeavors, he supported me with outstanding guidance while giving me the freedom to explore my own ideas. Furthermore, he inspired my work and taught me a lot about economics in general and economic research in particular. As a coauthor of two chapters, he also contributed directly to my dissertation. After several years as his research and teaching assistant I will always be grateful for his support and I look forward to many more fruitful discussions and intriguing research projects.

It would be hard to exaggerate how much I owe to my colleagues at the SIAW Institute in St.Gallen. Over the period of many years, this place provided me with a cordial and helpful environment to conduct my work. I am especially awed and indebted to Lukas Schmid who I first met as a fellow graduate student. Over the years, he turned from a nice office colleague into a formidable coauthor and great friend of mine. There is no doubt that this dissertation would not have been possible without his enduring help. Thank you, Lukas. For their help and many pleasant discussions, I would also like to thank Christine Benesch, Anna Brüderle, Irene Fensore, Johannes Fritz, Daniel Garraty, Berit Gerritzen, Florian Habermacher, Friedhelm Hentschel, Adrian Jäggi, Isabel Martinez, and Fabian Schnell. A special thanks to Alexa Tiemann who helped me a lot in my early days at the SIAW and is coauthor of my first paper. Furthermore, Michèle Klarer as well as Mark Schelker made my work at the institute much more enjoyable. Last but most certainly not least, I express my deep gratitude to Gabriela Schmid who has always been an outstanding help, especially during tough times.

During my PhD studies, I had the opportunity to visit the economics departments of both the University of California at Berkeley and Princeton University. In this context, I would like to acknowledge the Swiss National Science Foundation for the generous scholarship (P1SGP1-155369). For his invitation to Berkeley, I wish to express my gratitude to Emmanuel Saez. Multiple discussions with him refined and significantly added to my research projects. In addition, being able to work at Berkeley meant that I could meet many outstanding researchers. Most importantly, I remember the numerous discussions with Josef Zweimüller from the University of Zurich. His support and guidance greatly benefited my dissertation and I am very thankful to him for being a member of my thesis committee. Moreover, the discussions at Berkeley with Stefano DellaVigna, Devin Pope, Morten Saethre, and Justin Sydnor greatly contributed to my dissertation. Most of all, my time at Berkeley was enriched by sharing an office with Ole-Petter Moe Hansen. The discussions with him yielded not only a great amount of fun but also joint research projects. I look forward to continuing our work and his witty comments.

At Princeton, Alexandre Mas invited me to work at the Industrial Relations Section which is a phenomenal place to conduct research. During my visit, Alex provided me with excellent guidance for several of my research papers. Furthermore, his colleagues Henry Farber and Christopher Neilson were always available for very fruitful discussions. My thanks also go to Linda Belfield, Felipe Goncalves, Daniel Herbst, Dejan Kovac, Adrien Matray, Jakob Schlockermann, and all others who made my time at Princeton so memorable and enjoyable.

At the University of St.Gallen, I benefited greatly from many colleagues whose support is gratefully acknowledged. In particular, I want to thank Michael Lechner who has supported my work in St.Gallen from day one and also served as chairman of my thesis committee. Furthermore, my thanks go to Guido Cozzi, Philipp Denter, Simon Evenett, Roland Hodler, Martin

Huber, Christian Keuschnigg, Gebhard Kirchgässner, Winfried Koeniger, Giovanni Mellace, Manuel Oechslin, Anthony Strittmatter, and Uwe Sunde. I also look back and remember with great fondness numerous pleasant meetings with colleagues and friends for which I thank particularly Melanie Annaheim, Margaret Davenport, Nicolas Eschenbaum, Stephan Fretz, Hans Fricke, Alexander Gruber, Benjamin Guin, Katharina Hofer, Michael Knaus, Michael Kogler, my good friend David Pankoke, as well as Daniel Steinberg. Finally, I want to thank all those who helped me organize my studies. This comprises in particular Franziska Fawcett, Martina Flockerzi, Mirela Keuschnigg, and Kristin Schmidt.

Outside of the academic realm, my deep thanks go to numerous friends and relatives who have been a permanent source of sunshine in my life. In particular, I want to thank my siblings Viktoria and Michael, my relatives Resi and Heinrich Legge, Beata Sting, as well as many friends including Philipp Alles, Felix Bartels, Martin Breidel, Felix Heide, Vanessa and Dominik Jost, Lars Kettner, Steffen Löhr, as well as Manuel Ruby.

Despite my best efforts, this list of people I would like to thank will likely remain incomplete. But I could not leave out the greatest source of support for this dissertation. For more than a quarter of a century, I am lucky to have the unconditional support of my great parents, Angelika and Paul Legge. My academic journey has come at a significant cost to them but they never hesitated a moment to help me. There are no words to describe how much I owe them. Herzlichen Dank.

St.Gallen, July 19, 2016 Stefan Legge

Abstract

This dissertation uses empirical methods to study problems in the fields of macroeconomics, trade, political economy, and behavioral economics.

Chapter 1 provides an extensive literature review that dissects the large body of research on the determinants of income inequality. It shows that most explanations focus on technology, education, trade, and labor market regulations while less attention has been paid to the role of immigration, superstars, female labor supply, marital sorting, and demographic trends.

Chapter 2 investigates how population aging affects the rate of innovation. Assuming individuals must spend time on learning how to use new technology, the demand for innovative goods declines if the population becomes older. Data from OECD countries support these theoretical predictions.

Chapter 3 analyzes both empirically and theoretically how firms in developing countries respond to tariff reductions. In the presence of imperfect capital markets trade liberalization can hurt small and medium-sized firms. Financially constrained firms are more likely to either leave the market or reduce their R&D efforts when being exposed to lower tariff protection.

Chapter 4 studies voter preferences over immigration and redistribution in Europe. It shows that after 2002 support for income redistribution increased while attitudes over immigration polarized. A model that allows individuals to be motivated by altruistic motives is used to explain these trends.

Chapter 5 investigates the importance of ranking schemes and tests whether media attention affects perceptions about future events. Data on close World Cup tournaments in alpine skiing allows to estimate the causal effect of media attention. The results document that rankings generate sharp discontinuities in media attention. However, there is no evidence that biased media attention affects prices or quantities in the betting market.

Chapter 6 examines whether limited attention can affect individual risk-taking behavior and whether high stakes and individual experience mitigate behavioral biases. A unique feature of World Cup tournaments in slalom races allows to estimate causal effects. The results indicate that athletes misinterpret actual differences in race times by focusing on the leftmost digit, resulting in increased risk-taking behavior.

Zusammenfassung

Die vorliegende Dissertation verwendet empirische Methoden zur Untersuchung von Problemen in den Bereichen Makroökonomie, internationaler Handel, Politische Ökonomie und Verhaltensökonomie.

Kapitel 1 besteht aus einer Zusammenfassung der Literatur zur Frage, welche Faktoren die Einkommensverteilung beeinflussen. Es zeigt sich, dass die meisten Erklärungen auf technologischen Veränderungen, Bildung, Handel und Arbeitsmarktinstitutionen basieren. Weniger einflussreich sind Migration, Superstars oder Demographie.

Kapitel 2 untersucht, wie Bevölkerungsalterung auf Innovation wirkt. Wenn Individuen Zeit investieren müssen, um neue Technologien zu verwenden, sinkt die Nachfrage nach innovativen Gütern in einer alternden Bevölkerung. Daten der OECD Länder unterstützen diese theoretische Vorhersage.

Kapitel 3 analysiert, wie Firmen in Entwicklungsländern auf Handelsliberalisierung reagieren. Bei imperfekten Kapitalmärkten kann eine solche Liberalisierung kleinen und mittelgrossen Firmen schaden. Kreditbeschränkte Firmen neigen dazu, den Markt zu verlassen oder Forschungsausgaben zu reduzieren, wenn sie von Zollsenkungen betroffen sind.

Kapitel 4 studiert Wählerpräferenzen bezüglich Immigration und Umverteilung in Europa. Es zeigt, dass nach 2002 die Unterstützung für Einkommensumverteilung gestiegen ist während die Meinungen zur Immigration polarisierten. Ein theoretisches Modell, in dem Individuen potentiell altruistisch sind, wird verwendet, um diese Trends zu erklären.

Kapitel 5 untersucht die Bedeutung von Ranglisten und ob Medienaufmerksamkeit individuelle Erwartungen beeinflusst. Daten von Weltcup Skirennen ermöglichen die Schätzung von kausalen Effekten. Die Resultate zeigen, dass Ranglisten klare Unterschiede in der Medienaufmerksamkeit erzeugen, jedoch keinen Einfluss haben auf Quoten oder Nachfrage in Wettbörsen.

Kapitel 6 analysiert, ob beschränkte Aufmerksamkeit einen Einfluss hat auf individuelles Risikoverhalten und ob hohe Einsätze und Erfahrung Verhaltensfehler reduzieren. Im Rahmen von Weltcup Skirennen können kausale Effekte geschätzt werden, die zeigen, dass Athleten tatsächliche Zeiten falsch interpretieren und daher ein erhöhtes Risiko wählen.

Contents

Contents **vii**

1 What Drives Income Inequality? **1**
1.1 Introduction . . . 1
1.2 Empirical Evidence . . . 3
1.3 Theoretical Framework . . . 10
1.4 Theories on Inequality . . . 14
1.5 Interplay of Factors . . . 40
1.6 Conclusion . . . 42

2 Innovation in an Aging Population **49**
2.1 Introduction . . . 49
2.2 Demographic Trends . . . 52
2.3 Theory . . . 56
2.4 Empirical Evidence . . . 62
2.5 Conclusion . . . 70

3 Innovation and Trade in the Presence of Credit Constraints **86**
3.1 Introduction . . . 86
3.2 Theory . . . 92
3.3 Data . . . 100
3.4 Empirical Results . . . 104
3.5 Conclusion . . . 114

4 Trading off Welfare and Immigration in Europe **129**
4.1 Introduction . . . 129
4.2 Data . . . 132
4.3 Explaining Trends in Policy Preferences . . . 134
4.4 Theoretical Model . . . 145

4.5 Alternative Explanations . . . 157
4.6 Conclusion . . . 162

5 Media Attention and Betting Markets **200**
5.1 Introduction . . . 200
5.2 Data . . . 204
5.3 The Identification Problem . . . 210
5.4 Results . . . 217
5.5 Conclusion . . . 229

6 Limited Attention and Risk-Taking Behavior **262**
6.1 Introduction . . . 262
6.2 Theoretical Considerations . . . 266
6.3 Data . . . 274
6.4 Econometric Approach . . . 275
6.5 Results . . . 279
6.6 Conclusion . . . 293

References **315**

Chapter 1

What Drives Income Inequality?

1.1 Introduction

Income inequality is an old phenomenon and has existed throughout the history of mankind. In a recent study, Dow and Reed [2013] show that archaeological inequality could arise from the exclusion of outsiders by insider groups who controlled access to locations of unusually high quality. Today, income inequality is arguably the one of the most influential topics in both European and American politics. At Knox College in July of 2013, U.S. President Obama announced that he will focus his energy for the rest of his presidency on income inequality.[1] Similarly, the discussion about rising disparities in wages have had huge influence on various elections in Europe. Even the Pope addressed the issue, intending to push the church in a new direction.[2] In recent decades rising income inequality has also received a great deal of attention by economists. Discussions about the causes and consequences have been documented by Katz and Autor

[1] "Obama: Rest of my presidency is for working-class America", *CNN*, July 25, 2013.

[2] "Pope Francis denounces trickle-down economic theories in critique of inequality", *The Washington Post*, November 26, 2013.

[1999], Feng, Burkhauser and Butler [2006], Dustmann, Ludsteck and Schoenberg [2009], Kopczuk, Saez and Song [2010], or Atkinson, Piketty and Saez [2011]. The increase is at odds with simple theoretical models such as the Kuznets [1955] curve which would predict decreasing inequality for mature economies. However, as Williamson [1998] points out, there are several variables that can offset the forces generating the Kuznets curve. Demographic trends are one of these forces. Other factors are globalization, immigration, skill-biased technological change, a slowdown in educational supply, or product demand changes. When observing that inequality —i.e., the variance of incomes— is higher in 2010 than in previous decades, the question arises why this is the case. Answering this question, however, is complicated due to the multitude of influential factors as well as their interaction effects. This makes it difficult to disentangle the explanatory power of each separate theory. And since many factors are at work at the same time, their effects might offset or amplify each other. There is now a long literature aiming at identifying how much of the rise in inequality can be attributed to the various factors.

In this paper, I dissect the leading theories on why inequality has increased over the past couple of decades. This serves as the indispensable basis for any study on wage inequality. My work intends to update and expand previous work by Topel [1997], Katz and Autor [1999], and Acemoglu [2002*b*]. Moreover, I discuss the potential interplay of the various factors that shape the wage distribution. For one thing this may shed some light on new, promising explanations for changes in the income distribution. In addition, neglecting interaction effects may cause erroneous conclusions regarding the impact of individual factors.

The paper proceeds as follows. Section 1.2 provides some empirical evidence on income inequality over the past few decades. Next, Section 1.3 illustrates a theoretical framework guiding the overview of theories on inequality. Starting with technological change, Section 1.4 presents the various explanations for rising inequality including trade, labor market regulations, immigration, demography, female labor supply, and marital sorting. Section 1.5 examines the interplay of factors and Section 1.6 concludes.

1.2 Empirical Evidence

Before turning to the different theories that aim at explaining the evolution of wage inequality, this section addresses two important aspects. First, how income inequality can and should be measured. This includes a brief discussion of common fallacies with respect to income statistics. Second, I present some empirical evidence on the evolution of inequality and in particular the increase in (residual) wage inequality since the 1970s. Based on previous work by Autor [2009] and others, I indicate stylized facts which have to be explained by theory.

1.2.1 Measures of Income Inequality

There is a multitude of measures to express the magnitude of income inequality. All of these measures have their pros and cons and should therefore be regarded as complements rather than substitutes.[1] Detailed discussions of how income inequality should be measured can be found in seminal work by Atkinson [1970], Paglin [1975], as well as Formby, Seaks and Smith [1989].

Inequality is often expressed by the share of total income that is earned by a specific share of the population. Shares of income are also compared using the Kuznets ratio which is defined as the income share of the y % richest divided by the income share of the x % poorest people. It is common to use a ratio based on $y = 20$ and $x = 40$. Another, more frequently used measure of income inequality is the Gini Index. This index is based on the Lorenz curve which shows the proportion of the distribution assumed by the bottom x percent of the population. The Gini coefficient is then given by the area between the line of perfect equality (45 degree line) and the observed Lorenz curve, as a percentage of the area between the line of perfect equality and the line of perfect inequality. Hence the range of the Gini coefficient is from zero to one with the first meaning perfect equality and the latter maximum inequality. The Gini index, however,

[1] In the economics literature, starting with seminal contributions by Dalton [1920] four criteria have been postulated that any measure of inequality ought to satisfy: Anonymity (it does not matter who earns the income), Proportionality (the size of the population does not matter), Relativity (only relative incomes matter), and the Dalton priciple (regressive transfers increase inequality).

does not fully describe the curve. Thus very different distributions of income can yield the same Gini coefficient. Moreover, the Gini coefficient is not suited to compare countries of different population size [Taleb, 2015].

Another way to express income inequality is to focus only on the top income share. This refers to the share of total income that is earned, for example, by the richest one percent of the population. While this metric is commonly used both in academic and newspaper articles, it suffers from the fact that in different years very different individuals are in the specified top income bracket. Especially when comparing top income shares over time one has to bear in mind that individuals constantly move between brackets. Hirschl and Rank [2015], for example, find that 53.1% (11.1%) of the American population will spend one year of their life in the top tenth (one) percentile of the income distribution. Moreover, few people spend multiple years in the top one percent bracket. These concerns should also be kept in mind when using indicators such as the 90/10 wage gap. This ratio compares the wage of a worker at the 90th percentile of the income distribution with the one of a worker at the 10th percentile. Under full equality, this ratio is one. But with rising income inequality, the ratio exceeds one. Related to this measure is the so-called college premium which refers to the gap between (log) earnings of college and high school graduates. A recent report by Taylor, Fry and Oates [2014] documents the increasing disparity in all measures of economic well-being (e.g., income, unemployment rate, share in poverty) between college graduates and those with a high school diploma.

Finally, it is worth noting that many economic studies on inequality often refer to residual wage inequality. This refers to the inequality that is left after netting out (estimated) effects of education, potential experience, and gender. These factors explain a large fraction of variation in wages [Juhn, Murphy and Pierce, 1993]. Moreover, they are relatively easy to measure. Accounting for them, however, is debated in the literature [Autor, Katz and Kearney, 2008; Lemieux, 2006].

Pitfalls with Income Statistics — Like any statistic, all measures of income inequality should be used carefully. There are many pitfalls and fallacies with popular inequality statistics. A first order question in the quantification of inequality is whether annual income or hourly wages are actually the right measure. If individuals follow a typical path of life-cycle income, some retirees are not at all poor despite their current low annual earnings. Moreover, changes in the age structure will lead to mechanical changes in measured inequality.[1] Using detailed income statistics from Norway, Aaberge, Atkinson and Modalsli [2013] show how income inequality within a given cohort evolves over the life cycle.

To capture some of the distortions, Meyer and Sullivan [2011, 2013] suggest to differentiate between income and consumption. This way, non-monetary income (e.g., self-rented homes, household production, income from the informal sector, use of public goods, etc.) can at least to some extent be taken into account.[2] Another concern with respect to income inequality refers to measurement problems. Dealing with household (or family) income statistics requires additional caution since households vary in size and composition over time and across groups of the population. Moreover, aspects such as part-time work, capital income, fringe benefits, or transfers complicate the measurement of 'true' inequality.[3] Also, the adjustment for inflation is subject to controversial discussions.[4] If rich and poor individuals consume goods with different rates of inflation, this has to be taken into account in time series data. Moreover, one must not forget that income groups in statistical terms do not necessarily refer to flesh-and-blood people. Instead, over time individuals frequently move from one income bracket to another as explained above.

Finally, the definition of income is also subject to discussion. On the one hand, Hamermesh [1999] and Pierce [2001] find that accounting for non-wage

[1]In fact, a large degree of inequality simply results from the fact that income is a function of age (or experience). This has been illustrated by Sowell [2011] and examined by Lemieux [2006]. In Section 1.4.5, I discuss this issue in detail.

[2]Recent work by Aguiar and Bils [2015] finds that consumption inequality has followed a path that is very similar to the trend in income inequality.

[3]Katz and Autor [1999] document that supplements to wages and salaries increased steeply between 1959 and 1994. In addition, their importance varies substantially among jobs and individuals.

[4]A paper by Broda and Romalis [2009], for example, finds that standard price indexes rely on a representative consumer assumption which is only valid in a world with an identical basket of goods consumed by different income groups. The paper, however, suffered from controversies about its data.

benefits (e.g., safety provisions, health care, vacation, etc.) does not alter the finding that low-skill wages have declined in the past decades. On the other hand, Burkhauser, Larrimore and Simon [2012] describe the problem of using data on tax units rather than households. In addition, they argue that income data often do not include government transfer payments, are pre-tax rather than post-tax, do not properly adjust for changes in household size, or do not include non-taxable compensation such as employer-provided health insurance. Using the broadest measure of available resources, the authors find that the median income of individual Americans grew by 36.7 percent over the period from 1979 to 2007. This is in sharp contrast to the pre-tax and pre-transfer median real income which rose by only 3.2 percent.

Finally, when comparing income inequality over time, the problem arises how to adjust wages for inflation. The use of a simple consumer price index (CPI) for all income groups can be misleading. Moretti [2013] finds that the wage gap between college and high school graduates has grown significantly less when applying city-specific CPI.

All these technical issues make it very complicated to examine the evolution of income inequality. Even more difficult is the task to derive *normative* statements, especially when attributing a portion of the observed variation in incomes to preferences rather than ability [Lockwood and Weinzierl, 2015]. Whether income inequality is generally considered a problem has been subject to intense discussions [Bivens and Mishel, 2013; Mankiw, 2013].

1.2.2 Evolution of Income Inequality

The evolution of income inequality can be illustrated in many different ways. In Figure 1.1, the Gini indices and top income shares of six selected countries in the past decades are shown.

— Figure 1.1 about here —

Both measures show a significant increase in income inequality after 1980. Note that this is true even when taking into account taxes and transfers. Especially for the United States, a strong positive trend is evident. Kopczuk, Saez and

Song [2010] report that inequality started rising in the 1970s, plateaued from 1977 to 1980, then rose steeply until 1988, and grew at a slower pace through the end of their series in 2004. Figure 1.1 shows that in the U.S., most of the increase in the Gini coefficient took place before 2000. However, the top-1% income share kept rising. While the other countries had somewhat different trends it is notable that even in Sweden with its large public sector and welfare state, there is a significant increase observable.

The timing of the surge in inequality, beginning in 1980, coincides with inflection points in several other variables. This is what makes the identification of what caused the increase difficult. First of all, technological progress has been enormous in the past three decades. In 1980, for example, there was less than one personal computer per 100 people in the United States. This number increased to ten in 1985, twenty in 1990, thirty in 1995, sixty in 2000, eighty in 2005 and about 115 today. This technological change has altered the demand for skills in the labor market. Acemoglu and Autor [2011] document that the college/high school wage ratio declined during the 1970s but has been increasing since the early 1980s. This college premium can be regarded as the market price of skills, especially when netting out the effects of experience and gender. Adjusting for changes in the composition of the labor force, Acemoglu and Autor argue that the rising college premium resulted from the decreasing relative supply of male college versus non-college educated workers in the early 1980s. In contrast, Card and Lemieux [2001] attribute the trend in the college wage premium to male college completion rates, the Vietnam War, the baby boom cohort, and the incentive effects of a declining college premium during the 1970s.

Extending the comparison of college versus high school wages, Acemoglu and Autor [2011] document the diverging trends in real wages across educational groups. From the mid 1970s, wages increased considerably for highly educated (male) workers but decreased for workers without college diplomas.[1] Following a similar pattern, weekly and hourly wages of the 10th, 50th and 90th percentile increased between 1963 and 1973. Afterwards, however, there has been a substan-

[1] The pattern for female workers is notably different with real wages increasing for all educational groups except high school dropouts.

tial increase in the 90th percentile wage but very modest increases for the 50th and 10th percentile. Acemoglu and Autor [2011] show that for males at the 10th percentile, wages actually decreased after 1970.

When estimating changes in hourly wages by earnings quantiles (reflecting skills groups), Autor, Katz and Kearney [2008] document a U-shaped pattern for the time period after 1990. In a subsequent study, Autor, Dorn and Hanson [2013] find that this *polarization* of the labor market led to rising wages at the top and bottom of the skill distribution with very low gains in the middle. A similar pattern can be found for changes in employment shares [Acemoglu and Autor, 2011; Autor, Dorn and Hanson, 2013].

Although these findings are based on U.S. data, Goos, Manning and Salomons [2014] document similar trends for most European countries.[1] However, the evolution of inequality in general has been different in the United States and (continental) European countries. Okazawa [2013] examines these different developments and argues that different tax and education systems can explain why the wage distributions have evolved differently. In general, it remains a challenge to identify the contribution of single factors to the rise in income inequality because trends have been largely similar in Western countries. Hence many studies often focus on unique policy changes to address the identification problem. But given the multitude of stylized facts that have to be explained, it is unlikely that a theory based on one factor can account for every single trend.

A summary of stylized facts about wage inequality in the United States is provided by Autor [2009]. These facts focus on wage trends for specific groups of the population and apply to residual wage inequality as well. The most recent work in the field of income inequality has largely focused on top income shares. Studies by Autor, Katz and Kearney [2008], Lemieux [2008], Atkinson, Piketty and Saez [2011], as well as Föllmi and Martínez [2016] document trends for several countries. While the trends in the United States are more pronounced, an increase in the share of income collected by the top percentile (in a given year) has been found in several Western countries.

[1]For Germany, Dustmann, Ludsteck and Schoenberg [2009] provide detailed evidence of rising income inequality.

1.2.3 The Impact of Income Inequality

The interest in income inequality arises partly because the distribution of income affects a range of economic outcomes. In this subsection, I briefly address this topic.

Economic Growth — Income inequality can affect the growth rate of output through several channels in the short and long run [Halter, Oechslin and Zweimüller, 2014]. The most obvious channel through which income inequality can influence economic growth is through differential savings rates. If rich and poor individuals differ in their propensity to save, the income distribution affects aggregate capital and thus output [Kaldor, 1955]. The seminal work by Galor and Zeira [1993] highlights another crucial mechanism for how inequality can adversely economic growth. Poor individuals may face credit constraints which limit their ability to invest in human capital. Hence imperfect capital markets and indivisibilities in human capital investments cause the initial wealth distribution to affect both aggregate output and the long-run distribution of wealth. Additionally, income inequality has implications for product and process innovations, both of which affect economic growth [Föllmi and Zweimüller, 2006; Föllmi, Würgler and Zweimüller, 2014]. Finally, greater wage inequality can generate more redistribution and thus affect the incentive structure via the political process [Alesina and Rodrik, 1994; Persson and Tabellini, 1994].

While there are several further channels through which the distribution of income can affect growth, a quantitative assessment has proven to be complicated. Due to the fact that growth and income inequality are jointly determined, identification remains challenging.[1]

Income Inequality and Social Mobility — For many economists, social mobility is considered to be as important as income inequality.[2] There is empirical

[1] A similar problem has plagued attempts to assess the relationship between trade and growth [Grossman and Helpman, 2014].

[2] Social mobility, however, does not necessarily reflect opportunity: "Only by implicitly (and arbitrarily) assuming that a failure to rise must be due to society's barriers can we say that American society no longer has opportunity for upward social mobility." (Thomas Sowell,

evidence that income inequality and social mobility are related to each other. Clark [2014] documents that countries with a higher Gini coefficient have a larger intergenerational correlation in both earnings and education. This relationship has been called "The Great Gatsby Curve" [Corak, 2013].[1] Using rare surnames to track families and measure the intergenerational elasticity of wealth, Clark and Cummins [2015] find that wealth is very persistent over time. Wealth of families in Wales and England today is still correlated with their ancestors five generations before. Long and Ferrie [2013] document that social mobility in the United States used to be higher than in Britain but the difference disappeared in the 1950s. However, there has not been any significant change in social mobility since the 1970s [Chetty et al., 2014].

1.3 Theoretical Framework

Based on prior work by Topel [1997], I use a simple supply and demand framework to compare and organize the discussion of how each factor influenced the distribution of income. It is important to think of wages as prices for specific types of labor. The most common separation is between skilled and unskilled labor. However, recent work on the so-called polarization of the labor market suggests that there are at least three types: manual, routine, and abstract. I will describe this later in the text but for now it suffices to know that different types of labor exist and their prices are determined by supply and demand.

The Demand Side — The most important factor affecting the demand for specific types of labor is arguably technology. In the production process, broadly speaking, labor and capital are complementary inputs. Depending on what kind of machines (i.e., technology) are used, different types of labor are required. Technological changes like the introduction and refinement of computers and robots —which function as substitutes to low-skill labor— have effectively lowered the

"Economic Mobility", March 6, 2013)

[1] The name for the relationship was coined by former Council of Economic Advisers staff economist Judd Cramer. In the figure, the likelihood that an individual will inherit its parents' relative position of income level is plotted against income inequality.

demand for workers with low skills.[1] Similarly, globalization and the emergence of international production networks have altered production in Western countries. Due to low labor costs poorer countries have a comparative advantage in those steps of the production chain which rely on low-skill labor inputs. Labor market institutions like the minimum wage, unions or employment protection laws also influence firms' demand for workers. Finally, product demand can alter the demand for specific types of labor. Leonardi [2015], for example, uses Consumer Expenditure Survey data to show how changing demand for high-skill-intensive and low-skill-intensive services affected the labor market in the United States and United Kingdom.

The Supply Side — The primary source of skills is the educational system. Acemoglu and Autor [2011] document a sharp deceleration in the relative supply of young college graduate males starting in 1975. This trend has been attributed to the Vietnam War, the declining college premium of the 1970s, the baby boom cohort, and the female college completion rate [Card and Lemieux, 2001]. As a result, the supply of skilled workers did not keep pace with the rising demand for skilled labor due to technological changes [Goldin and Katz, 2008]. Irrespective of the educational system, demographic shifts can mechanically alter the distribution of income. In particular the large baby boom cohort (born 1946-1964) has been found to play a significant role [Autor, Katz and Kearney, 2008; Lemieux, 2006]. Changes in the composition of the (domestic) labor force can also be caused by immigration. Several studies have examined how the inflow of low- or high-skilled migrants affects wages and employment of natives.

Model Framework — In order to compare how various factors affect the distribution of income, I suggest a simple framework based on the model that was provided by Acemoglu and Autor [2011]. Suppose the economy is populated by low- and high-skilled workers, indicated by L and H. The production function of

[1] More precisely, the demand for routine labor has declined [Autor and Dorn, 2013].

this economy is given by

$$Y = \left[(A_L L)^{\frac{\sigma-1}{\sigma}} + (A_H H)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{1.1}$$

where A_L and A_H are factor-augmenting technology parameters and $\sigma \in [0, \infty)$ is the elasticity of substitution between the two types of labor. If we assume that labor markets are competitive, each worker receives a wage that is given by the value of the marginal product:

$$w_H = \frac{\partial Y}{\partial H} = A_H^{\frac{\sigma-1}{\sigma}} \left[(A_L)^{\frac{\sigma-1}{\sigma}} (H/L)^{-\frac{\sigma-1}{\sigma}} + A_H^{\frac{\sigma-1}{\sigma}} \right]^{\frac{1}{\sigma-1}} \tag{1.2}$$

$$w_L = \frac{\partial Y}{\partial L} = A_L^{\frac{\sigma-1}{\sigma}} \left[(A_L)^{\frac{\sigma-1}{\sigma}} + A_H^{\frac{\sigma-1}{\sigma}} (H/L)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{1}{\sigma-1}} \tag{1.3}$$

To express income inequality in this economy we can use the skill premium ω, the wage of high-skill workers divided by the wage of low-skill workers. Taking logs, a convenient form for the skill premium is given by

$$\ln \omega = \frac{\sigma - 1}{\sigma} \ln \left(\frac{A_H}{A_L} \right) - \frac{1}{\sigma} \ln \left(\frac{H}{L} \right). \tag{1.4}$$

This equation can be used to illustrate how technology, education, or international trade affect income inequality. Skill-biased technological change increases the skill premium by raising A_H/A_L. This effect, however, can be dampened if the supply of skills (H/L) increases as well. These two forces are described as a race between technology and education [Goldin and Katz, 2008; Tinbergen, 1975]. International trade effectively affects H/L, too. If the United States, for example, intensifies its trade relationship with a low-wage country, implicitly L increases and the wage premium (in the U.S.) becomes larger. Low-skilled immigration to a rich country has a similar impact.

Necessary Extensions — In order for the model to capture the role of other determinants of income inequality, the framework must be extended. To account for labor market institutions, the labor market in the economy must be described in more detail. The framework outlined above assumes full employment and a perfect labor market. As a result, wages are simply given by the marginal product. Studying the impact of unions, however, requires a bargaining process between firms and workers. Moreover, worker productivity must take more than just two values (w_L, w_H) to allow for a meaningful analysis of minimum wages. Ideally, the productivity distribution across wagers also follows a hump-shaped path over the life cycle of a worker. With this extension, demographic trends can influence measured income inequality as well.

In order for the model to capture how a rising share of female labor supply affects the distribution of wages, the model must feature two entirely separate distributions of experience and skills for men and women. The model would also be more realistic and suited for the analysis of changing product demand—which itself can be the result of increasing income inequality—if the economy featured more than one good. In addition, new goods should be invented with property and patent rights generating large incomes for those at the top of the income distribution. Capital as a further input of production is also necessary to account for non-labor income. A government taxing incomes and redistribution resources is a first step to analyze the impact of public policy. The second step is to include individual preferences over consumption and leisure to account for behavioral responses to policy changes.

This long list is not exhaustive but already reflects the near-impossible task to account for several determinants of income inequality at the same time within one model. The literature has thus restricted itself to the analysis of one factor at a time. However, it is important to keep in mind the complexity outlined above if only to be cautious with policy conclusions.

1.4 Theories on Inequality

1.4.1 Technological Change

The dominant explanation for changes in income inequality is based on technological change. Following this theory, changes in the production process have influenced the demand for specific types of labor. While some skills are complementary to new technology, others can be considered substitutes. This is insofar relevant as the cost of computer technology has fallen dramatically. Nordhaus [2007] describes the tremendous increase in computer power that has been achieved in the past one and a half centuries. Relative to the price of labor, computation has become cheaper by a factor of 7.3 x 10^{13}. This caused both households and firms to increasingly use computers for the completion of routine and increasingly complex tasks. Figure 1.2 illustrates the rapid diffusion of new technology. In 1980, the number of personal computers in American households was almost zero. Ever since, however, the number has increased dramatically, reflecting how quickly the new technology spread. Notably, the time trend has been very similar in European countries.

— Figure 1.2 about here —

Given the fact that the rise in income inequality began around 1980 as well, many theories developed in the early 1990s suggested that technology is a major cause of increasing wage disparities [Bound and Johnson, 1992]. Several empirical studies suggested a complementarity between the new technology and high-skill labor [Acemoglu, 1999; Berman, Bound and Machin, 1998; Berman, Bound and Griliches, 1994; Doms, Dunne and Troske, 1997; Katz and Murphy, 1992]. In a larger historical perspective, one has to keep in mind that technological change can be both skill-replacing and skill-biased. There is a broad consensus that it was skill-*replacing* in the 19th century but skill-*biased* in the twentieth century. The idea of technological developments raising the demand for skills goes back to studies by Jerome [1934] and others in the first half of the twentieth century [Acemoglu, 2002*b*]. Up until today, a substantial part of the literature on inequality suggests skill-biased technological change (SBTC) as its primary cause

[Guvenen and Kuruscu, 2012]. The theoretical foundation of this literature has been termed "canonical model" by Acemoglu and Autor [2011]. It is based on the assumption that different skill groups produce different goods or perform distinct and imperfectly substitutable tasks. Factor-augmenting technology is skill-biased in the sense that it generates *demand* shifts and thus affects inequality. At the same time, however, various factors may change the *supply* of skills, thus either mitigating or exacerbating inequality changes. These factors will be examined in detail later in this paper.

1.4.1.1 Skill-Biased Technological Change

According to Card and DiNardo [2002], most economists in the late 1980s regarded skill-biased technological change (SBTC) as the primary cause of the increase in measured inequality.[1] That is, the rise in new technologies such as computers and robots has benefited high-skilled compared to low-skilled workers. Given the available data, this view was in line with observed time trends. In the 1970s, the college-high school wage gap narrowed which Freeman [1975, 1976] describes as a result of an oversupply of educated workers, coining the term "The Overeducated American". By 1985, however, wage inequality had increased again. This was also the time when personal computers were introduced in the workplace which made it a seemingly perfect explanation.

In general, the benefits of (unskilled labor saving) technological change can be heterogeneous across factors of production. This should trigger distributional effects. The term "directed technological change" refers to the attempt to endogenize the origin and (the direction of the) bias of new technologies that are developed and adopted. Acemoglu [1998, 2002*a*,*b*] argues that in the U.S. the relative supply of skills increased in the late 1960s. This led to a decline in the wage premium for college graduates in the 1970s (along with compositional effects shown by Carneiro and Lee [2011]).[2] But it also triggered an endogenous technological change which was skill-biased and favored highly skilled individuals.

[1]See, for example, Bound and Johnson [1992]; Katz and Murphy [1992]; Levy and Murnane [1992]; Juhn, Murphy and Pierce [1993]; Krueger [1993].

[2]The increase in college enrollment also led to a decline in the average quality of college graduates.

Had it not been for the substantial skill bias in technology, the increased supply of educated workers should have depressed the skill premium.[1]

Empirically, the theory of SBTC has found substantial support. Several studies document that within firms, computer investments and R&D expenditures predict subsequent rates of skill upgrading [Autor, Katz and Krueger, 1998; Bernard and Bradford Jensen, 1997; Machin and Van Reenen, 1998; Wolff, 1996]. Following substantial investments in computer technology, organizational practices decentralize decision-making and raise the demand for high-skilled workers [Bresnahan, Brynjolfsson and Hitt, 2002; Katz and Autor, 1999].

1.4.1.2 Revisionist Literature

In the early 2000s, the so-called revisionist literature (Card and DiNardo [2002], Lemieux [2006]) argued that SBTC was a premature explanation. According to this strand of the literature, the 1980s surge in wage inequality was an "episodic" event caused by institutional and compositional forces. Furthermore the modest inequality growth in the 1990s is inconsistent with a key role for SBTC. Using more recent data than previous studies, Card and DiNardo [2002] find that most of the increase in wage inequality between 1979 and 1999 had occurred prior to 1985, that is before computers were broadly used. Moreover, SBTC (i.e., the rise of computer technologies) can neither explain why inequality did not continue to grow in the 1990s nor why gender and racial wag gaps or the gradient in the return to education have changed. This led the authors of the revisionist literature to suspect that the increase in inequality in the early 1980s has to be explained by other factors. In particular, the fall in the real value of the minimum wage and the declining unionization seem to have had an important impact. I will discuss these factors in Section 1.4.4 in more detail.

However, these explanations also fall short to explain many developments such as the closing of the gender gap. In addition, Autor, Katz and Kearney [2008] highlight that the revisionist literature neglects the differential evolution of inequality in the upper and lower tail of the income distribution. Since inequality

[1] In continental Europe, the technological change took place during the 1980s and the 1990s and appeared to be capital-biased [Acemoglu, 2002*a*].

did not increase much in the lower tail after 1985, the minimum wage does not seem to be a proper explanation. The paper by Autor, Katz and Kearney also introduced the idea of a modified version of SBTC in order to capture the observed polarization in employment and wages.[1]

1.4.1.3 Polarization

The papers by Autor, Katz and Kearney [2006, 2008] document three major findings. First, wage inequality in the top half of the distribution has exhibited an unchecked increase for 25 years while inequality in the bottom half has ceased growing. Second, employment grew rapidly at the bottom and at the top of the distribution. And third, a model with computerization of routine tasks can explain this pattern of polarization. Since the authors employed U.S. data only, Goos, Manning and Salomons [2009] used data on 16 European countries to show that job polarization took place in Europe in the 1990s as well. In almost all European countries, high-paid and low-paid employment grew disproportionately (relative to medium-wage routine jobs).

In some sense, the polarization literature is an extension of the canonical model. Acemoglu and Autor [2011] argue that the canonical model neglects several issues such as the polarization in employment and wages, offshoring, or the effects of technology on moderately-skilled workers. To explain recent trends in inequality more successfully, the authors provide a task-based model. They remove the (implicit) assumptions of the canonical model that each skill group is assigned to a fixed task or product, and that technology is simply factor-augmenting. In the task-based model, the assignment of skills to tasks is endogenous and new technology can enable machines to perform tasks that were previously performed by labor.

In a recent contribution to the polarization literature, Autor and Dorn [2013] present a spatial model in which polarization results from the interaction between consumer preferences and the decreasing cost of automating routine tasks. Their model features three types of labor: low-wage manual, medium-wage routine,

[1] The term "polarization" was originally coined in a working paper by Goos and Manning [2003], later published as Goos and Manning [2007].

and high-wage abstract labor. Following Autor, Katz and Krueger [1998], Autor, Levy and Murnane [2003], Lewis [2011], as well as Goos, Manning and Salomons [2014], technology is hypothesized as being a substitute for low-skill routine tasks while complementing creative tasks of high-skilled workers. The adoption of new technology is endogenous as in Beaudry, Doms and Lewis [2010]. As a result of technological progress, the model predicts a considerable decline in wages for routine tasks. This induces a reallocation of low-skill workers to service occupations. Employment and wages of low-skilled workers in the service industry increase. Autor and Dorn use spatial data from the U.S. between 1980 and 2005 to support their theoretical model.[1]

1.4.1.4 Ability-Biased Technological Change

Apart from skill-biased technological change (SBTC), there is also a literature on ability-biased technological change. Empirical evidence provided by Bartel and Sicherman [1999] supports the idea that technological change increases the demand for innate abilities of more-educated workers in the United States during the 1980s. This strengthens the importance to differentiate between education and abilities. In terms of theory, Galor and Moav [2000] provide a model of ability-biased technological change in which they distinguish between ability and skills. Prior theories of SBTC, the authors show, fall short to explain the observed increase in inequality *within* skill groups because they treat them as homogeneous. According to Galor and Moav, technological progress has three effects. First, the "erosion effect" which describes the reduction in the adaptability of existing human capital for a new technological environment. Second, the "productivity effect" which describes the increase in productivity as a result of a superior level of technology. Third, the "composition effect" which describes the decline in the threshold level of ability above which people choose to become skilled. The latter effect explains why technological progress raises the number of skilled workers and

[1] Building on prior work by Autor and Dorn, Boehm [2015] further investigates the impact of routine-biased technical change (RBTC). Using data on the U.S. labor market, the author provides evidence of task prices being polarized in the 1990s and 2000s.

decreases the number of unskilled workers.

Galor and Moav also describe the positive feedback loop between education and technological progress: the so-called ”acceleration hypothesis”. If the economy experiences a large increase in the *rate* of technological progress, we first observe a slowdown of productivity growth because of the erosion effect. But as the rate of technological progress reaches a new steady state, productivity growth is higher than initially because the erosion effect disappears whereas the productivity effect does not. This argument is in line with previous work by Galor and Tsiddon [1997] who argue that the life cycle of technology shapes the evolution of wage inequality. At first, inventions raise the return to skills and thus inequality increases. When technologies become more accessible, however, the return to skills as well as wage inequality declines.

1.4.1.5 Education and Technical Progress

Implicit in the canonical model as well as other technology-related explanations is the role of education in shaping the distribution of income. While many theoretical contributions simply assume a given number and size of skill groups, education cannot be regarded as exogenous. In fact, Topel [1997] and more recently Goldin and Katz [2008] argue that there is a race between education and technology. Goldin and Katz's book follows the idea by Tinbergen [1975] who first suggested that technological development and increased access to schooling are two dominant yet opposing forces shaping income ratios. In this framework, investments in human capital play a major equalizing role. The recent increase in inequality is seen as a consequence of a slowing rate of accumulation of human capital, which has not kept pace with skill-biased technological progress [Acemoglu and Autor, 2011; Card and Lemieux, 2001]. This is in line with the *steady-demand hypothesis*, as described by Acemoglu [2002*b*].[1]

Acemoglu and Autor [2012] summarize and extend the work of Goldin and Katz. They highlight an important limitation, namely that Goldin and Katz ignore the multi-dimensionality of human capital. To address this, Acemoglu and

[1] In contrast, the *acceleration hypothesis* by Galor and Tsiddon [1997], Galor and Moav [2000], and others holds that there has been an increase in the skill bias after 1970.

Autor set up a new theoretical framework based on the allocation of different types of human capital to distinct tasks. Thus they combine the work of Goldin and Katz with their task-based model presented in Acemoglu and Autor [2011].

Yet it remains a puzzle why the supply of education did not keep pace with technical change. To explain this, Abraham [2008] proposes a general equilibrium OLG model. His argument is that only (the small fraction of) young people can still decide about their educational attainment. Moreover, the cost of acquiring higher education is a negative function of ability. Thus only a small part of the population chooses to attend college. As a result, if skill-biased technological change is the driving force behind rising inequality, high wage inequality will prevail in the foreseeable future. This prediction, however, depends on the substitutability of labor of different skill levels, as emphasized by Lindley and Machin [2011]. They show that within-graduate inequality has increased, and that postgraduate and college-only workers are imperfect substitutes. This supports a key assumption of the canonical models discussed earlier.

Unlike in the standard human capital model [Ben-Porath, 1967], the model by Guvenen and Kuruscu [2009, 2012] distinguishes two labor inputs: While individuals are endowed with a fixed amount of "raw labor" (strength, health) they can accumulate "human capital" (skills, knowledge) during their life. The ability to accumulate human capital, however, is heterogeneous across individuals. Calibrating this model with SBTC, Guvenen and Kuruscu can match several moments of the wage distribution after 1970: a general increase in wage inequality, an initial fall in the college premium followed by a sharp increase, rising within-group inequality, stagnating median wage growth, and increasing consumption inequality.

Finally, Mincer [1997] and Lemieux [2006] point out that heterogeneity in school quality may be higher at the university and college than at the high school level. As more and more students attend higher education, unobserved heterogeneity in wages should increase. This argument is used by Lemieux [2006]. He finds that compositional changes with respect to education (and experience) play a key role in the surge of residual wage inequality.

Concluding the section on technological change and income inequality, there is a broad consensus that new technology has shaped the labor market and the distribution of income. While a large body of literature describes in detail the impact of technological change, two questions remain. First, how much can other determinants of income inequality contribute to understand trends of the past? Second, what is to expect from rapid improvements in robotics and artificial intelligence? A recent study by Autor [2015] sheds some light on the latter question, emphasizing that automation not only substitutes for but also complements (some types of) labor. Addressing the first question, is the goal of the following sections.

1.4.2 International Trade

Research on international trade has long been linked to the distribution of income. Since the seminal work by Ricardo [1817], the concept of comparative advantages is key to thinking about the impact of trade on domestic wages an employment. If a rich country intensifies its trade relationship with a poor country, both are expected to specialize in the production of products with lower relative opportunity costs. A low wage country such as China will produce goods which require a high degree of low-skill labor input. As a result, the intensified globalization of the past decades has been linked to the loss of jobs and decreased earnings of low-skill workers in Western countries.

Figure 1.3 shows the development of trade for six selected countries. The importance of trade is measured by the total exports plus imports as a share of GDP. We see that despite differences in levels, all Western countries have experienced an upward trend. This has been the result of both lower political trade barriers and technological advancements.

— Figure 1.3 about here —

In the simple models of classic trade theory, the lowering of trade barriers should lead to factor price equalization. The increased (effective) supply of low-skilled labor in advanced economies tends to lower prices of goods with low skill intensity. With wages of low-skilled jobs declining in advanced economies, industries substitute toward this type of labor. Moreover, changes in domestic supply of skills should not have a large effect because each country is relatively small compared to the world market which determines factor rewards.

For a long time fears of cheap foreign labor have been influential in political debates. Early contributions by economists indicated that domestic low-skilled workers may be hurt by trade liberalization, especially in the short run [Heckscher, 1919; Ohlin, 1933; Stolper and Samuelson, 1941]. Starting in the early 1990s, international trade was again described as a major force behind de-industrialization and rising income inequality in Western countries. This view is supported by Borjas and Ramey [1994] as well as Wood [1995].[1] Katz and Autor [1999] argue that intensified international trade has an adverse effect on low-skilled workers only to the extent that import-competing industries disproportionately employ low-skilled workers while exporting sectors are relatively skill-intensive. Although this pattern can be empirically supported for U.S. trade with poor countries, it cannot in industries that have high imports and exports with developed countries [Borjas, Freeman and Katz, 1997]. Related to this it is important to note the emergence of global production networks in which discrete activities are allocated across countries. If only some parts of the production are outsourced, most of the effect of trade will be seen *within* industries [Feenstra and Hanson, 1996, 2003; Helpman et al., 2015].

In line with their comparative advantage, most industrialized countries import goods that are intensive in low skills.[2] Thus increased trade openness should implicitly raise the supply of low-skill labor. Therefore globalization is expected to cause wages of low-skill workers to decline relative to high-skill workers. However, as Topel [1997] points out, this effect depends crucially on the substitutability of high-skill and low-skill workers. As industrialized countries specialize in skill-

[1] Wood [1995] points out that within each sector there is a wide distribution of factor proportions and labor productivity. Imports from least developed countries are likely to be most directly competing with the segment of an industry using the most unskilled-labor intensive production techniques.

[2] Manufacturing imports of the United States from less-developed countries increased from 0.8% of GNP in 1970 to 2.3% in 1980 to 2.8% in 1990 to 4.1% in 1996 [Borjas, Freeman and Katz, 1997].

intensive goods, the question is to what extent low-skill workers can switch to skill-intensive industries. If substitution is easy, international trade is unlikely to cause rising income inequality. While classic trade theories assume perfect mobility, empirical evidence suggests a moderate magnitude of substitutability [Topel, 1997]. However, when workers are induced to switch occupation because of exposure to international trade, empirically this has been associated with significant wage losses [Ebenstein et al., 2014].

Although a study by Autor, Levy and Murnane [2003] finds overall very limited support for the hypothesis that globalization (and offshoring) played an important role in shaping employment trends, more recent papers do find a substantial effect of trade on income inequality. Harrison, McLaren and McMillan [2011] as well as Haskel et al. [2012] document how studies published between 1990 and 2010 revealed the inconsistency of Heckscher-Ohlin trade models and observed changes in inequality. However, the authors also review the new trade theory triggered by Melitz [2003] which is based on heterogeneous firms. Although still lacking empirical support, the new theories—such as in Grossman and Rossi-Hansberg [2008], Helpman, Itskhoki and Redding [2010], Egger and Etzel [2012], or Blanchard and Willmann [2013]— provide detailed insights into the effects of trade on inequality. A notable exception with respect to empirical evidence is the study by Goos, Manning and Salomons [2014]. They use data on 16 European countries and conclude that in addition to the routinization hypothesis, offshoring plays a significant role in explaining employment trends.[1]

In terms of theory, several recent papers successfully extend standard trade models. Helpman, Itskhoki and Redding [2010] introduce search and matching frictions of the Diamond-Mortensen-Pissarides type in a Melitz [2003] model. They show that trade openness initially increases inequality but subsequent liberalization can have ambiguous effects on inequality. Monte [2011] combines skill-biased technological change and trade in a model of heterogeneous technology and individuals. Using this model, he can explain the asymmetric effects of trade and technological shocks across the wage distribution, in line with data from the

[1] To some extent, offshoring may augment the effects of technological change [Autor and Dorn, 2009; Autor, Dorn and Hanson, 2013]. A discussion of this is provided by Autor, Dorn and Hanson [2015].

past fifty years in the United States. Finally, Egger and Kreickemeier [2012] show how a model of heterogeneous individuals (with respect to managerial ability) and rent sharing at the firm level can explain why aggregate welfare gains after trade liberalization come along with rising inequality between and within population groups.

A key policy question is to what extent ongoing globalization can affect industries and jobs. Blinder [2009] as well as Blinder and Krueger [2013] argue that essentially any job that does not need to be done in person (i.e., face-to-face) can ultimately be outsourced, regardless of whether its primary tasks are abstract, routine, or manual. Acemoglu and Autor [2011] provide a measure of occupational offshorability and show that it is highest in clerical and sales occupations. Moreover, offshorability is considerably higher in professional, managerial and technical occupations than in either production/operative or in service occupations. This reflects the fact that many white-collar job tasks primarily involve generating, processing, or providing information. Hence they can potentially be performed from any location. As a result, intensified international trade will continue to affect employment and wages in developed countries.

1.4.3 Immigration

Large waves of migration typically affect the composition of the immigration country's labor force with respect to education or experience. If the bulk of immigrants supplies low-wage labor, often fears arise that immigration may have a negative effect on native low-skilled workers. This illustrates how immigration can affect the distribution of income. The question, however, is whether this fear is supported by empirical evidence.

While Blau and Kahn [2015] argue that such compositional effects can indeed have a substantial impact on wages and employment, most of the empirical literature does not provide much evidence for a large effect of immigration on wage distributions [Butcher and Card, 1991; Ottaviano and Peri, 2012]. This finding can be rationalized in two ways. First, the group of immigrants usually makes up only a small fraction of the total labor force.[1] Second, immigrants and natives are often imperfect substitutes in the labor market even if they possess equal levels of education and experience. That view is supported by Card [2009] using city-level data from the United States.[2] Cortes [2008], Manacorda, Manning and Wadsworth [2012] as well as Ottaviano and Peri [2012] obtain similar results for the substitutability and emphasize that the only major negative effect of immigration is on the wages of previous immigrants.[3] Often immigrants downgrade their wages upon arrival. As a result, domestic workers in the upper part of the income distribution can gain from low-skill immigration [Beerli and Peri, 2015; Dustmann, Frattini and Preston, 2013].

Apart from international migration, Bound and Holzer [2000] sheds some light on how within-country migration might affect wages and employment. The authors argue that people within the United States can move to other places if their area of residency is affected by a negative labor demand shock. However, empirical research documents a lower geographic mobility among low-skilled workers. This is one reason, the authors argue, why some less educated workers in the U.S. suffered from a decline in real wages.

1.4.4 Labor Market Institutions

While technology, trade or immigration affect either the demand or supply of (skilled) labor, labor market institutions usually play a different role. Before discussing this, it is important to note that I define the term 'labor market institutions' to include labor unions, minimum wages, wage setting practices and social norms. While each item should be analyzed separately, they all tend to

[1] Topel [1997] points out that during the 1970s, immigration added only 2 million workers to the American labor force while during the same decade 20 million native workers entered the labor market. Nevertheless, immigration can be concentrated on some cities, a fact that has been exploited for identifying the impact of immigration [Card, 2009].

[2] Card [2012] discusses in detail why some authors do and others do not find a sizable effect of immigration on wages and employment.

[3] Following the negative-selection hypothesis by Borjas [1987], this might be because similarly skilled people migrate from one country to another. A discussion of this hypothesis is provided by Chiquiar and Hanson [2005].

dampen income inequality [Akerman et al., 2013]. Moreover labor market institutions are particularly relevant for explaining cross-country differences in income inequality. With technological change and international trade being phenomena affecting all countries roughly equally, significant differences in union density and labor market regulations allow for the investigation of what explains different inequality trends across countries.[1] And although a recent study by Goos, Manning and Salomons [2014] uses data on 16 European countries and finds labor market institutions not to be a significant factor, several other studies do find an impact. In this section, I discuss each institution separately.

Before I review the related literature, I provide some descriptive statistics. Figure 1.4 shows time trends for the two prominent labor market institutions. On the one hand, the trade union density has been falling in almost all Western countries. On the other hand, the (real) minimum wage was increased in several places.

— Figure 1.4 about here —

In the United States, union density was already comparably low in 1970. Since then it has declined from about thirty to less than 15 percent. Similar trends are observed in several European countries. The minimum wage in the United States declined steeply in the 1980s and has remained roughly stable since then (infrequent adjustments explain the bumpy path). However, several U.S. cities recently introduced significantly higher minimum wages.[2] In Europe, especially France's governments constantly increased the statutory minimum wage. The United Kingdom introduced a minimum wage in 1998 which quickly increased until the financial crisis in 2008. Note that neither Switzerland nor Germany had a minimum wage in the period shown in the Figure. Thus I plot trends in Japan and Spain. A large body of research has investigated the impact of minimum wages on employment. In this paper, however, I only discuss how it affects income inequality.

[1] A problem with minimum wages, however, is that they are endogenous in the sense that regulators respond to other market factors [Katz and Autor, 1999].

[2] For a discussion, see the article "A Reckless Wager", *The Economist*, July 25, 2015.

Labor Unions — A key objective of labor unions is to increase the labor share and to reduce the spread of incomes.[1] While this is clearly beneficial for some workers it can be costly for others. Acemoglu, Aghion and Violante [2001] focus on this point and argue that in the United States and Great Britain skill-biased technological change (SBTC) caused a de-unionization, that is a decline in the share of private sector workers in unions. The main idea of their model is that with SBTC, for skilled workers the benefits provided by unions no longer outweigh the costs of wage compression. Therefore they leave unions and the decline in unionization leads to a reduction in the overall compression effect on wages that unions typically exert [Card, 1996]. As result, inequality increases. Empirically, this view is in line with previous work by DiNardo, Fortin and Lemieux [1996] who argue that about 40 percent of the increase in the gap between the 90th and 50th percentile can be attributed to de-unionization. Further empirical support is provided by Champagne and Kurmann [2013] who use data from the Current Population Survey (CPS) to show that across-worker volatility of hourly wages has increased over the past 25 years. By means of a dynamic stochastic general equilibrium (DSGE) model, they argue that de-unionization and new performance pay schemes are responsible for a large fraction of the observed increase in volatility.

When investigating the impact of unions it is important to note that unionization rates evolved differently among different educational groups. Katz and Autor [1999] write that the unionization rate fell from 1973 to 1993 by 20.8 percentage points for those with less than 12 years of schooling, 14.8 percentage points for those with exactly 12 years of schooling, and actually increased slightly for college graduates. In a study related to this, Checchi, Visser and van de Werfhorst [2010] find that union density is highest among mid-income workers. The authors argue that rising inequality can itself be a source of union decline. If the dispersion of wages increases (e.g., due to technological change) high-wage workers may want

[1] Card [1996] finds that the effect of unions on wages are largest for low-skill and less-educated workers. This suggests that higher unionization should reduce income inequality. On the other hand, unions can cause an increase in wage differentials if insider-outsider effects matter. However, this depends on whether there are positive spillover effects on wages of non-union members.

to quit the union. This changes the composition of unions significantly.

Minimum Wages — In order to reduce income inequality and to ensure sufficiently high wages at the bottom of the distribution many governments set minimum wages. When examining their impact it is crucial to use the real value of the wage floor. In the United States, for example, the nominal federal minimum wage was fixed at $3.35/hour from 1981 to 1990. Due to inflation the real minimum wage, however, declined substantially throughout this period [Lee, 1999]. Several papers investigated whether this decline contributed to the rise in income inequality. In a seminal study, DiNardo, Fortin and Lemieux [1996] use data from the CPS and find that —next to the decline in union density— the decreasing real value of the minimum wage is an important factor for explaining rising inequality. Especially the 50-10 log wage differential was strongly affected by the fall in the real minimum wage during the 1980s. In a subsequent seminal study, Lee [1999] uses data on state-level variation in the real value of the federal minimum wage in the United States. His analysis finds an even larger effect of statutory wage floors on income inequality when allowing for spill-over effects of minimum wages. In fact, Lee argues that if it had not been for the decline in the minimum wage inequality would have fallen. Further research by Lemieux [2006, 2008] finds that minimum wages had a strong impact on residual wage inequality between 1973 and 2003. Moreover, Lemieux argues that the theory of SBTC falls short to explain why inequality in several European countries did not increase as much as in the United States. As he points out, institutional changes including union density, minimum wages, social norms, or performance pay schemes play an important role for explaining cross-country trends in wage inequality. The evidence from the United States is supported by several studies using data on other countries. Bosch and Manacorda [2010], for example, investigate whether the minimum wage in Mexico affected earnings inequality between the late 1980s and early 2000s. The authors find that almost all of the growth in inequality can be attributed to a steep decline in the real minimum wage.

The impact of statutory wages floors, however, is not unquestioned. Using data from the United States, Autor, Manning and Smith [2016] address the topic in a recent paper. They note that after DiNardo, Fortin and Lemieux [1996] and Lee [1999] surprisingly few studies have examined the role of the minimum wage for income inequality. Using more data than was available previously, Autor, Manning and Smith reassess the impact of the minimum wage. They document problems with the estimation using only observations from the 1980s. Exploiting a significantly larger data set, the authors find only modest inequality-reducing effects of minimum wages in the lower tail of the wage distribution. This is not surprising given that between 1979 and 2012, there is no year in which more than ten percent of aggregate work hours were paid at or below the federal or applicable state minimum wage.

A more fundamental problem with minimum wages is that they do not necessarily benefit poor workers. Often low-paid individuals are their family's second earners. In fact, the Congressional Budget Office (CBO) estimates that only one fifth of the income benefits are received by people beneath the poverty line. In addition, the increase in consumer prices that results from a higher minimum wage is found to be more regressive than a typical sales tax [MaCurdy, 2015]. Finally, a very high minimum wage may reduce observed income inequality simply because several low-wage workers become unemployed.

Wage Setting and Social Norms — A relatively vague argument why income inequality may have increased in the past decades is that social norms prevented a large dispersion prior to the 1980s. These norms, however, are said to have declined, allowing for a larger spread of wages and some exceptionally high salaries for CEOs and superstars. In the economics literature, linking cultural factors to economic outcomes has long been considered problematic [Guiso, Sapienza and Zingales, 2006]. Nevertheless a couple of studies investigate how pay schemes affect the distribution of income.[1] Piketty and Saez [2003], for example, argue that social norms play a significant role with respect to top wage shares. Especially through their impact on tax policy, norms have shaped the distribution of income. In a subsequent paper, Lemieux, MacLeod and Parent [2009] suggest

[1] Note that the reverse effect, from income inequality on social norms has also been investigated [Knack and Keefer, 1997].

that performance-based payment provides a channel through which changes in returns to skill get translated into higher wage inequality. Using data from the Panel Study of Income Dynamics (PSID), the authors find that about a fifth of the increase in wage variation between 1976 and 1998 can be attributed to performance-based pay schemes. Especially for the upper tail of the income distribution, this channel seems to have great influence on inequality. Despite these studies, there is limited research on how social norms affect the distribution of income. Most of the work in this field remains rather vague.[1] One reason for this is the lack of convincing data to measure norms.

Overall, the empirical evidence suggests that labor market institutions in fact influence the distribution of incomes. However, the debate on the importance of labor market institutions has not been settled. Recent policies like the introduction of the minimum wage in Germany or the increase of the minimum wage in the United States provide an opportunity for future research.

1.4.5 Demography and Compositional Changes

During the 20th century, all countries in Europe and the Western offshoots experienced substantial changes with respect to the composition of their population. For the United States, Hobbs and Stoops [2002] document the effect of fertility and mortality trends as well as immigration on the country's age structure. Most importantly the pyramid structure of the age distribution has been replaced by a more rectangular shape. But it is also important to note the large differences not just over time but also across subgroups of the population.There has been, for example, a growing gap of about 4-7 years between the median age of blacks and whites. Moreover, the median age of whites increased between 1900 and 1940, stagnated until 1970, and has increased steeply ever since the 1970s.

In the labor market, a notable shift occurred after 1980. Figure 1.5 shows the ratio of senior to young workers. The former are defined as 45-64 year olds while the latter comprise 15-34 year-old workers.

[1] Examples include Levy and Temin [2007] as well as Lindsey [2009].

— Figure 1.5 about here —

In all six countries shown, the ratio is strongly influenced by the baby boom generation born in the 1950s and early 1960s. This cohort entered the labor market in the 1970s and thus lowered the ratio of senior-to-young workers. However, after reaching a very low level in the 1980s the ratio has surged in virtually all OECD countries.

How has this compositional change affected the income distribution? If wages follow a Mincer type function of education and experience at least mechanically income inequality can increase as a result of population aging. There is ample empirical evidence that the dispersion of wages is larger among older workers. Hence one would expect the demographic trends to cause increasing income (and especially wealth) inequality. According to Lam [1997] there are several further trends causing compositional effects, including changing family structures, increases in female labor force participation or part-time work, as well as changes in marital sorting.

Several studies have examined how these compositional effects affected wage inequality. The first strand of the literature has examined how cohort sizes affect wages. Most studies (e.g., Freeman, 1979; Macunovich, 1998, 1999; or Higgins and Williamson, 2002) find that a large cohort size leads to the compression of wages. Higgins and Williamson argue that large mature working age cohorts are correlated with lower aggregate inequality, while large young adult cohorts are associated with higher inequality. The core idea is that fat cohorts tend to receive low rewards. If these cohorts are in the middle of the age-earnings curve it flattens the overall age-earnings curve. As a result, income inequality is reduced. However, when the large cohort is either young or old, the slope of the curve is heightened and inequality augmented.[1]

This line of argument follows earlier research on the effects of demography on income dispersion. Among the first studies, Welch [1979] finds that age-earnings profiles steepened when the baby boom cohort entered the labor market. In other words, the return to experience increased during the 1970s. Dooley and

[1] However, Topel [1997] adds that more experienced workers are usually specialized in certain tasks or jobs. This reduces intra-cohort competition.

Gottschalk [1984] argue that the baby-boom cohort —as a result of sharply increasing and then declining labor force growth rates— faced a rise in the return to human capital. This temporarily increased the variance of investments in human capital and thus income inequality. Murphy and Welch [1992] describe the impact of the baby boom cohort on the composition of the labor force with respect to education and experience. They emphasize that the assumption of a stable demand—which would allow wage changes to be driven exogenously by supply shifts—only fits the data for the period 1963-1979 but not afterwards. This makes it difficult to disentangle supply shifts (i.e., compositional effects) from demand shifts as a result of technological change, trade, or other factors.

An important contribution to the research on how demographic shifts affect the distribution of incomes was provided by Lemieux [2006]. His work adds to the finding by Juhn, Murphy and Pierce [1993] that residual wage inequality — that is inequality among individuals with the same education and experience— accounts for the bulk of the increase in overall inequality. In Lemieux's view an increase in residual wage inequality can result from changes (i) in skill prices, (ii) in the dispersion of unobserved skills (linked, for example, to school quality), or (iii) in measurement error. While all three play a role, Lemieux argues that the dispersion of unobserved skills has been largely neglected in the literature. His key argument is based on two assumptions. First, he assumes that unobserved skills gain importance with labor market experience.[1] As a result, it is straightforward to expect compositional effects from population dynamics on measured inequality. The second assumption holds that unobserved skills also gain importance with the average level of education. The argument is that the dispersion of school quality is larger at the college or university level than at the high school level. Moreover, varying investments in human capital lead individuals to follow different income paths which increasingly diverge over time.

Based on data from 1973-2003, Lemieux emphasizes the impact of composi-

[1] This builds upon a study by Chay and Lee [2000] who document that the variance of wages in general grows with labor market experience. This is in line with Mincer [1974] who also argues that wage dispersion is positively related to experience. He regards heterogeneous investments in on-the-job training as primary cause of this relationship. Farber and Gibbons [1996] provide a different explanation by assuming that wages reflect expected productivity. By accumulating experience, employers learn more about the effective productivity of workers.

tional effects on inequality. He finds that wage dispersion is larger for older and more educated workers even within narrowly defined groups. Hence a large fraction of the increase in residual income inequality is a spurious result of population aging after 1980. The initial decrease in education during the 1970s, triggered by the baby boom cohort, offset the increase in education. This is why composition effects became important only after 1980. Holding the composition of the labor force fixed at its 1973 level, and re-weighting the CPS data from 2003, Lemieux finds only a very modest increase in wage dispersion.

A potential shortcoming of the paper by Lemieux [2006] is that it assumes away general equilibrium effects on skill prices. That is, he neglects the effect of cohort sizes on wages. Furthermore, his data set and thus his analysis is limited to the United States and the time until 2003. Moreover, the compositional effects in his analysis do not explain the polarization pattern observed by Autor, Katz and Kearney [2006] and others. This issue has been emphasized by Autor, Katz and Kearney [2005, 2008]. They argue that Lemieux's composition hypothesis falls short to explain the divergent trends in the upper and lower tails of the wage distribution. Neither changes in the aggregate nor the residual wage distribution between 1973 and 2003 are in line with changes in the composition. Autor et al. suggest that the impact of the labor force composition is limited to the lower tail of the distribution. Only by neglecting the different trends in the upper and lower tails, Autor et al. argue, is it possible to "fully explain the aggregate trend in residual inequality during the 1990s".[1]

Except for the studies by Lemieux [2006] as well as Autor, Katz and Kearney [2005, 2008], there are only a few papers on the demography-inequality nexus. Deaton and Paxson [1994, 1995, 1997, 1998] use a life cycle model and data from the U.S. and Great Britain to argue that within-cohort inequality should increase with the age of the cohort. With an aging population this should lead to greater national income inequality. In particular, both within-group and between-group inequality should surge. This is also based on Mincer [1974] who argues that depending on individual human capital accumulation people follow different

[1] In particular, Lemieux is said to over-explain the increase in the bottom of the wage distribution and under-explain it at the top.

slopes in age-earnings profiles.

1.4.6 Concentration of Top Incomes: Superstars and CEOs

Most of the increase in the dispersion of incomes took place in the upper part of the distribution. Especially incomes at the very top have surged in the last few decades [Atkinson, Piketty and Saez, 2011]. In Figure 1.1, I show that in the United States the Top 1% of income earners today receive an income share of almost twenty percent, twice as high as in 1980. Piketty and Saez [2003] as well as Atkinson, Piketty and Saez [2011] document the U-shaped development of top income shares in several English speaking countries during the twentieth century. For Continental Central European countries as well as Japan, the authors find an L-shaped development. The pattern for Nordic and Southern European as well as some developing countries is less clear. For Switzerland, a recent study by Föllmi and Martínez [2016] finds a significant increase in both the level and volatility of top income shares between 1981 and 2008.

A growing literature has investigated this trend. Possible culprits include technical advancements, globalization, the increasing size of corporations, changing social norms as well as new payment schemes.[1] Following the seminal work by Rosen [1981], both internationalization and technological changes —in particular in information and communications— tend to favor the emergence of superstar markets. In these markets, exceptionally talented individuals see an increased relative productivity. Today, athletes and artists increasingly operate on a global market and via TV and internet can reach a significantly larger audience than in the past. Hence they can sell the product of their labor to a vastly larger group of customers. However, Bakija, Cole and Bradley [2012] find that occupations in media, arts and sports only account for 1.7% of the top one percent of the income distribution. The majority of top earners are executives and managers in finance (13.2% in 2005) and non-finance (30.0%) corporations as well as medicals (14.2%) and lawyers (7.7%).

[1] As discussed before, the problem with changes in social norms (e.g., "greed is good") is that their impact is difficult to test empirically.

Surging top incomes are a challenge to most theories on rising income inequality. The supply of education, for example, as relevant as it may be for the overall distribution cannot explain why a few employees receive outstanding salaries. However, there is a growing literature on top incomes for superstars and chief executive officers (CEOs). With respect to superstars, a long literature following Rosen [1981] has documented the concentration of earnings on small groups of people. Gabaix and Landier [2008] provide empirical evidence that CEO salaries have increased sixfold between 1980 and 2003. This tremendous surge, Gabaix and Landier argue, can be fully rationalized by the sixfold increase in market capitalization of large firms during that period. This link between firms size and CEO compensation, however, emerged after 1980 and was absent in prior decades [Frydman and Saks, 2010]. Despite rapid firm growth after World War II, executive salaries remained flat until the mid-1970s. Hence, manager's ability to extract rents from the firm is discussed as a potential explanation for rising CEO salaries [Bebchuk and Fried, 2003; Kuhnen and Zwiebel, 2009]. Moreover, Piketty, Saez and Stantcheva [2014] find that pre-tax incomes of CEOs are lower in the presences of high top tax rates.

In related work, Kaplan and Rauh [2013] investigate wealth inequality at the top of the distribution. They document that the wealthiest Americans in the Forbes 400 today are less likely to have inherited their wealth or to have grown up in a wealthy family. Instead, these individuals possess a good education and apply their skills to the most scalable industries, including technology and finance.

1.4.7 Female Labor Supply

In all advanced economies, female labor supply has increased over the past decades [Aguiar and Hurst, 2007]. Figure 1.6 shows the steep increase in female labor force participation after 1970. This trends is not unique to any country but a general phenomenon throughout the Western countries. It is also notable that male labor force participation decreased slightly after 1970. These shifts had an effect on both competition and the composition of the labor force.

— Figure 1.6 about here —

The group of women in the labor force changed the overall composition with respect to age, experience, industry, and education. Moreover, women provide competition to men in various sectors, particularly in low-wage jobs because the median of the female wage distribution is typically below the male median wage. The competition effect depends on the substitutability between men and women in the labor market. Topel [1994, 1997] discusses this in detail. On the one hand, empirical analyses provide support for the hypothesis that increasing female labor supply affected the incomes and employment of low-wage men. On the other hand, women are typically occupied in different industries. Moreover the timing of changes in the female labor supply and income inequality does not fit very well.

The effect of compositional effects have been described by Mulligan and Rubinstein [2008]. They argue that increased wage inequality *induced* educated women to participate in the labor market. This explains the observed decline in the wage gap between men and women as well as the growing inequality within gender as documented by Katz and Autor [1999]. In a recent study by Blau and Kahn [2007], the authors use CPS data to document the large rightward shift in labor supply of women. In addition, they report a substantial decline in married women's responsiveness to their husbands' wages. Related to theories of marital sorting, this may have contributed to increasing inequality of household incomes. This is in line with Burtless and Karoly [1995] who find that the rise in female incomes after 1979 has occurred mainly in high-income households which further boosted inequality.

1.4.8 Sorting

Several theories discuss the effects of marital and educational sorting on inequality. They are based on the observation that today individuals increasingly choose partners with similar education. This leads to lower within-household income differences but larger between-household inequality. Moreover, inequality could rise over time if children's income potentials are a function of parental education.[1]

Empirically, an early study by Kremer [1997] finds only minor effects of marital and educational sorting on inequality. Fernández and Rogerson [2001], however, disagree with this findings and provide a dynamic model which incorporates two skill groups, marriage, fertility, education, and the determination of income. The difference between their model and Kremer's lies in the interaction between changes in the skill distribution and the prices of skills. In addition, Fernández and Rogerson allow for a nonlinear relationship between parental and their children's education, a negative correlation between education and fertility, as well as wages to depend on the skill distribution. With these modifications, the authors find marital sorting to have a substantial effect on income inequality. In a subsequent paper, Fernández [2002] builds an OLG model which also features intergenerational transmission of education: Children are more likely to develop skills when having more educated parents. Calibrating the model with data from the United Kingdom, the author shows how marital sorting leads to increased inequality between skilled and unskilled individuals.[2] Most recently, Greenwood et al. [2014] find that assortative matching has had a great impact on inequality. If matching between husbands and wives in 2005 had been random, the Gini coefficient would have fallen from the observed 0.43 to 0.34, reflecting a tremendous reduction in inequality. Additional work by Eika, Mogstad and Zafar [2014] finds that primarily low educated individuals increasingly sort themselves into internally homogeneous marriages.

It is important for theories on marital sorting to take into account within-household re-allocations. Lise and Seitz [2011] provide empirical evidence for the importance of within-household consumption allocations. While between-household inequality has risen in the past, an offsetting reduction in within-household inequality took place. Marital sorting appears to be the most important explanation for this trend. This gives rise to the question which factors actually

[1] Related to the literature on sorting is a study by Burtless and Karoly [1995] on the effects of family structures on inequality. The authors find that the increase in the proportion of single-head families boosted inequality between 1959 and 1989.

[2] Ermisch, Francesconi and Siedler [2006] provide further empirical evidence for the importance of marital sorting in both Germany and Great Britain.

determine the prevalence and timing of marital sorting. Friesen and Krauth [2007] use data from Canada and argue that sorting can already occur in schools. If students are sorted across schools by parental education, existing inequality can be amplified. Charles, Hurst and Killewald [2013], on the contrary, focus on sorting among adults. They use data from the Panel Study of Income Dynamics (PSID) and examine whether spouses sort on the basis of parental wealth. This differs from previous studies which focused on spouse's education. The authors find a significant correlation between parental wealth of married individuals in the United States.

1.4.9 Consumer Preferences and Product Demand

When individuals purchase goods and services they implicitly pay for specific types of labor that went into the production of these goods and services. Thus changing consumer demand for products with varying skill-intensity can affect the distribution of income. As an example described by Topel [1997], an increase in the demand for skill-intensive luxury goods may boost the dispersion of wages.

For the most part, however, empirical research has found little evidence for a significant impact of changes in the industry composition on inequality. This does not imply, however, that product demand plays no role at all. Being a side effect of the structural change, it may explain the rise in the service industry which has been documented by Buera and Kaboski [2012] as well as Autor and Dorn [2013]. Consumer preferences are also crucial for the extent to which technological change affects employment and wages. If we assume that individuals do not admit close substitutes for the tangible outputs of service occupations (e.g., restaurant meals, house-cleaning or security services) we can explain the surge in service employment [Autor and Dorn, 2013]. Furthermore, the effect of technological change on inequality is mitigated because consumer preferences limit the number of tasks and jobs which will be automated.

A few recently published studies investigate the impact of product demand on income inequality directly. Mazzolari and Ragusa [2013] use the idea that skilled workers face higher opportunity costs of time and thus spend more on market substitutes for home production activities. If the skill premium increases —for example due to technological changes— the employment of low-skill workers in service occupations increases. Mazzolari and Ragusa use data from the United States to support this prediction empirically. Leonardi [2015] extends this work and argues that highly educated individuals prefer to consume goods and services which are relatively skill-intensive. Under this assumption, if the supply of skilled workers increases the demand for skills is heightened. Using data from the Consumer Expenditure Survey, Leonardi documents also that the share of highly educated heads of households increased from 27 to 62 percent between 1972 and 2012. Over the same period, the total expenditure share for high-skill services (e.g., health and education) increased while demand for low-skill services (e.g., food and apparel) declined. In total, this consumption mechanism can explain about 6.5 percent of the shift in relative labor demand both in the United States and United Kingdom.

1.4.10 Fiscal and Monetary Policy

The government can influence the distribution of income in many ways. In this paper, I already discussed interventions in the labor market and education. Here, I discuss the impact of fiscal and monetary policy.

A recent strand of literature investigates how tax policies affect the distribution of *pre*-tax income. As already mentioned, Piketty, Saez and Stantcheva [2014] find that pre-tax incomes of CEOs are lower in the presence of high top tax rates. More generally, Guvenen, Kuruscu and Ozkan [2014] argue that progressive taxes distort the incentive to accumulate human capital which in turn reduces the cross-section dispersion of before-tax wages. Although most of the literature on the determinants of inequality trends deals with wages before taxes, it is important not to neglect the effects of taxation and public spending on inequality. Davies and Hoy [2002], for instance, document a trend towards flat rate tax systems. The authors argue this typically benefits the upper and lower tail of the income distribution at the expense of the middle class.

In addition to shaping tax policies, governments can affect income inequality through spending decisions as well. Cozzi and Impullitti [2010] argue that government expenditures can explain up to 15 percent of the observed increase in wage inequality between 1976 and 1991. This is the result of increased public investments in high-tech sectors. During the 1980s and 1990s, the share of investment in equipment and software rose from twenty to fifty percent, causing a re-allocation of demand from low- to high-skill industries. Simply put, policymakers increased the pace of skill-biased technological change in the United States.

The impact of *monetary* policy on income inequality has often been neglected. This is because *fiscal* policy is considered to be the appropriate tool to alter the distribution of income. Recent work by Coibion et al. [2012] as well as Airaudo and Bossi [2015], however, finds that contractionary monetary policy can significantly raise the Gini coefficient. This is the result of many potential channels. Low- and high-income individuals differ, for example, in their portfolios, asset market participation rates, primary source of income, or savings rates. All this plays a role when investigating how monetary policies affect their relative well-being.

1.5 Interplay of Factors

After the discussion of which factors influence the distribution of income, I now discuss in some detail how the impact of one factor may depend on others. These interaction effects are largely neglected in the literature. The lion's share of studies who consider several factors dismiss all but one. And even if a study admits that more than one factor is found to have a significant effect, the interplay is largely neglected.[1] In contrast, this section sheds some light on potential interaction effects which may offer interesting perspectives for future research.

Consumer Preferences and Technological Change — By and large, technological change is considered to be the most important driver of changes in the distribution of wages. However, the extent to which new technologies actually af-

[1] Autor and Dorn [2013], for example, investigate the impact of factors other than unbalanced technological progress. However, all other variables —including offshoring, rising demand for home production substitutes, and growing low-skill immigration— are said to be negligible with respect to their findings.

fect labor markets also depends on consumer preferences. Following the seminal work by Autor and Dorn [2013], technological change leads to a polarization of the labor market with a notable increase in service-sector jobs. There is, however, no reason why service jobs are immune to the impact of technology. Restaurants, for instance, can use touch-screen computers to replace its cashiers.[1] Due to consumer demand, however, this practice has been limited in its application to low-price fast food restaurants despite the fact that it could be used in virtually all restaurants. Similarly, banks could replace most of their employed tellers by installing additional ATMs. This is not optimal though if customers value direct interaction with the banks' staff. Hence, the employment boom in the service sector documented by Autor and Dorn [2013] has not only be caused by technological change but because of consumer preferences.

Mitigating Impact of Labor Market Regulation — Technological change and import competitions have been found to significantly affect labor markets in developed countries. Their impact, however, depends on the institutions—minimum wages, union density, etc.—of the labor market. If, for example, a new technology reduces the value of some task, the effect on wage inequality depends on whether there is a statutory wage floor. In the absence of a minimum wage, technological change lowers wages and within-firm inequality increases. In contrast, if there is a binding minimum wage we should observe employment effects and within-firm inequality might decrease. Similarly, the presence of a high union density might spread the benefits of technological improvements. However, if disturbances caused by labor market institutions are too large, high-skill workers might decide to leave unions [Acemoglu, Aghion and Violante, 2001]. As a result, labor market regulations can potentially mitigate the effects of technological change or international trade on income inequality. However, such regulations are endogenous and respond to wage pressure caused by SBTC or globalization. This makes estimating the interaction effect particularly challenging.

[1] One of the first restaurants to use this technology was McDonald's in the late 1990s according to the article 'McDonald's Testing Self-Serve System', published by *The New York Times* on August 12, 1999.

Technology and Trade — Following work by Autor, Dorn and Hanson [2015], both technological change and international trade are major forces shaping the distribution of wages. Their impact, however differs substantially. While the authors find little employment effects in industries affected by routine-task specialization, import competition is associated with sharp declines in manufacturing employment.

1.6 Conclusion

Growing income inequality is not just a concern for normative discussions. If the concentration of income and wealth leads to a concentration of political power, there is reason for discussion on positive grounds. Moreover, income inequality can also effect economic development through a multitude of channels.

A large and growing body of literature has been devoted to examine which factors shape the distribution of income. In general, there appear to be two ways to identify which factor has contributed how much to the changes in the income distribution: First, one can exploit time trends. The surge of new technology such as computers or robots, for example, has been an ongoing trend since the 1980s in virtually all countries. In contrast, tax policies or minimum wages have been altered at specific points in time. The second path for identification is to exploit the fact that countries differ in their policies. While all countries have experienced technological progress and globalization, tax and labor market policies often differ significantly across countries. Precisely matching a country's trends in the various measures of income inequality to its policies, however, remains a challenging task for future research.

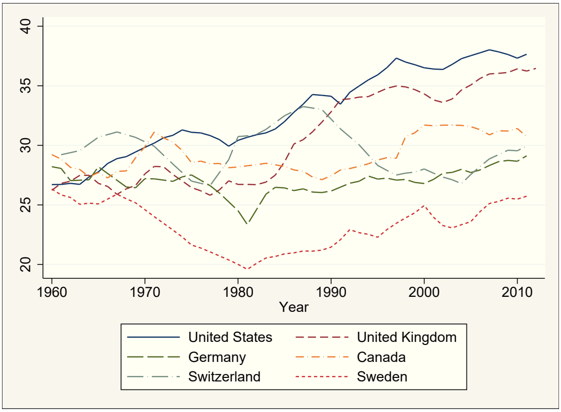
One broad conclusion of the literature review in this paper is that no single factor was solely responsible for rising income inequality [Autor, Manning and Smith, 2016]. The most significant contributing factors include a change in the demand for skills that was driven by technological advancements as well as international trade. In the United States, this shift in demand met with a slowdown in the supply of college graduates. In other countries, the increase in the supply of skills did not keep pace with demand either. Second, declining union densities and real minimum wages can explain, to some extent, the sharp increase in wage inequality during the 1980s. Third, the increasing dispersion of incomes in the upper part of the income distribution is a result of superstar markets and, to a lesser extent, altered tax policies. Finally, the literature finds a significant yet small impact of immigration, demographic changes, female labor supply, sorting, and consumer preferences.

In terms of policy conclusions, my research does not intend to provide any particular recommendations. Instead, the goal is to improve the understanding of the determinants of inequality in order to help overcome several puzzles in the literature. In this, my research is in line with Deaton and Paxson [1997] who wrote that a sound understanding of what drives income inequality is necessary "if only to avoid the imposition of unnecessary policies designed to correct it" (Deaton and Paxson, 1997, p.97).

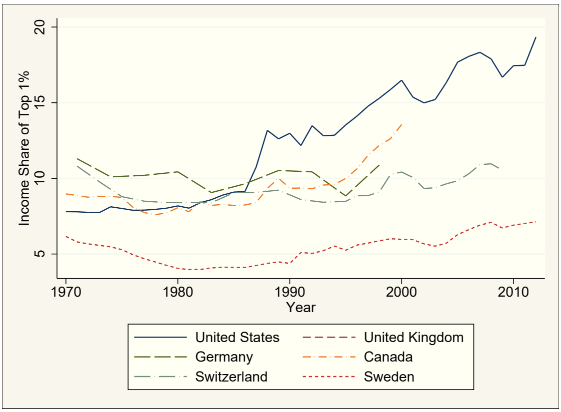
Figures

Figure 1.1: Gini and Top Income Shares in Selected Countries

(a) Gini Indices

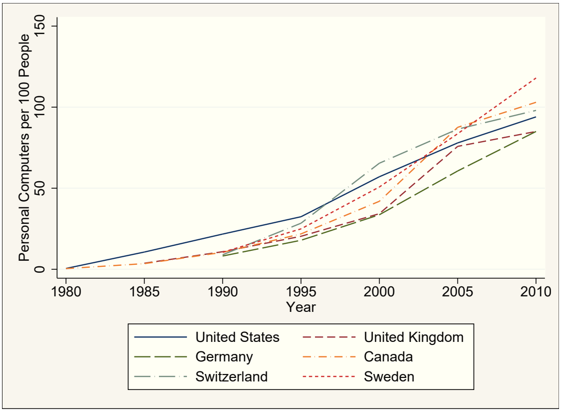


(b) Top Income Shares



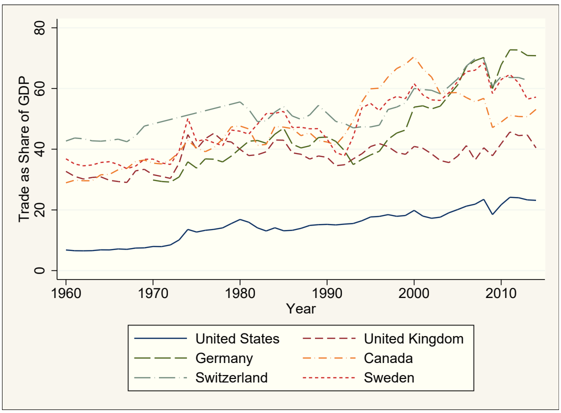
Note: The figures show the evolution of income inequality in six selected countries. In Panel (a) the Gini indices based on net income post taxes and transfers are plotted. In Panel (b) the income shares of the top 1 percent are shown. Data Sources: Solt [2009]; Alvaredo et al. [2014]

Figure 1.2: Diffusion of Personal Computers in Selected Countries



Note: The figure shows the number of personal computers (PC) per 100 people in six selected countries. Data Source: International Telecommunication Union (ITU)

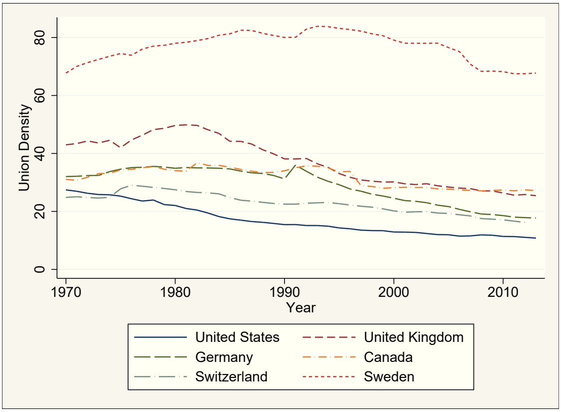
Figure 1.3: Trade as a Share of GDP for Selected Countries



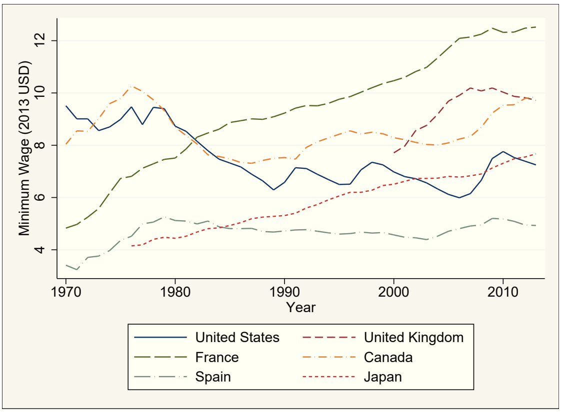
Note: The figure shows the evolution of total trade (imports and exports) over GDP for five selected countries. Data Source: World Development Indicators

Figure 1.4: Unions and Minimum Wage for Selected Countries

(a) Trade Union Density

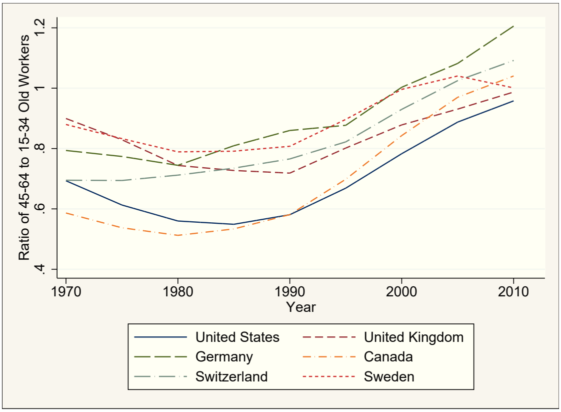


(b) Real Minimum Wage



Note: The figures show the evolution of trade union density (a) as well as the real minimum wage (b) for selected countries. Trade union density corresponds to the ratio of wage and salary earners that are trade union members, divided by the total number of wage and salary earners. Statutory minimum wages are converted into a common hourly and annual pay period. The resulting estimates are deflated by national Consumer Price Indices (CPI) and then converted into a common currency unit using USD current exchange rates. Data Source: OECD Labor Force Statistics

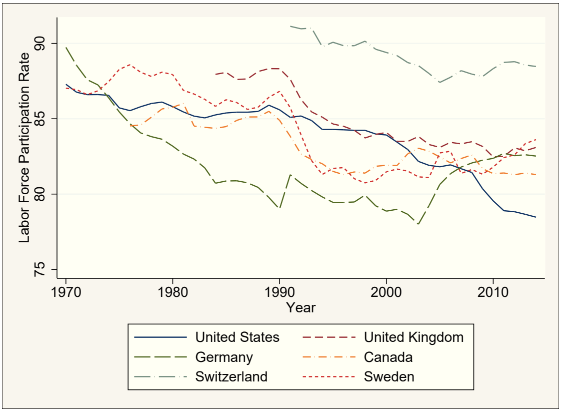
Figure 1.5: Senior to Young Workers in Selected Countries



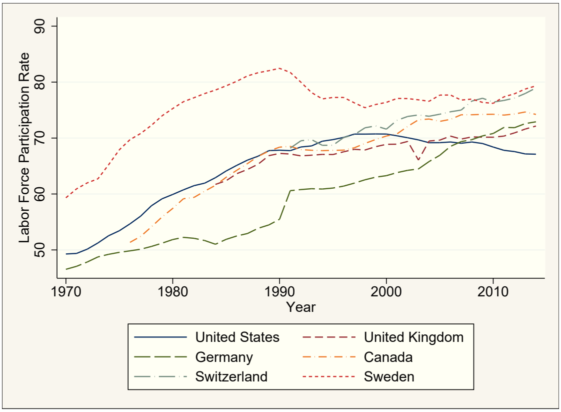
Note: The figure shows the ratio of 45-64 to 15-34 year-old individuals in six selected countries. Data Source: UN Population Division

Figure 1.6: Labor Force Participation for Selected Countries

(a) Male



(b) Female



Note: The figures show the evolution of labor force participation rates among men (a) and women (b) for selected countries. The population is restricted to 15-64 year olds. Data Source: OECD Labor Force Statistics

Chapter 2

Innovation in an Aging Population

"Anything that is in the world when you are born is normal and ordinary and is just part of the way the world works. Anything that is invented between when you are fifteen and thirty-five is new and exciting and revolutionary and you can probably get a career in it. Anything that is invented after you are thirty-five is against the natural order of things." — Douglas Adams, The Salmon of Doubt (2002)

2.1 Introduction

What is the impact of population aging on an economy's rate of innovation? In the coming decades virtually all Western countries face significant demographic changes. This is the result of the baby boom in the 1950s and early 1960s, followed by historically low fertility rates afterwards. Tempo-adjusted total fertility rates (TFR), which account for the postponement of childbearing, are way below the replacement level in most of the OECD and EU countries.[1] Irrespective of whether past fertility trends continue, there will be tremendous changes in the age structure of most rich countries. While the first major economies already face a decline in population size the various repercussions of demographic trends have received new attention [Last, 2013].[2] A central question in this regard is what are the implications of this demographic shift for the economic development of affected countries. Whereas some studies suggest that low fertility rates favor the standard of living [Lee and Mason, 2014], others provide a more ambiguous outlook [Bloom, Canning and Fink, 2010; Prettner, 2013].

This paper investigates how population dynamics affect an economy's rate of innovation. In a first step, I show that in the past three decades the share of the working-age population has not changed substantially but hovered around 65 percent in OECD countries. However, while this group was dominated by young workers in 1980, the demographic structure of 2010 looks remarkably different. The baby-boom cohort entered the labor market in the 1970s or 1980s and is now about to leave. Computing the ratio of senior (aged 45 to 64 years) to young (15 to 34 years) workers, I document a tremendous increase after 1980. To study how this affects the economy, my theoretical model assumes that individuals live for two periods and have to invest in knowledge in order to possess the necessary skills for using innovative goods. For young and old individuals in the model there are different, limited time windows for such skill investments to pay off. As a result, people of different age have different preferences with respect to innovative products. While young individuals demand innovative goods, the old cohort abstains from learning how to use the latest technology. It follows that in an aging population, aggregate demand for innovative goods declines. As a result, the economy's rate of innovation slows down.

While these predictions arise from a theoretical model, I examine them using a novel data set that contains information about patent applications for all

[1] For Japan (1.47) or the EU-27 (1.77), the adjusted TFR is already way below replacement level. The United States (2.14) is exactly at the threshold. More detailed statistics are provided by Bongaarts [1999] and in Table 2.1 in the appendix.

[2] To illustrate the magnitude of demographic trends, in an article entitled 'The incredible shrinking country' (published on March 23, 2014), *The Economist* finds that Japan lost roughly the population of Jersey City (or 244,000 people) in 2013.

OECD countries for the years 1978 to 2010. The empirical results show that growth in patent applications per capita is lower in countries with an older workforce. For almost all countries, an aging population has been associated with less innovative activity. In line with the idea of the theoretical model, the decline in patent applications appears to be driven by patents on information and communications technology (ICT). In order to test the robustness of the negative effect of population aging on innovation, I apply an instrumental variables approach using demographic information from earlier periods as instrument. The results confirm the finding that population aging has a negative impact on the growth rate of patent applications. Finally, I show that those countries that experienced the largest demographic shifts between 2000 and 2010 are the ones with the most negative trend in patent growth.

My research is related to different strands of the literature. First, a couple of studies have investigated how demographic trends affect the economy in general and innovation in particular. The work by Acemoglu and Linn [2004] is most closely related to my paper. The authors show that due to demographic trends the size of various age groups changes over time. Since most pharmaceutical drugs are mainly used by a specific age group, demographics affect the market size for each drug. Acemoglu and Linn use this mechanism to establish empirically a positive market size effect on innovation. A related paper by DellaVigna and Pollet [2007] finds that demographic trends can be used to predict changes in demand for age-sensitive sectors. Linking consumer preferences and product demand to innovation, Föllmi and Zweimüller [2006] as well as Föllmi, Würgler and Zweimüller [2014] provide theoretical models which explain firms' decision to undertake either product or process innovations. The idea that demand affects the rate of innovation is also supported empirically by Miao and Popp [2013]. In particular, the authors show that natural disasters lead to an increase in risk-mitigating innovations. Finally, research by Strulik, Prettner and Prskawetz [2013] provides a model that predicts the changing correlation between population growth and innovative activity in the process of economic development.

A large body of literature has established that innovation is key for economic growth (e.g., Aghion and Howitt, 1992). Although economic growth is not a primary concern of this paper, I show how the incentives to innovate are affected by an aging population. A negative effect of demographics on growth, however, is not clear [Becker, Glaeser and Murphy, 1999]. An empirical study by Ahituv [2001] finds that a decrease in population growth is associated with higher GDP per capita growth. In general, demographic trends can affect the economy through its implications on, for example, social security, savings rates and capital accumulation, the business cycle, or education [Cutler et al., 1990; De La Croix and Licandro, 2013; Jaimovich and Siu, 2009]. In this regard, Krueger and Ludwig [2007] show that the demographic transition towards an older population reduces return to capital while having a positive effect on wages.

My paper adds to the literature by suggesting a new mechanism which links population aging to innovation. Notably, this mechanism as well as my theoretical model addresses the *demand* side of innovation and remains silent about the *supply* side. Following Acemoglu [2002*a*], I argue that in an aging population the demand for certain innovative goods declines. The reduced market size for such goods then leads to a reduction in the respective R&D efforts. While I also find some empirical evidence for supply side effects in the data, I do not address this in my theoretical model as it is not the focus of this paper.

The paper proceeds as follows. In Section 2.2, I illustrate demographic trends in the past decades. Section 2.3 provides a simple theoretical model which illustrates the mechanism through which population aging affects the rate of innovation. In Section 2.4, I test the model's predictions empirically. Finally, Section 2.5 concludes.

2.2 Demographic Trends

In this section, I document several demographic trends which occurred in major economies over the past couple of decades. Particular emphasis is put on changes in the age composition of the workforce. For this I define a new measure to capture population aging: the senior-to-young worker ratio.

2.2.1 Fertility Rate

Most Western countries experienced substantial fertility shocks in the past and have had very low levels of fertility since the 1970s. In Figure 2.1, the fertility rates of six selected countries since 1960 are shown.[1] It was during the 1970s when the fertility rate fell below 2.1 (i.e., the replacement level) in several countries. And despite some variation afterwards, most countries countries kept a sub-replacement rate of fertility. Even France and the United States have a fertility rate that is barely at the replacement level.

— Figure 2.1 about here —

Historical data on fertility rates reveals substantial shocks over time. Figure 2.1 in the appendix indicates that the fertility rate in Germany declined sharply after 1910, experienced a boom between 1950 and 1965, and has been below the replacement level since 1970. The baby boom is clearly visible and can be observed for many countries. For the empirical analysis in Section 2.4, I exploit this boom as exogenous variation. Jones and Tertilt [2008] document similar demographic trends for the United States: The total fertility rate (TFT) fell steeply from about 5.5 in 1850 to 2.4 in 1940, reaching a temporary high of 3.5 during the baby boom period around 1960, but fell afterwards to about 2.0 in 1990. Since then, the TFR has remained roughly stable although this is largely driven by the relatively high fertility rate of Hispanics (2.4) and Blacks (2.0). For white non-Hispanic Americans, the TFR is about 1.8 and thus below the replacement level.

Both observed trends in fertility —temporary shocks as well as the fall below replacement levels— caused substantial changes in the composition of the population. Today almost all of the world's population lives in countries with declining fertility rates. This finding does not change when taking into account the postponement of childbearing. The so-called tempo-adjusted total fertility rates are also found to be substantially below replacement levels. While this adjustment

[1] Note that the depicted fertility rates are not tempo-adjusted. However, such adjustments increase the fertility rate by about 0.2 and do not alter the trend. Fertility data for a larger set of countries is provided by Table 2.1 in the appendix.

usually raises the fertility rate by about 0.2, for Japan (1.47), Switzerland (1.69), as well as the EU-27 (1.77), the adjusted TFR is already very low. And the United States (2.14) is about to pass the threshold to sub-replacement levels.

These low fertility rates already caused significant demographic shifts and if trends of the past continue, larger compositional changes will occur. This in turn will have implications that have been described as the root cause of many economic and social problems [Last, 2013]. Most notably, with a continuous sub-replacement level of fertility, the population will shrink in size and become older. Both developments can affect the rates of innovation and investment as I argue in this paper. However, at least since the Industrial Revolution there has not been any country experiencing sustained, structural population decline. Today, only Russia and Japan with their already shrinking populations may indicate the impact of sustained below-replacement levels of fertility.[1]

2.2.2 Compositional Changes

Fertility rates of the past decades altered the composition of the population in all countries. Due to the baby boom cohort, now in their 50s and 60s, the demographic shifts are particularly sizable.

Shares of Young, Adults, Retirees — The most straightforward way of illustrating compositional changes in the population is to define three age groups: children aged 0-19, adults aged 20-64, and retirees aged 65 or more. Figure 2.2 in the appendix shows how the share of each group changed after 1950. In addition, projections for the future are shown assuming a constant fertility rate. The plots illustrate the tremendous demographic shifts. Most important is the impact of the baby boom period from 1950 to the early 1960s. First it increased the share of 0-19 year olds. Then, between 1970 and 2020, the share of working-age people is extraordinarily high. Starting around 2020, however, the share of retirees will

[1] There are two major historical incidents of shrinking populations. The first occurred in the Roman Empire between A.D. 200 and 600 and marked the descent into the Dark Ages. The second incident was caused by the Black Death between 1340 and 1400 with world population shrinking from 443 to about 374 million.

increase sharply. This share has always increased due to rising life expectancy but the positive trend will increase once the baby boom cohort enters retirement age. Much of the economic research on population aging has focused on the implications of this latter group. In particular, financing retirement schemes has been subject to intensive research. However, for this paper I focus on the working-age population and how it changed due to demographic trends in the past.

Shares of Young and Senior Workers — Splitting the population into children, adults, and retirees shrouds a remarkable shift among the adult group: The shares of senior (45-64 year old) and young (15-34) workers have changed substantially in OECD countries. In order to document this, I define senior-to-young worker ratio (henceforth *S2YWR*) as

$$S2YWR = \frac{\text{share } (45 - 64 \text{ year old})}{\text{share } (15 - 34 \text{ year old})} \tag{2.1}$$

This ratio is motivated in two ways: First, in most countries people usually work when their age is somewhere between 15 and 65. While careers differ greatly across individuals, it is possible to argue that people outside this age bracket account for a tiny fraction of the total labor force. With this information, we can state that the age group of 15–34 year olds broadly covers the youngest people in the labor market. Conversely, the 45–64 year olds mark the most senior group. Using the ratio defined above is interesting for another reason. In the United States, the baby boom period lasted from 1946 to 1964. Thus in the year 1980, the baby boom cohort fell exactly in the group of “young workers”. Thirty years later in 2010, the baby boomers are “senior workers”. As a result, the senior-to-young worker ratio was very low in 1980 and has increased ever since. This is shown in Figure 2.2, not only for the United States but also for China, Germany, and Japan.

— Figure 2.2 about here —

Since the baby boom was a phenomenon observed in many countries, there are similar trends in the *S2YWR* for all four of the world’s largest economies.

Together these four countries account for roughly half of the world GDP. For the set of OECD countries we observe very similar trends. There is not a single country in which the ratio did not increase.[1] Between 1980 and 2010, the overall mean of the *S2YWR* changed from 0.63 to 0.96, a 52 percent increase in thirty years.

— Figure 2.3 about here —

The remarkable shift in the demographics of the working-age population is illustrated by Figure 2.3. The entire distribution of the *S2YWR* shifted to the right. Most notably, there was no OECD country with a *S2YWR* exceeding unity in 1980. In contrast, today the majority of countries has a larger share of senior workers than young workers. The next section presents a theoretical model which describes a particular mechanism how the shift in the *S2YWR* affects innovation.

2.3 Theory

This section describes the setup and steady state equilibrium of an overlapping-generations model that features age-dependent preferences arising from necessary investments in skills for consumption. The model builds upon prior work by Acemoglu and Linn [2004] but adds a novel mechanism linking population aging to an economy's rate of innovation. Furthermore, in my model individuals choose a preferred level of product innovation and quality.[2]

2.3.1 Setup

Population and Demography — The economy is populated by a discrete number of generations denoted by $t \in \mathbb{N}^+$. All individuals live for two periods: young

[1] Figure 2.3 in the appendix shows the trends in the *S2YWR* for every OECD country between 1980 and 2010.

[2] In the model by Acemoglu and Linn, individuals do not choose a specific quality but are indifferent between the best and second-best quality level while spending a constant share of their income on the innovation-related good. The rate of innovation is determined by the market size for each drug which in turn affects firm competition and R&D efforts to be the firm with the highest quality.

and senior adulthood. For simplicity, I assume away childhood and retirement. Population growth is determined by the exogenous fertility rate n_t. In particular, young adults L_t^Y of period t have $(1+n_t)L_t^Y$ children. These will then become young adults at date $t+1$ and senior adults in period $t+2$. By assumption, raising children comes at no cost. All decisions are made at the beginning of adulthood when individuals decide about how much to consume and how much to invest in skills.

In any period t, there are L_t^Y young adults as well as L_t^S senior workers. It holds that $L_{t+1}^Y = (1+n_t)L_t^Y$ and $L_{t+1}^S = L_t^Y$. This results from the fact that all individuals live for exactly two periods. The total population grows for $n_t > 0$, shrinks for $n_t < 0$, and remains constant for $n_t = 0$. The ratio of senior to young workers ($S2YWR$) is given by

$$\frac{L_t^S}{L_t^Y} = \frac{1}{1+n_{t-1}} := S2YWR_t. \tag{2.2}$$

A baby boom period can be illustrated by an increase in n_t. First, this leads to a decline in $S2YWR_{t+1}$ when the large cohort enters the labor market. Subsequently, however, there is an increase in $S2YWR_{t+2}$ when the baby boom cohort turns into senior workers.

Utility and Types of Goods — The economy features two different types of goods. First, a basic good denoted by y. This can be consumed, used for production or for research expenditures. Second, there is a sophisticated good x which can be produced at different quality levels q and requires skills for consumption. Each individual has an exogenously given endowment y_t in both life periods.[1] Preferences are given by

$$U_t = u_t + r^{-1}u_{t+1} \tag{2.3}$$

[1] An alternative is to assume inelastic supply of one unit of labor each period at wage rate y_t. Note that for simplicity I abstract from modeling senior workers to earn more than young workers. A study by von Weizsäcker [1996] provides a discussion of the impact of population aging on income inequality.

$$\text{with} \qquad u_t = c_t^{1-\gamma}(q_t x_t)^\gamma \qquad \text{and} \qquad u_{t+1} = c_{t+1}^{1-\gamma}(q_{t+1} x_{t+1})^\gamma \tag{2.4}$$

where r is the discount rate of consumers (and the economy's interest rate) and $\gamma \in (0,1)$. Consumption of the basic good (y) is denoted by c_t while x_t is the amount of the sophisticated good consumed in period t. In order to simplify the analysis, it is assumed that $x_t \in \{0,1\}\ \forall t$. Given the utility function, this implies that individuals always consume one unit of the sophisticated good and only choose the quality thereof. This assumption as well as the Cobb-Douglas functional form are for simplicity.

The price of the basic good is normalized to 1 in all periods (numeraire) while p_t denotes the (relative) price of the sophisticated good. In the absence of bequests, savings, or borrowing, the budget constraints for the two periods read

$$y_t \geq c_t + p_t(q) + e_t \qquad \text{and} \qquad y_{t+1} \geq c_{t+1} + p_{t+1}(q) + e_{t+1} \tag{2.5}$$

where $p_t(q)$ and $p_{t+1}(q)$ are the price of the chosen quality q, and e_t as well as e_{t+1} reflect investments in skills in the two periods. These are necessary for the consumption of higher quality versions of the sophisticated good.

Investments for Innovative Goods — Each individual in the economy can only consume higher qualities of the sophisticated good if she has the necessary skills. While it takes no learning to consume quality q_{t-1} for someone born at time $t-1$, it requires an investment of $e_t = \phi(q_t - q_{t-1})$ with $\phi > 0$ to be able to use the state-of-the-art quality q_t.[1] This assumption can be motivated by an example. For every innovative good (e.g., computers) adults have to spend time learning how to use it. This investment is not necessary for more established goods (e.g., telephone) if individuals grew up at a time when these were already available. An individual born in 1960, for instance, grew up in a world with widespread use of telephones but no computers. When the latter were introduced during the 1990s, this person was in his thirties and had to spend a considerable amount of time on learning how to use computers.[2] This kind of investment is

[1] The way I model the acquisition of knowledge is similar to Garicano and Rossi-Hansberg [2015] who assume a fixed cost of learning a unit length of solving problems.

[2] Sometimes the terms 'digital natives' and 'digital immigrants' are used to distinguish indi-

crucial for the model's dynamics. A simplifying assumption that I make is that the costs of learning does not depend on an individual's age. This is in line with recent research by on increases in healthy life expectancy (cf. Reuter-Lorenz and Park, 2014 as well as Strulik and Werner, 2016). Note that it is possible to extend the model to allow skills to deteriorate over time:

$$E_t = e_t \quad \text{and} \quad E_{t+1} = E_t(1-\xi) + e_{t+1} \quad \text{with } \xi \geq 0 \tag{2.6}$$

However, for the baseline model, I assume that skills, once obtained, are neither lost nor unlearned. This appears to be the more relevant case. Thus I assume $\xi = 0$ throughout the model. Moreover, the cost of learning how to use new technology is assumed to be the same for an individual in her young and senior period.

R&D and Production — At any time t there is one firm with the technology to produce the best quality q_t of the sophisticated good. This firm can produce one unit of x_t at quality q_t using one unit of the basic good y_t. The marginal cost of producing x are one irrespective of the quality level. In order to achieve one unit increase in quality, a firm has to spend $\delta > 0$ units of the basic good. If a firm develops a new quality, it receives a patent for one period. The total R&D spending for the quality improvements demanded by consumers in period t is given by $z_t = \delta(q_t - q_{t-1})$. By assumption there is free entry into R&D and each (potential) firm has access to the same research technology.[1]

Given that the marginal costs of production for all quality levels is equal to one, the firm with the best quality level at time t faces a competitive fringe by competitors. Hence it will only be able to charge $p_t(q_t) = 1 + \delta(q_t - q_{t-1})$ for its quality. This results from the fact that in period t the price of the quality level q_{t-1} drops to one as the patent expires. Any better quality, first available in period t, is offered at price $p_t(q_t)$ which increases in q_t.

viduals who grew up in a digital world from those growing up in an earlier time.

[1]Note that following Aghion and Howitt [1992] it is assumed that the firm with the best quality does not invest in R&D itself.

Choosing Quality and Utility Maximization — In order to illustrate how individuals optimize their lifetime utility, consider the cohort of individuals born in period $t-1$. Upon entering adult life, this group chooses an optimal strategy for both periods t and $t+1$. They determine the consumption of the basic good in both periods, c_t and c_{t+1}. Moreover, they choose how much to invest in skills: e_t and e_{t+1}. When deciding whether to purchase a quality higher than q_{t-1} —which can be purchased at a price of one and consumed without skill investments— the gain in utility must outweigh the costs. In particular, individuals can choose between $\{c_t, q_t\}$ and $\{c_t + (\delta+\phi)\Delta q_t, q_{t-1}\}$, where $\Delta q_t = q_t - q_{t-1}$ and $\delta(q_t - q_{t-1})$ as well as $\phi(q_t - q_{t-1})$ indicate the higher price and the necessary skill investments, respectively. If the first combination, $\{c_t, q_t\}$, is strictly superior, individuals will invest $e_t > 0$ in skills. In the second period $t+1$ the cohort is in senior age. For them to abstain from further skill investments ($e_{t+1} = 0$) and continue consuming q_t it must hold that $\{c_{t+1} + (\delta + \phi)\Delta q_{t+1}, q_t\}$ is preferred over $\{c_{t+1}, q_{t+1}\}$. To satisfy both conditions, I impose the following assumption.

Assumption 1. *It is assumed that*

$$c_{t+1}\left(q_{t+1}^{\gamma/(1-\gamma)} q_t^{-\gamma/(1-\gamma)} - 1\right)/\Delta q_{t+1} < \delta + \phi < c_t\left(q_t^{\gamma/(1-\gamma)} q_{t-1}^{-\gamma/(1-\gamma)} - 1\right)/\Delta q_t.$$

where I use the fact that that $p_{t+1}(q_t) = 1$. The assumption implies that the cost of innovating (δ) and learning (ϕ) are not too large to prevent individuals from choosing q_{t-1} over q_t but large enough to not invest in skills as senior adults. The optimality of investing in skills only when young (i.e., $e_t > 0$ and $e_{t+1} = 0$) requires that

$$\left(c_t^{1-\gamma} q_t^{\gamma}\right) + \frac{1}{r}\left(c_{t+1}^{1-\gamma} q_t^{\gamma}\right) \geq \left(\tilde{c}_t^{1-\gamma} q_t^{\gamma}\right) + \frac{1}{r}\left(\tilde{c}_{t+1}^{1-\gamma} q_{t+1}^{\gamma}\right) \tag{2.7}$$

with $c_t = \tilde{c}_t = y_t - 1 - (\delta + \phi)\Delta q_t$, $c_{t+1} = y_{t+1} - 1$ and $\tilde{c}_{t+1} = y_{t+1} - 1 - (\delta + \phi)\Delta q_{t+1}$. For condition (2.7) to hold, I impose the following assumption.

Assumption 2. *It is assumed that* $\delta + \phi \leq \Delta q_{t+1}^{1/1-\gamma}/(\Delta q_{t+1}^{\gamma/1-\gamma} - y - 1)$.

This implies that, in general, a cohort born at $t-1$ lives and consumes quality

q_t in periods t and $t+1$. It follows that only the young generation invests in skills and demands a higher quality level. In order to maximize lifetime utility, individuals choose a quality level given by

$$q_t^* = \frac{\gamma}{\delta + \phi} y_t + \gamma q_{t-1}. \tag{2.8}$$

It is straightforward to see that the preferred quality is increasing in γ and y_t but decreasing in the cost of innovating (δ) and learning (ϕ). Using equation (2.8) and the fact that a firm has to invest $z_t = \delta(q_t - q_{t-1})$ for the quality improvement, we get that

$$z_t^* = \delta\left(\frac{\gamma}{\delta + \phi} y_t - (1-\gamma) q_{t-1}\right) \tag{2.9}$$

which shows that total R&D spending is increasing in y_t. Profits of the firm with the best technology in period t are given by

$$\pi_t(q_t) = \delta(q_t - q_{t-1}) L_t^Y = \delta\left[\frac{\gamma}{\delta + \phi} y_t - (1-\gamma) q_{t-1}\right] L_t^Y. \tag{2.10}$$

Note that patent protection expires after one period. Thus the firm with the best quality in one period receives no profits in subsequent periods: $\pi_{t+1}(q_t) = 0$. This is because both the price and the marginal cost of quality q_t are one in period $t+1$.

2.3.2 Equilibrium

In equilibrium, aggregate demand for both vintage (q_{t-1}) and innovative goods (q_t) depends on the composition of the population. While young adults of period t purchase the latest quality (q_t), senior adults only demand the basic good as well as last-period's best quality of the innovative good (q_{t-1}). When the share of senior relative to young adults increases, aggregate demand shifts in favor of vintage goods. This is summarized in the following proposition.

Proposition 1. *Controlling for population size, if the population share of senior individuals increases, the demand for innovation declines.*

Proof. Total demand for the innovative quality q_t in period t is given by the number of young adults L_t^Y as every young individual consumes one unit of x_t at quality q_t. □

As a result of the change in demand, total R&D spending as well as the economy's rate of innovation decreases in an aging population.

Proposition 2. *If the senior-to-young worker ratio is higher, per capita total spending on R&D as well as the number of patent applications per capita decreases. Formally, if the $S2YWR_t$ is higher both z_t/L_t and $\Delta q_t/L_t$ are reduced.*

Proof. Total spending on R&D is given equation (2.9). Dividing both sides by L_t and replacing population size on the right-hand side by the $S2YWR$ from equation (2.2) shows that $\partial z_t^*/\partial S2YWR_t < 0$. □

Concerning the dynamics of the model, an increase in the fertility rate n_t of period t causes population aging in $t+2$. During this process, the rate of innovation decreases. While the model abstracts from several potentially important factors, it illustrates the main mechanism through which population aging can affect the rate of innovation. Extensions of the model could include international demand for innovative goods. However, in the empirical part, I already control for trade openness and find that it does not affect my findings. Another potential issue with the model is that it contains only one sector with R&D. In a multi-sector model, however, innovation could shift from one sector with declining demand to other sectors like health care in which demand might increase if the population becomes older. In the empirical analysis, I discuss this idea in more detail.

2.4 Empirical Evidence

This section puts the predictions of my theoretical model to the data. The idea of is to investigate whether the remarkable demographic shift in the composition of the labor force —documented in Chapter 2.2— is associated with economic outcomes of the kind suggested by the model. In particular, the theory suggests that with an aging population (or labor force), there should be a decline in the demand for innovative goods (Proposition 1) as well as a reduced number of patents (Proposition 2).

Concerning the first proposition, the lack of detailed data on product demand by age group obstructs a direct test. However, survey data from the Pew Research Center shows that, for example, smartphone ownership is highest among younger Americans. While 85% of 18–29 year olds use a smartphone, only about half of the 50–64 year olds do so. Similarly, data from Statistica indicates that more than half of iPhone users in the United States are aged 13-34 while only 21% are aged 45-64. Furthermore, among smartphone owners it is the younger age group which uses the most sophisticated services like online banking or turn-by-turn navigation. However, it remains to be shown that —beyond this anecdotal evidence— population aging generally affects a country's rate of innovation.

2.4.1 Data

For the empirical analysis, I use a novel data set that contains country-level statistics on both demographics as well as innovative activities. I compile the data set by drawing on three sources. First, for demographic information the UN Population Division provides data from 1950 to 2010 for a large set of countries. This not only includes fertility rates and age group shares for the past but also projections for the future based on different assumptions with respect to the fertility rate. Second, in order measure innovative activity I follow the recent literature (e.g. Aghion et al., 2014 and 2015) and use the number of patent applications per capita.[1] The OECD publishes information on such applications for each country and year. This data is publicly available for the years 1978–2012 and covers mostly OECD countries but has information for a selected group of non-OECD members as well. Field-specific information on patent applications (e.g., in ICT) is available for the same set of countries and the period 1999-2012. In the empirical analysis, I use this information to measure innovative activity in

[1] An alternative measure for a country's innovative activity is given by its R&D expenditures as a share of GDP. While such expenditures can be considered a proxy variable for R&D efforts, patent applications closer resemble the quality improvements that I discuss in the model.

each country. Finally, the Penn World Table (PWT, Mark 8.1) provides detailed data on economic indicators for 167 countries between 1950 and 2011. Throughout my analysis on innovative activities, I use data from 1978 to 2010 unless otherwise indicated. The focus is on OECD countries.[1] In order to illustrate the data set, Table 2.1 provides summary statistics on the main variables employed in the regressions.

— Table 2.1 about here —

The data set includes all 34 OECD countries for the period 1978-2010. Hence, the number of observations is 34×33 years $=$ 1,112. The number of patent applications, however, takes the value zero in some years and countries. Thus the growth rate in patent applications is missing in a few cases. The senior-to-young worker ratio (henceforth *S2YWR*) in this data set ranges from 0.31 to 1.21, reflecting the large variation across countries and time. Innovative activity is measured as the total number of patent applications per one million people which ranges from zero to 345.

2.4.2 Descriptive Evidence

The main prediction of the model is that the rate of innovation declines in an aging population. In order to examine this effect in the data, I first plot the time trend in the senior-to-young worker ratio for all OECD countries. In Figure 2.4, it is apparent that there was not much of a trend prior to the late 1990s. However, afterwards there was a steep increase.

— Figure 2.4 about here —

Similarly, there was not much of a trend in the average growth rate in patent applications per capita prior to 2000. In contrast, in the period between 2000 and

[1] As of 2015, there are twenty founding member countries and fourteen that joined later. The former group joined in 1961 and consists of Austria, Belgium, Canada, Denmark, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, United Kingdom, and the United States. Subsequently, the following countries joined: Japan (1964), Australia (1971), New Zealand (1973), Finland (1969), Mexico (1994), the Czech Republic (1995), Hungary (1996), Poland (1996) and Korea (1996).

2010, there was a notable decline in R&D-related activity. In the econometric analysis of Section 2.4.4, I will thus concentrate on this period.[1] This descriptive evidence suggests that population aging has rather a contemporaneous effect on innovation with little indication of anticipation.[2] In Figure 2.5, I plot the annual growth rate in patent applications per capita against the senior to young worker ratio. The graph is based on data from all current OECD countries and the time period 1978 to 2012.

— Figure 2.5 about here —

The growth rate of patent applications is significantly lower in those country-year observations with a larger senior-to-young worker ratio. As a notable finding, the plot suggests that there is a large variation in patent growth rates at S2YWR below one. Once the share of senior workers dominates, however, the mean growth rate in patent applications converges to zero. On average the growth rate drops from 0.26 to 0.01 when moving from a S2YWR of 0.75 or lower to a ratio above 1.10. A similar pattern is also found for most OECD countries when they are analyzed separately. Figure 2.6 provides two examples.

— Figure 2.6 about here —

In both the United States and Germany, the rate of innovative activity declined after 1978 along with a large increase in the *S2YWR*. Whether this observation just illustrates a correlation between two variables that follow a time trend or whether there is in fact a causal relationships requires further investigation.

2.4.3 Empirical Specification

In the first step of the econometric analysis, I run a regression of patents per capita on the senior-to-young worker ratio. This uses data from all 34 OECD countries and control variables for the GDP per capita, population size, trade

[1] Another benefit of focusing on the time after 1999 is that I have data on field-specific patent applications.

[2] The missing anticipation effect is discussed by DellaVigna and Pollet [2007]. To explain this phenomenon, the authors suggest a model of inattention to available information about the future.

openness and the share of 65+ year olds. Moreover, I control for the initial level of total patent applications per capita to take into account convergence effects [Barro and Sala-i Martin, 1992]. Finally, I add country- and year-fixed effects to the regressions. The baseline specification is given by

$$\mathrm{PAT}_{c,t} = \gamma_c + \delta_t + \tau \mathrm{S2YWR}_{c,t} + \mathbf{X}_{c,t}\beta + \varepsilon_{c,t} \tag{2.11}$$

where $\mathrm{PAT}_{c,t}$ is the growth in patent applications per capita, γ_c as well as δ_t are country- and year-fixed effects, $\mathbf{X}_{c,t}$ is a vector of control variables, and $\varepsilon_{c,t}$ denotes the time-varying country-specific idiosyncratic standard error which is clustered at the country level. The coefficient of interest is given by τ and indicates the impact of the senior-to-young worker ratio on the rate of innovation. Following the theoretical model of Section 2.3, I expect τ to be negative. This would indicate that population aging negatively affects R&D-related activities.

Identifying the causal effect of population aging on the rate of innovation remains a challenging task even though trends in demography can be considered exogenous. The primary source of concern is that some unobserved factor reduces R&D-related activities. The simultaneous trend in demography could then be spuriously correlated with trends in innovation. In order to mitigate this concern, I add country- and year-fixed effects to the regression. They pick up two disturbing factors. First, time-fixed country-specific factors which are not explicitly included on the right-hand side of equation (2.11). This is crucial as the literature points out the role of long-run, persistent determinants of innovation which differ across countries. These include, for example, political institutions, property rights, or cultural traits. All these factors vary across countries and affect the rate of innovation. However, in my empirical model, such determinants are absorbed by the time-invariant country-fixed effect. Second, the inclusion of year-fixed effects controls for common shocks such as the economic slowdown caused by the financial crisis in 2007-08. In addition, δ_t also reduces the impact of spurious time trends and panel error correlations. The downside of including fixed effects is that adding $N + T - 2$ dummy variables to the model creates a particularly demanding environment. It removes a large share of both the cross-country and within-country variation of the data. Hence, when reporting the results I show estimates of equation (2.11) with and without fixed effects.

2.4.4 Results

The results shown in Table 2.2 confirm the pattern found before and suggest again that a higher share of senior (45-64 year old) workers is associated with less growth in patent applications. This finding does not depend on whether control variables are added to the specification. Moreover, the addition of country- and year-fixed effects does not alter the negative coefficient on the *S2YWR* either.

— Table 2.2 about here —

Considering the magnitude of this effect, the estimates of Table 2.2 suggest that an increase of the senior-to-young worker ratio by 1% reduces the the growth rate of patents per capita by about 0.9%. In the last three columns of Table 2.2, patent data for specific sectors is used as the dependent variable. First, I consider the field of information and communications technology (ICT). According to my theoretical model, the negative impact of population aging is expected to be larger in this case. The estimates shown in Table 2.2 confirm this prediction. While the negative coefficient on $\text{S2YWR}_{c,t}$ remains highly significant its magnitude increases from about 0.9 to 1.2 when focusing on ICT patents. In contrast, no such significant negative effect is found among patents in medical or pharmaceutical technology. In line with research by Acemoglu and Linn [2004], research in an aging population might shift toward the development of new drugs and medical technology.

Supply versus Demand Effect — To this point, the empirical analysis has supported the hypothesis that the growth rate in patent applications per capita decreases in an aging population. This raises the question whether R&D-related efforts are reduced because of a supply or demand effect. The former would arise if innovative activity is mostly carried out by young workers. Hence their absence would lead to a decline in patent applications. The demand effect, in contrast, implies that due to the absence of young workers there is less demand for innovative goods. In the theoretical model of Section 2.3, the demand effect explains why the economy's rate of innovation is reduced if the share of senior workers increases. The data set allows to test this idea. In Column (5) of Table 2.2, the senior-to-young worker ratio is replaced by its two components: the population share of young (15–34 year old) and senior (45–64) workers.[1] The estimates suggest that both supply and demand factors play a role. The larger the fraction of young workers, the higher the rate of innovation. In contrast, an increase in the share of senior workers reduces patent growth.

In a second test, I examine whether there is a decline in the growth rate of patent applications *per young worker*. If this can be found in the data, it would serve as evidence of a demand effect. One problem with estimating this, however, is that any change in the number of patent applications per young or senior worker is driven largely by changes in the nominator. Demographic changes —even on a five-year basis— are minor compared to changes in R&D-related activity. Hence, it is not surprising to see that Figure 2.5 in the Appendix shows results that are very similar to the ones obtained when using patent applications *per capita* as in Figure 2.7.

2.4.5 Robustness Tests

The analysis so far indicates a negative impact of population aging on the rate of innovation. In order to explore the robustness of this findings, I conduct two sets of robustness checks.

Historical Demography — To further test whether the *S2YWR* has an effect on innovation, I suggest an instrumental variable (IV) approach using demographic information from earlier periods. In particular, I instrument, for example, the $S2YWR$ of the year 2000 by the ratio of 35–54 year olds to 5–24 year olds in 1990. The latter ratio can be referred to as medium-age worker to children ratio ($MAW2CR$). The validity of the IV approach rests upon two conditions,

[1]Note that using the S2YWR as an explanatory variable imposes the assumption that the share of young and the share of senior workers have opposite effects on innovation. Having both shares separately in the specification allows them to have independent effects.

namely that (i) demographic patterns within each country are not too distorted by immigration, and that (ii) historical population shares are uncorrelated with unobserved determinants of innovative activity. Empirically, I provide supportive evidence for the first condition.[1] Testing the second assumption, however, is not possible.

— Table 2.3 about here —

The use of $MAW2CR_{c,t-10}$ as an instrumental variable for $S2YWR_{c,t}$ is supported by large first stage F-tests shown in Table 2.3. The results with respect to how population aging affects innovative activity remain similar compared to the OLS estimation shown in Table 2.2. The point estimates are smaller but still highly significant. This provides further evidence that an aging population is associated with less innovation.

Time Trends — In a final test, I further explore whether changes in the senior-to-young worker ratio are not just correlated with R&D-related activities within each OECD country. If there is indeed a causal effect of workforce aging on patents, we should expect to see a sharper decline in the growth of patent applications in countries with larger increases in the *S2YWR*. Figure 2.7 supports this idea using data from 2000–2010.

— Figure 2.7 about here —

The graph plots the estimated time trends in the patent growth rate against the ten-year change in S2YWR. Countries like Korea and Greece that experienced the largest demographic shift between 2000 and 2010 are the ones that saw the largest decline in their patent growth rates over the same time period. The relationship shown in the figure can also be explored using regression analysis. Columns (4) and (5) of Table 2.3 provide the regression results. Notably the negative coefficient on the time trend in the *S2YWR* does not disappear when

[1] In the absence of migration, the $MAW2CR_{1990}$ would be identical with the $S2YWR_{2000}$. Immigration usually accounts for a tiny percent of the change in the composition of the workforce of a given age. Using data from the OECD countries, I find a correlation of about 0.9 between $S2YWR_t$ and $MAW2CR_{t-10}$.

controlling for time trends in the GDP per capita, population size, trade openness and the share of 65+ year olds. Despite the small number of observations (34 OECD countries), the coefficient on the change in *S2YWR* is highly significant and negative.

Extending the Time Period — The results of the econometric analysis suggest a negative impact of population aging on the rate of innovation. Motivated by findings of Figures 2.4 and 2.5, I use data for the most recent period 2000-2010 in the regressions. As is visible in the figures, there is not much of a trend in either the S2YWR or innovation prior to the year 2000. Hence, when extending the sample to the period of 1980–2010 or 1990–2010, I obtain similar estimates as shown in Table 2.2 with reduced significance.

2.5 Conclusion

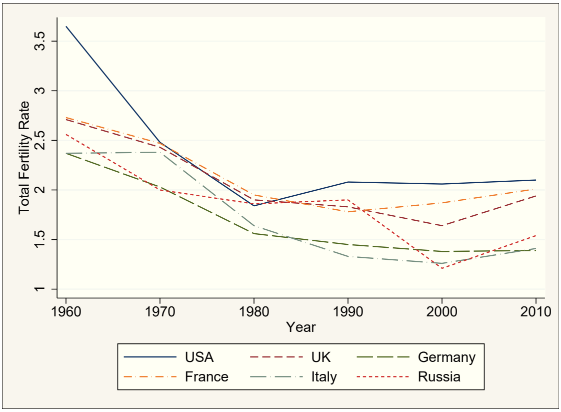
Given that basically all major countries experienced a substantial decline in their fertility rates, population aging will be a major trend in future decades. If present trends persist, by 2050 the median age in Europe will be significantly above fifty, up from about thirty-five today. In this paper, I explore a novel mechanism through which population aging affects innovation. Supported by the empirical analysis, my findings suggest that the rate of innovation is reduced in countries with aging workforces. Potentially the mechanism described in this paper also affects the rate of economic growth. This is because economic growth often occurs through improvements in quality as new models of consumer goods replace older ones [Bils, 2009].

It is important to emphasize that the mechanism described in this paper does not apply to all kinds of innovative activity. Following Acemoglu [2002*a*], the declining demand for innovative goods as in my model will reduce R&D with respect to such goods. However, innovative activity might shift to other sectors in an aging population. Most importantly, there might be more research in the field of health care due to the increased market size for such goods and services. Moreover, the shortage of young workers will affect prices and thus spur innovation with respect to labor-saving technologies.

Overall, my research aims at improving our understanding of the effects of an aging population on innovation. A misunderstanding of this link can have severe consequences if it leads to the imposition of misguided policies. These may arise, for example, from the neglect of slower economic growth in long-term projections which are relevant for retirement schemes [Poterba, 2014; Rojas, 2005]. In terms of policy conclusions, my research does not intend to provide any particular recommendations. Nevertheless, it is important to point out that the effects of population aging are not restricted to the impact on innovation as discussed in this paper. Research by Razin, Sadka and Swagel [2002*b*] finds that an increasing dependency ratio is correlated with lower taxes and less generous social transfers. However, several policies have been suggested to raise fertility rates. These comprise changes in social security, infrastructure investments, or improving education. While such policies been have been found to increase fertility in some studies [Bauernschuster, Hener and Rainer, 2016], there is no consensus [Bick, 2016]. In any case, they will not prevent population aging. Moreover, altering the fertility rate will not affect the composition of the workforce for at least two decades. Hence, if anything policy interventions must address the rate of innovation directly. Recent empirical evidence provided by Thomson [2015], however, suggests that the effectiveness of tax incentives for R&D remains obscure.

Figures and Tables

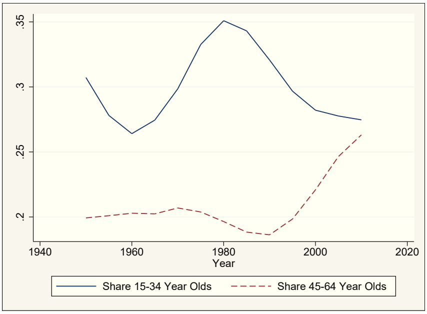
Figure 2.1: Historical Trends in Fertility Rates



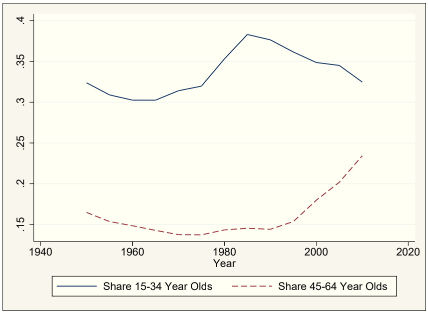
Note: The figure plots trends in fertility rates for six countries and the time period after 1960. Sources: Bundesinstitut für Bevölkerungsforschung and U.S. Census Bureau.

Figure 2.2: Young and Senior Workers, 1950–2010

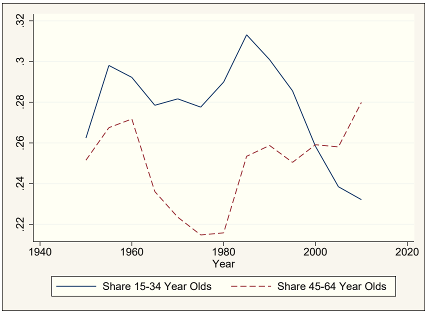
(a) United States



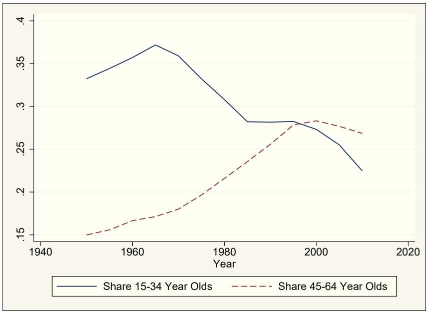
(b) China



(c) Germany



(d) Japan

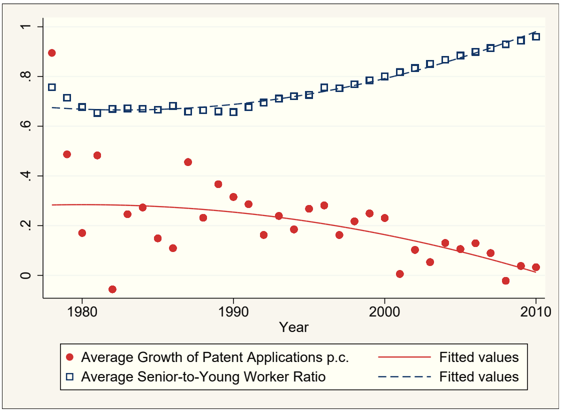


Note: The figures plot the share of young workers (15-34 years old) and senior workers (45-64 years old). The time period ranges from 1950 to 2010. Data obtained from the UN Population Division.

Figure 2.3: Senior-to-Young Worker Ratio, 1980 and 2010

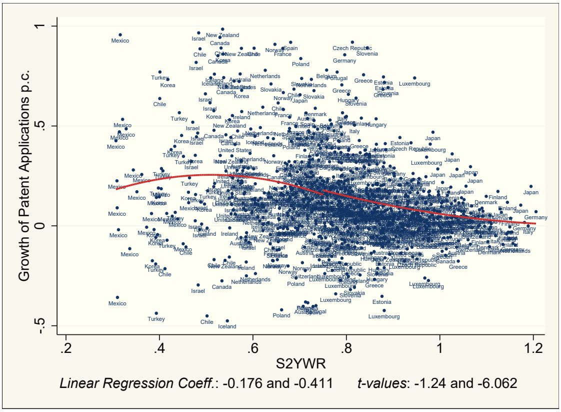
Note: The figure plots the kernel density of the senior-to-young worker ratio in all OECD countries. The solid blue line shows data from 1980 while the dashed red line is based on 2010. Data obtained from the UN Population Division.

Figure 2.4: Average Senior-to-Young Worker Ratio and Innovation



Note: The figure plots the average senior-to-young worker ratio (dashed blue line) as well as the average growth in patent applications per capita (solid red line) for the time period 1978–2010. The data includes all OECD countries. A quadratic fit is shown.

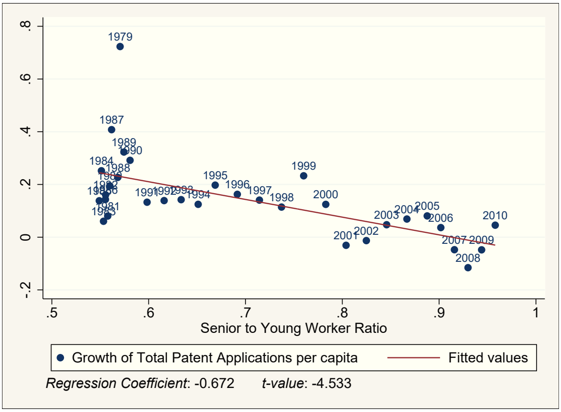
Figure 2.5: Senior-to-Young Worker Ratio and Innovation



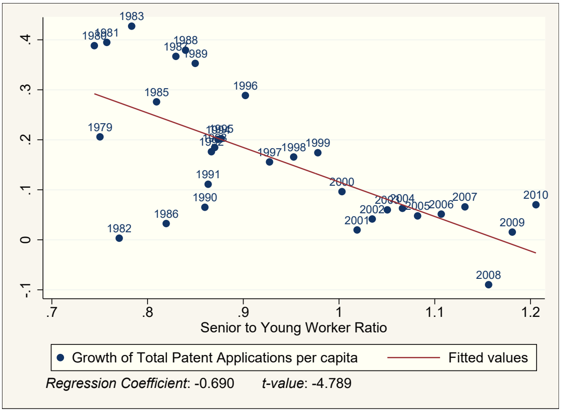
Note: The figure plots the growth in patent applications per capita against the senior-to-young worker ratio. Each dot represents a country-year observation. The data includes all OECD countries for the period 1978 to 2012. A quadratic fit is shown for S2YWR below and above 0.75. The coefficients and t-values for a split linear regression are indicated below the figure.

Figure 2.6: Senior-to-Young Ratio and Innovation

(a) United States

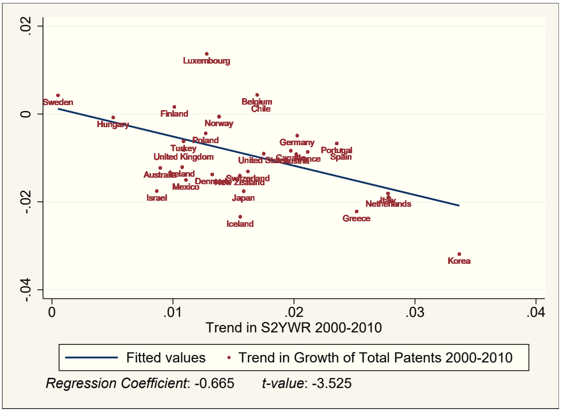


(b) Germany



Note: The figures plot the growth in patent applications per capita against the senior-to-young worker ratio. Plot (a) is based on data from the United States and Plot (b) from Germany for period 1978 to 2012. A linear fit is shown.

Figure 2.7: Ten-Year Trends in Patent Applications Per Capita



Note: The figure plots a country's ten-year trend in total patent growth against its ten-year trend in the senior-to-young worker ratio (S2YWR). The data includes all OECD countries for the period 2000 to 2010.

Table 2.1: Summary Statistics

Variable	**Mean**	**SD**	**Min**	**Max**	**N**
S2YWR	0.75	0.19	0.31	1.21	1,122
Total Patent Applications	49.21	71.77	0	345.51	1,122
Growth of Patent App. p.c.	0.51	5.32	-1	162.14	986
GDP p.c.	22,644.28	10,529.94	4,380.1	82,814.16	1,074
Population	32.51	50.82	0.22	312.25	1,122
Share 65+ Year Olds	12.9	3.53	3.83	22.96	1,054

Note: The table shows descriptive statistics for the data set used in the empirical analysis. Patent Applications are expressed per one million people in the population. GDP per capita is given in 2005 US Dollar. Population size is measured in million. Trade is defined as the sum of imports and exports divided by total GDP.

Table 2.2: Population Aging and Growth in Patent Applications

	Dep. Variable: Growth in Patent Applications							
			All Industries			ICT	MedTec	Pharma
Mean value			(0.08)			(0.12)	(0.22)	(0.19)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
S2YWR	-0.240***	-0.198*	-0.863***	-0.923***		-1.154**	-1.251	-1.128
	(0.088)	(0.112)	(0.270)	(0.332)		(0.520)	(1.567)	(1.121)
Share 65+		0.004	-0.045*	-0.035		-0.079*	0.113	0.030
		(0.007)	(0.023)	(0.026)		(0.046)	(0.120)	(0.083)
Share 15–34					4.081**			
					(1.523)			
Share 45–64					-2.813*			
					(1.584)			
Controls	-	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	-	-	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	-	-	-	Yes	Yes	Yes	Yes	Yes
Observations	374	374	374	374	374	373	368	372
R-squared	0.039	0.070	0.072	0.148	0.152	0.076	0.037	0.055

Note: The table shows the result of eight separate regressions using dependent variables as indicated in the top row. Control variables include the GDP per capita, population size, and trade openness. The sample includes all 34 OECD countries and the time period 2000-2010. Standard errors are clustered at the country level and shown in parentheses. Significance at the 10% level is indicated by *, at the 5% level by **, and at the 1% level by ***.

Table 2.3: Instrumental Variables and Trend Regressions

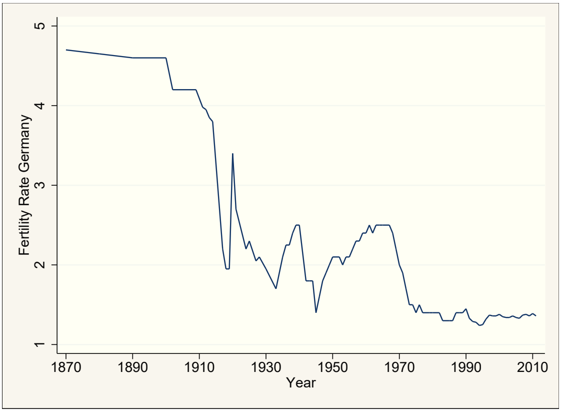
	IV Estimation			Trend Regression	
Mean of dep. var.	All Patents (0.08)		ICT Patents (0.12)	Patents 2000-2010 (0.011)	
	(1)	(2)	(3)	(4)	(5)
S2YWR	-0.613***	-0.501***	-0.621**		
	(0.103)	(0.159)	(0.288)		
Share 65+		-0.016	-0.035		
		(0.017)	(0.024)		
Trend in S2YWR				-0.954***	-0.762***
				(0.220)	(0.215)
Share 65+ in 2000				0.001	
				(0.001)	
Trend in Share 65+					-0.004
					(0.014)
Control Variables	-	Yes	Yes	2000	Trends
Country FE	-	Yes	Yes	-	-
Year FE	-	Yes	Yes	-	-
1st Stage F-stat	213.8	87.9	87.4	-	-
Observations	374	374	373	34	34
R-squared	0.039	0.031	0.015	0.332	0.345

Note: The table shows the results of three separate IV regressions in columns (1)–(3) with the SY2WR being instrumented by the lag-10 of the MAW2CR (ratio of 35–54 to 5–24 year olds). In columns (4) and (5), the time trend in patent applications is regressed on the time trend in the S2YWR. Control variables include the GDP per capita, population size, and trade openness. Standard errors shown in parentheses are Jacknife in columns (1)–(3) and Huber-White in (4) and (5). Significance at the 10% level is indicated by *, at the 5% level by **, and at the 1% level by ***.

Appendix

Additional Tables and Figures

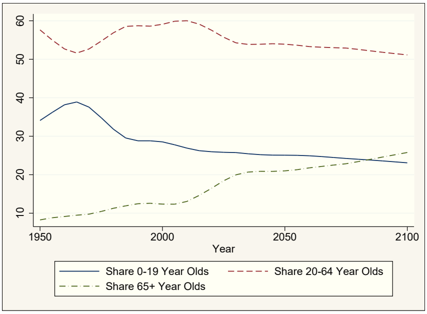
Figure 2.1: Historical Fertility Rates in Germany



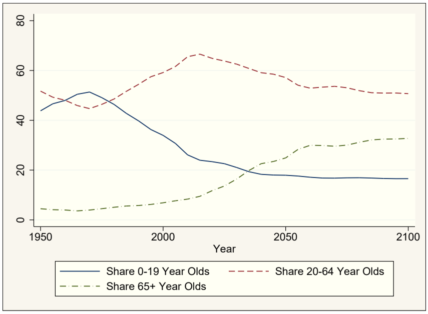
Note: The figure plots the fertility rate in Germany since 1870. Data obtained from the Bundesinstitut für Bevölkerungsforschung.

Figure 2.2: Age Group Shares in Selected Countries, 1950–2100

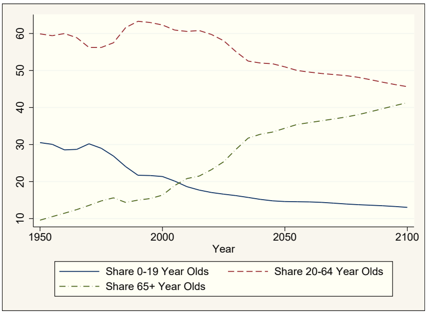
(a) United States



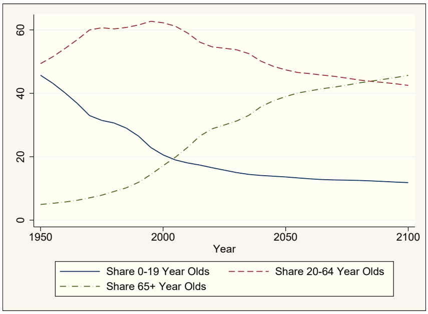
(b) China



(c) Germany

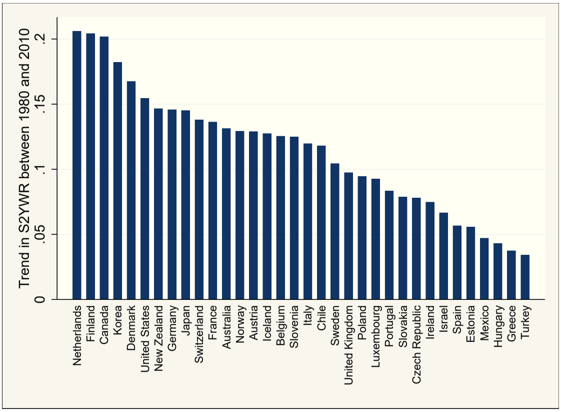


(d) Japan



Note: The figures plot the share of children (0-19 year old), working-age adults (20-64), and retirees (65 plus). Estimates for the years after 2010 are based on a constant fertility rate (CFR). Data obtained from the UN Population Division.

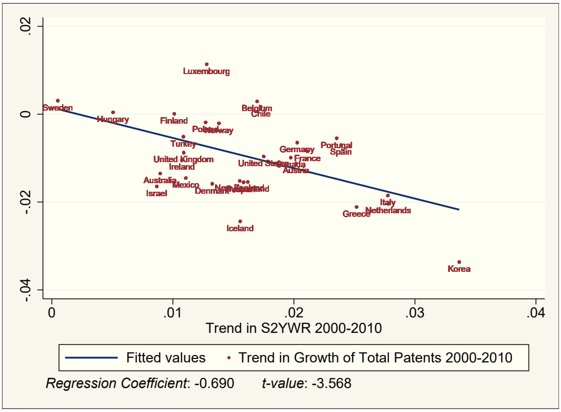
Figure 2.3: Time Trends in the S2YWR in OECD Countries



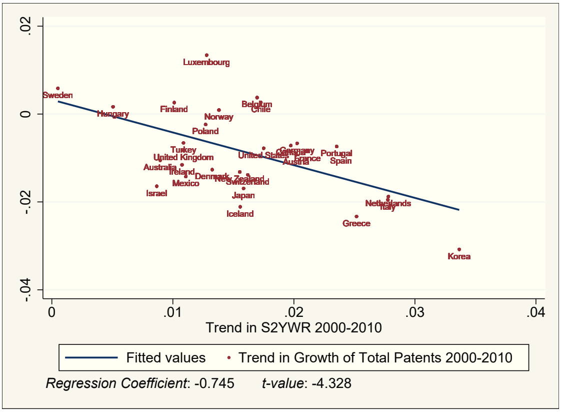
Note: The figure plots the estimated coefficients from regressing the senior-to-young worker ratio (S2YWR) on a year variable. The coefficients are estimated for each country separately and multiplied by ten. The time period is from 1980–2010.

Figure 2.5: Trends in Patents Per Young and Senior Worker

(a) Per Young Worker



(b) Per Senior Worker



Note: The figures plot a country's ten-year trend in total patent growth against its ten-year trend in the senior-to-young worker ratio (S2YWR). Patent growth is measured per young worker (aged 15-34) in Panel (a) and per senior worker (45-64) in Panel (b). The data includes all OECD countries for the period 2000 to 2010.

Table 2.1: Trends in Fertility Rates

Country	Total Fertility Rate					
	1960	1970	1980	1990	2000	2010
Australia	3.45	2.86	1.89	1.90	1.76	1.95
Austria	2.69	2.29	1.65	1.46	1.36	1.44
Belgium	2.54	2.25	1.68	1.62	1.64	1.84
Canada	3.90	2.33	1.68	1.71	1.49	1.63
Chile	5.58*	3.95	2.72	2.59	2.05	1.89
Czech Republic	2.11	1.91	2.10	1.89	1.14	1.49
Denmark	2.54	1.95	1.55	1.67	1.77	1.87
Estonia	1.98*	2.17*	2.02	2.05	1.36	1.72
Finland	2.71	1.83	1.63	1.79	1.73	1.87
France	2.74	2.48	1.95	1.78	1.87	2.02
Germany	2.37	2.03	1.56	1.45	1.38	1.39
Greece	2.23	2.40	2.23	1.40	1.27	1.47
Hungary	2.02	1.97	1.92	1.84	1.33	1.26
Iceland	4.26	2.81	2.48	2.31	2.08	2.20
Ireland	3.76	3.87	3.23	2.12	1.90	2.06
Israel	3.87*	3.78*	3.14	3.02	2.95	3.03
Italy	2.41	2.42	1.68	1.36	1.26	1.41
Japan	2.00	2.13	1.75	1.54	1.36	1.39
Korea	6.00	4.53	2.82	1.57	1.47	1.23
Luxembourg	2.28	1.98	1.50	1.62	1.78	1.63
Mexico	6.78	6.72	4.71	3.36	2.65	2.28
Netherlands	3.12	2.57	1.60	1.62	1.72	1.80
New Zealand	4.24	3.17	2.03	2.18	1.98	2.17
Norway	2.91	2.50	1.72	1.93	1.85	1.95
Poland	2.98	2.20	2.28	1.99	1.37	1.38
Portugal	3.10	2.83	2.18	1.56	1.56	1.39
Slovakia	3.07	2.40	2.31	2.09	1.29	1.40
Slovenia	2.18	2.21	2.11	1.46	1.26	1.57
Spain	2.86	2.90	2.22	1.36	1.23	1.37
Sweden	2.20	1.94	1.68	2.14	1.55	1.98
Switzerland	2.44	2.10	1.55	1.59	1.50	1.54
Turkey	6.40	5.00	4.63	3.07	2.27	2.06
United Kingdom	2.72	2.43	1.90	1.83	1.64	1.92
United States	3.65	2.48	1.84	2.08	2.06	1.93

Note: The table shows total fertility rates for all OECD countries for the period 1960–2010. The total fertility rate in a specific year is defined as the total number of children that would be born to each woman if she were to live to the end of her child-bearing years and give birth to children in alignment with the prevailing age-specific fertility rates. It is calculated by totaling the age-specific fertility rates as defined over five-year intervals. Assuming no net migration and unchanged mortality, a total fertility rate of 2.1 children per woman ensures a broadly stable population. The data is taken from OECD (2015): Fertility rates (indicator).

Chapter 3

Innovation and Trade in the Presence of Credit Constraints

This chapter is based on joint work with Reto Föllmi and Alexa Tiemann from the University of St.Gallen.

3.1 Introduction

A large body of trade literature provides empirical support for the superior performance characteristics of exporting firms relative to non-exporters (e.g., Bernard and Jensen, 1999; Bernard et al., 2007). One particular explanation for this superiority points to a complementarity between firms' exporting status and their investments in productivity-enhancing activities. Investigating this link follows a key insight by Schmookler [1966] that inventors channel their efforts into those lines of activity with high prospective profits. If the size of any market that firms operate in becomes larger, potential profits increase and so do firms' R&D-related efforts. This can be linked to trade policy in the sense that any removal of trade frictions effectively enlarges the size of the market firms are serving. Operating in a globalized market makes it more profitable for firms to invest in R&D as a means to remain competitive and capture a large share of the international market.

We examine both theoretically and empirically the decision of firms, who operate under financial constraints, to invest in R&D when tariff rates change. We believe this setup to be of particular interest as some firms may not capture the increased market size due to trade liberalization if credit constraints restrict their ability to adjust the production. Previous research by Banerjee [2004] as well as Banerjee and Duflo [2014] documents the importance of credit constraints for small and medium-sized enterprises in developing countries. Aghion et al. [2012] and Gorodnichenko and Schnitzer [2013] show theoretically that financial frictions affect a firm's decision to invest in innovative activities in a closed economy setup. In their empirical contribution, Aghion et al. (2012) find that French firms' R&D investments turn procyclical in the presence of binding credit constraints. We augment this argument and illustrate how trade liberalization can worsen access to finance for small firms, thereby affecting innovative activities. Poor entrepreneurs in our model are particularly susceptible to trade openness because reducing tariff protection has heterogeneous effects across firms. While unconstrained enterprises increase their innovative activities, credit-rationed firms —which are in general small or medium-sized— either cut spending on R&D or leave the market altogether.

Our research adds to several recently published studies documenting the so-called market size effect on innovation with respect to trade liberalization. Bernard, Jensen and Schott [2006] show that U.S. manufacturing industries exhibit strong productivity growth after experiencing large declines in trade costs. Atkeson and Burstein [2010] provide a general equilibrium model and show how trade liberalization affects firms' decision to invest in R&D. For a sample of Argentinean firms, Bustos [2011*a*,*b*] documents that tariff reductions were associated with technology upgrading. Aw, Roberts and Winston [2007] as well as Aw, Roberts and Xu [2011] also consider research investments and exporting behavior as joint decisions. Using plant-level data from Taiwan they find evidence that both activities increase productivity.[1] Finally, Lileeva and Trefler [2010] draw on Canadian firm-level data and provide evidence that after a reduction in tariff rates, firms that started exporting or exported more increased their productivity and became more innovative.

This literature, however, has been largely silent about firm-level responses to trade liberalization in the presence of imperfect capital markets. Traditional trade theory assumes that resources are allocated efficiently across firms by means of a well-functioning capital market. While this takes into account that firms face upfront expenses for production and R&D efforts, it abstracts from capital market imperfections found in most developing economies.[2] Recent advances in trade theory incorporate credit constraints. Foley and Manova [2015] provide a survey of this literature. In particular, lacking access to capital can preclude some firms from engaging in trade although this could be profitable [Manova, 2013; Manova, Wei and Zhang, 2015]. This adverse effect of financial frictions is particularly severe in sectors with high dependency on external finance [Rajan and Zingales, 1998].

In our theoretical model we follow previous work by Föllmi and Oechslin [2010, 2014] but we allow for quality-upgrading. In particular, we assume that entrepreneurs differ in their initial wealth endowments. As a result, poor firm owners must rely on external finance for their business activities. In doing so, they use their local monopoly power as a collateral. However, after trade liberalization the decline in mark-ups as well as the rise in interest rates limit their ability to borrow. While trade liberalization would induce them to adjust their production, investments are restricted by financial constraints. This gives rise to two results. First, quality-upgrading is limited in the sense that we observe fewer investments in R&D among financially constrained firms exposed to tariff reductions. Second, severely credit-constrained firms are likely to leave the market after

[1] Verhoogen [2008] finds similar evidence in a panel data set on Mexican manufacturing plants. In a related study, Costantini and Melitz [2008] build a dynamic model of plant-level adjustments to trade liberalization capturing the joint innovation and exporting decision, selection into exporting, and learning-by-exporting.

[2] Aghion et al. [2010] analyze the implications of financial frictions with respect to volatility and growth. They provide theoretical and empirical support for adverse effects of credit constraints on long-term investments.

trade liberalization.

In addition to our theoretical work, we examine empirically the responses of firms to trade liberalization conditional on a number of firm and market characteristics that may explain heterogeneous behavior. We focus on the combined effect of financial constraints and tariff reductions on market exit as well as several R&D-related activities. The latter are measured by product and process innovations as well as filed patents over a three year period. For our empirical analysis we use a novel data set on seven Latin American countries for the years 2006 and 2010. We merge firm survey data from the World Bank's Enterprise Surveys with tariff data from the World Integrated Trade Solution database. The former contains a large set of variables on firm characteristics while the latter allows us to apply a precise measure of treatment for each firm on the four-digit ISIC classification level. Our data set is preferable to those used in many other studies as it contains direct measures of innovation and financial constraints. Hence we do not have to rely on proxies for outcome and treatment variables in the econometric specification.

Although tariff reductions in our data have a positive impact on productivity-enhancing activities, we find evidence showing that credit constraints partly drive firms' responses to liberalization. The results indicate that tariff cuts worsen small and medium-sized firms' access to finance. Furthermore, financially constrained firms experience substantial declines in annual sales if they were subject to tariff reductions. We also find that market exit is more pronounced among financially constrained firms. Our estimates suggest that a tariff reduction for these firms is associated with a significant increase in the probability of leaving the market. Moreover, we find that among surviving firms, those reporting financial constraints in the initial period are associated with a lower probability of introducing innovative products or production processes if they are subject to trade liberalization. These findings are in line with our theory and shown to be robust to the inclusion of various control variables as well as country- and industry-fixed effects. Moreover, the results are not driven by any single country or by an underlying correlation of firm size and credit constraints. Furthermore, the results are similar when using simple average or weighted average tariff rates and the impact is generally magnified in less developed countries.

Our work adds another dimension to the literature on how financial constraints distort reallocations within firms after trade liberalization. In this we contribute to recent studies by Hsieh and Klenow [2009] as well as Song, Storesletten and Zilibotti [2011] who examine how credit market frictions lead to resource misallocation in low-income countries. At the macro level, Caselli [2012, 2013] shows that among developing countries gains from trade openness depend inversely on the degree of wealth inequality prior to liberalization. Closely related to our work, Chesnokova [2007] presents a model with necessary investments and credit constraints. In this model, specialization after trade liberalization can be welfare-reducing if specialization in agriculture affects the wealth distribution such that credit constraints become more binding. Amiti and Weinstein [2011] examine firm-specific shocks to trade finance supply in the setting of Japan's systemic crises from 1990 through 2010. Their findings suggest that liquidity shocks hurt firms' export growth even more than domestic sales.

We also contribute to previous research on international trade and product choice. Acemoglu and Zilibotti [2001] document that firms in poor economies generally tend to produce less innovative goods. Hence these firms typically adjust their production after liberalization [Fan, Li and Yeaple, 2015; Fieler, Eslava and Xu, 2014]. In particular, exporting firms in low-income countries produce higher-quality goods for export than for the domestic market [Verhoogen, 2008]. However, producing superior quality requires R&D investments. Hence, innovative firms may be exposed to higher survival risks if they do not retain diversified sources of finance [Fernandes and Paunov, 2015]. Our paper adds to these studies, suggesting that imperfect capital markets can be a source of market exit and limited product upgrading after trade liberalization.

This finding can be linked to research on the determinants of gains from trade. In the past decade, several studies revived the idea that countries specialize according to their comparative advantages and benefit from trade because they have access to different technologies [Eaton and Kortum, 2012]. Simultaneously, trade theory has been enriched by the idea of heterogeneous firms within countries. Following the seminal work by Melitz [2003], much research has analyzed the differential effects of trade liberalization across firms and workers. Moreover, there is an ongoing discussion about the magnitude of welfare gains from trade. While Eaton and Kortum [2002] as well as Arkolakis, Costinot and Rodríguez-Clare [2012] find only modest gains, Caliendo and Rossi-Hansberg [2012] as well as Melitz and Redding [2014] argue that trade openness can induce a reorganization of production which raises domestic productivity and may cause welfare gains from trade to become arbitrarily large. Average welfare effects, however, shroud large heterogeneity in welfare gains across countries, in particular the mixed experiences of developing countries [Galor and Mountford, 2008; Greenaway, Morgan and Wright, 2002]. In this regard, our work addresses the question under which circumstances heterogeneous responses and financial constraints at the firm level can limit overall gains from trade.

More broadly, our study is related to previous research on trade liberalization and income inequality [Goldberg and Pavcnik, 2007] as well as the dispersion of mark-ups across firms [Epifani and Gancia, 2011]. In our model, entrepreneurs differ in their initial capital endowments. Due to capital market frictions, the reduction of tariff rates has an adverse effect on poor entrepreneurs. In contrast to their better-endowed peers, not only can they not benefit from new export opportunities but they also suffer from rising borrowing costs.

The different strands of the literature lead to our hypothesis that access to finance plays a key role in determining firm-level responses to trade liberalization. We investigate this relationship by considering two observable firm decisions: market exit and productivity-enhancing activities. Our theoretical model predicts less product-upgrading and increased market exit among credit-constrained firms that are subject to trade liberalization. The empirical results in our paper support these predictions. Adjustments after liberalization appear to be impaired at the firm level if access to finance is limited. Both theoretically and empirically these findings add to previous work by Peters and Schnitzer [2015]. In addition we contribute to research following Aghion, Caroli and García-Peñalosa [1999] who show that in the presence of limited borrowing capacities, the distribution of wealth affects firms' production possibilities.

The remainder of the paper is organized as follows. In Section 3.2 we present our theoretical model to illustrate the impact of capital market frictions in a setting of heterogeneous firms facing trade liberalization. Section 3.3 describes our data set and provides descriptive statistics. Our empirical strategy as well as the results are shown in Section 3.4. The final Section 3.5 concludes and discusses policy implications.

3.2 Theory

3.2.1 The Setup

The model, in particular the design of the credit market, follows Föllmi and Oechslin (2010). We consider a static economy, populated by a continuum of (potential) entrepreneurs with population size 1. The individuals are heterogeneous with respect to their initial capital endowment $\omega_i, i \in [0,1]$. The capital endowments are distributed according to the distribution function $H(\omega)$. Aggregate capital endowment, $\int_0^\infty \omega dH(\omega)$, is denoted by K.

Each individual owns a specific skill (a "business idea") that makes him a monopoly supplier of a single differentiated good. All goods are produced with a simple technology that requires physical capital as the only input into production. Following the trade literature (e.g., Melitz [2003]), starting production needs a fixed outlay of f capital units. Formally, the production function reads $y_i = a(k_i - f)$ where a is a productivity parameter and y_i and k_i denote, respectively, output and capital invested. In addition to the initial business idea, the entrepreneur has the option to invest in R&D which raises the quality of the product from level 1 to $q > 1$, to model it in the most simple way.[1] Investment into quality upgrading requires additional $f(q-1)$ capital units.

The individuals' utility function takes the familiar CES-form and consumers

[1]Empirical support for the effect of trade liberalization on product quality upgrading is provided by Fernandes and Paunov [2013].

treat the two different quality versions of a good j as perfect substitutes

$$U = \left[\int_0^1 (c_{1j} + qc_{qj})^{(\sigma-1)/\sigma} \, dj \right]^{\frac{\sigma}{\sigma-1}}, \ \sigma > 1, \tag{3.1}$$

where $c_{\cdot j}$ denotes consumption of good j at quality level 1 or q. Each individual i maximizes the objective function (3.1) subject to the budget constraint

$$\int_0^1 (p_{1j}c_{1j} + p_{qj}c_{qj}) \, dj = m(\omega_i), \tag{3.2}$$

where $p_j(q_j)$ is the price of good j (for quality 1 or q) and $m(\omega_i)$ refers to individual i's nominal income depending on the initial capital endowment, ω_i.

Since both quality levels have the same marginal costs in production, we may simplify the exposition assuming that $p_{1j} \leq p_{qj}/q$. Under these conditions and if only one quality per good j is consumed, individual i's demand for good j reads

$$c_j(m(\omega_i), p_j(q_j), q_j) = q_j^{-1} \left(\frac{p_j(q_j)/q_j}{P} \right)^{-\sigma} \frac{m(\omega_i)}{P}, \tag{3.3}$$

where $P \equiv [\int_0^1 p_{1j}^{1-\sigma} dj]^{1/(1-\sigma)}$ is the familiar CES price index.

Individuals take the equilibrium borrowing rate as given but there may be credit-rationing. The reason for an upper bound on borrowing is the imperfect enforcement of credit contracts. Following Foellmi and Oechslin (2013), we assume that – in case of default – borrower i loses only a fraction $\lambda \in (0, 1]$ of the current firm revenue, $p(y_i, q_i)y_i$. Hence the parameter λ indicates how well credit contracts can be enforced.[1] If λ is close to zero, the borrowers do not lose much when they do not honor their debt. In that case, the incentives for lenders are small to provide high levels of external finance.

The lender will give credit only up to the point where the borrower still has an incentive to pay back. Formally, the size of the credit cannot exceed $\lambda p(y_i, q_i)y_i/r$,

[1] Alternatively, we can assume that the lender can only recover a fraction λ of current profits in case the entrepreneur defaults.

where r denotes the interest rate. As there is no default in equilibrium, the borrowing rate r must be the same for all agents. To calculate the amount of credit needed, note that you need $k_i = y_i/a + f$ capital units to produce y_i. To produce at quality level q_i additional $f(q_i - 1)$ capital units are needed. With equity ω, you need to borrow $y_i/a + f + f(q_i - 1) - \omega$ capital units. Taking that into account, borrower i will repay the debt if

$$\lambda p(y_i, q_i) y_i / r \geq y_i/a + f q_i - \omega_i. \tag{3.4}$$

3.2.2 Effects of International Trade

We assume that the home economy is a developing country (the "South"). The trading partner, the rest of the world, is an advanced economy and referred to as the "North". Trade costs take the usual "iceberg" formulation and we assume that $\tau \geq 1$ units of a good have to be shipped in order for one unit to arrive at the destination. As in Föllmi and Oechslin [2014] we assume that the North differs from the South in that its markets function perfectly. In particular, the northern credit market is frictionless so that there are no credit constraints. Moreover, in the North, each variety in both qualities is produced by a large number of firms so that the northern goods market is perfectly competitive. Regarding access to technology and preferences, there are no differences between the two regions. Further, for the sake of simplicity, the North produces the same spectrum of goods as the South does.

These assumptions imply that all northern firms charge a uniform price for a given quality, equal to the marginal cost. We normalize the northern price level for products of quality q to one. This normalization implies that all goods prices in the North (as well as the northern marginal cost) are also equal to one.

What does this mean for the market structure in the South? Although entrepreneur i has a domestic monopoly, he faces a competitive fringe by Northern producers and cannot set a price above τ when supplying the high quality, and τ/q when supplying the low quality. We assume that the market is sufficiently integrated such that all entrepreneurs face the competitive fringe (for cases with intermediate values of τ, see Föllmi and Oechslin [2014]).

We are left to determine the borrowing rate. Since we are looking at an equilibrium in which a positive mass of entrepreneurs is credit-constrained and cannot serve the whole market, the economy imports goods from abroad. This, in turn, implies that there must be positive aggregate exports with balanced trade in a static model. The marginal product of capital equals a. There if an entrepreneur exports one unit of an arbitrary good, this needs τ/a units of capital and generates an income of 1. The entrepreneur compares the return from exporting $(\tau/a)^{-1}$ with the returns when acting as lender on the domestic market. Arbitrage requires therefore that the domestic borrowing rate r must equal a/τ.

For credit-constrained firms, the maximum output $\overline{y}$ is determined by $\lambda p(\overline{y}, q_j)\overline{y}/r \geq \overline{y}/a + fq_j - \omega$, where we use (3.4) and $p(\overline{y}, q_j) = \tau q_j/q$. Note that the price the firm can charge is given by $p(\overline{y}, q_j) = \tau q_j/q$. That is, if the firm invests into quality upgrade it can charge a price of τ, otherwise the price is τ/q. We get $\overline{y} = a(\omega - fq_j)/(1 - \lambda\tau^2 q_j/q)$. Firms not facing the credit constraint serve the whole market. Using (3.3) and taking into account that high-quality industry output is given by $y_{\max} = q_j^{-1}(\tau/q)^{-\sigma} P^{\sigma-1}Y$, where $P^{\sigma-1}Y$ is uniquely determined in the macroeconomic equilibrium (see Foellmi and Oechslin, 2013). To sum up, domestic output in sector j is given by

$$y_j = \min\left\{\frac{a(\omega - fq_j)}{1 - \lambda\tau^2 q_j/q}, q_j^{-1}(q/\tau)^{\sigma} P^{\sigma-1}Y\right\} \tag{3.5}$$

Note that firm output increases in initial wealth for the credit-constrained, poorer entrepreneurs. The reason is the credit market imperfection: an increase in ω means an entrepreneur has more resources to invest and – in addition – it allows for higher borrowing since the entrepreneur has more own collateral which he would lose by not honoring the credit contract. In that sense, $\left(1 - \lambda\tau^2/q_j\right)^{-1}$ may be interpreted as credit multiplier. Note that the credit multiplier falls and the firm size of constrained entrepreneurs necessarily falls if τ decreases. This is due to two effects: First, a decrease in trade costs lowers the maximum price monopolists can charge which erodes profits serving as collateral. Second, a lower τ increases the borrowing rate $r = a/\tau$, since exporting is more attractive. Higher borrowing rates make it more difficult that equation (3.4) holds.

3.2.3 Decision on Exit and R&D

An entrepreneur seeks to maximize his nominal income. It is given by revenues minus interest payments or $p(y_i, q_i)y_i - r\,(y_i/a + fq_i - \omega_i) = (1-\lambda)p(y_i, q_i)y_i$, for active credit-constrained entrepreneurs (using (3.4)). The entrepreneur compares the entrepreneurial income with and without quality upgrading, m_{eq} and m_e, respectively, and the income he would get if he decides to exit and become a lender, earning $m_l \equiv r\omega = a\omega/\tau$. Thus, he maximizes nominal income

$$\max_{\{e,q\}} \{m_{eq}(\omega), m_e(\omega), m_l(\omega)\}$$

where

$$\begin{aligned} m_{eq}(\omega) &= (1-\lambda)\tau\frac{a\,(\omega - fq)}{1-\lambda\tau^2} \\ m_e(\omega) &= (1-\lambda)\tau\frac{a\,(\omega - f)}{q-\lambda\tau^2} \\ m_l(\omega) &= \frac{a\omega}{\tau}. \end{aligned}$$

To have an interesting problem where all three occupations (l, e, eq) are possible outcomes, we make the following assumption on trade costs, which is necessary and sufficient such that occupation e exists.

Assumption 1. *We assume the following condition to hold:* $\tau^2 > 1 + q$

Given Assumption 1 we see directly that $m'_{eq}(\omega) > m'_e(\omega) > m'_l(\omega)$. On the other hand $m_{eq}(0) < m_e(0) < m_l(0) = 0$. Hence, the poorest agents will choose to become lenders, for medium levels of ω agents become entrepreneurs without investing into quality upgrading, and for high levels of ω the entrepreneurs invest in R&D as well. The critical wealth level ω_1 where agents are indifferent between becoming entrepreneur or lender equals $\omega_1 = f(1-\lambda)\tau^2/(\tau^2 - q)$. The critical wealth level ω_2 where agents are indifferent between investing into quality or not investing equals $\omega_2 = f\left(1 + q - \lambda\tau^2\right)$. Obviously, occupation e exists only iff $\omega_2 > \omega_1$. It is easy to check (by insertion) that $\omega_2 > \omega_1$ holds iff $\tau^2 > 1 + q$.

Intuitively, the product market imperfections make entrepreneurship more profitable than being lender. As entrepreneurship entails fixed costs, this option is only preferred to being lender if the firm size is large enough. A fortiori this argument holds for investment in R&D. Since firm size and wealth are positively correlated with each other, poorer entrepreneurs are more likely to become lenders and are less likely to invest into high-quality production.

– Figure 3.1 about here –

Individuals with an initial endowment below ω_1 decide to be lenders while those with a larger endowment become entrepreneurs. If the initial endowment is larger than ω_2, individuals become entrepreneurs and invest in quality upgrading.

The following two propositions state that financially constrained firms are more likely to exit the market and less likely to invest in R&D when trade liberalization occurs.

Proposition 1. *A decrease in trade costs τ induces severely credit-constrained firms to exit the market.*

Proof. The derivative of ω_1 with respect to τ reads $\partial\omega_1/\partial\tau = -2fq(1-\lambda)\tau/(\tau^2-q)^2 < 0$. A decrease in τ increases the range of entrepreneurs who choose to become lenders. $\square$

Proposition 2. *A decrease in trade costs τ reduces investment into quality upgrading by financially constrained firms.*

Proof. The minimum wealth level necessary to invest in quality, $\omega_2 = f\left(1+q-\lambda\tau^2\right)$, decreases in τ. A lower level of ω_2 reduces the range of credit-constrained entrepreneurs who invest in R&D. $\square$

Things look different for unconstrained entrepreneurs. Trade liberalization raises, ceteris paribus, the incentives to invest into quality upgrading. The reason is that for unconstrained entrepreneurs the high-quality output increases when trade costs τ fall, as market demand is higher with lower prices and eventually higher real income due to lower price distortions. A financially constrained entrepreneur, instead, is incapable to serve the full market because of limited access to credit. This market-size effect makes the option to invest in R&D more attractive for the unconstrained entrepreneur. The income of a financially unconstrained entrepreneur, producing high-quality products, is given by

$$(q/\tau)^{\sigma-1} P^{\sigma-1}Y - (a/\tau)\left(q^{-1}(q/\tau)^{\sigma} P^{\sigma-1}Y/a + fq - \omega\right).$$

Income when producing low-quality is given by

$$(q/\tau)^{\sigma-1} P^{\sigma-1}Y - (a/\tau)\left((q/\tau)^{\sigma} P^{\sigma-1}Y/a + f - \omega\right).$$

The difference between the two expressions equals $(a/\tau)\left(q^{-1}(q/\tau)^{\sigma} P^{\sigma-1}Y/a - f\right)(q-1)$. Hence, the incentive to invest in high quality rises when $\tau^{-\sigma}P^{\sigma-1}Y$ is larger.

Gross capital supply equals demand in the capital market equilibrium condition,

$$\begin{aligned} K = {} & [1 - G(\omega_3)]\left[q^{-1}(q/\tau)^{\sigma} P^{\sigma-1}Y/a + fq\right] \\ & + \int_{\omega_2}^{\omega_3}\left[\frac{\omega - fq}{1-\lambda\tau^2} + fq\right] dG(\omega) + \int_{\omega_1}^{\omega_2}\left[\frac{\omega - f}{1-\lambda\tau^2/q} + f\right] dG(\omega), \end{aligned}$$

where $\omega_3 = fq + \left(1-\lambda\tau^2\right) q^{-1}(q/\tau)^{\sigma} P^{\sigma-1}Y/a$ denotes the wealth level such that the entire market demand can be served. When τ falls, the gross capital demand of entrepreneurs falls and more entrepreneurs become lenders. Hence, whenever some entrepreneurs are credit-constrained, $\tau^{-\sigma}P^{\sigma-1}Y$ must rise such that the capital market equilibrium condition holds. (It stays constant if all entrepreneurs are unconstrained.) The key difference to the constrained entrepreneurs is that the output $y_{\max}(q) = q^{-1}(q/\tau)^{\sigma} P^{\sigma-1}Y$ increases when trade costs τ fall, as market demand is higher with lower prices and eventually higher real income. A financially constrained entrepreneur, instead, is incapable to serve the full market because of limited access to credit. The rise in firm output makes the option to invest in R&D more attractive for the unconstrained entrepreneur.

Proposition 3. *For firms not facing financial constraints, a decrease in trade costs τ increases the probability to invest in R&D.*

Proof. The income of a financially unconstrained entrepreneur, producing high-quality products, is given by

$$(q/\tau)^{\sigma-1} P^{\sigma-1}Y - (a/\tau)\left(q^{-1}(q/\tau)^{\sigma} P^{\sigma-1}Y/a + fq - \omega\right).$$

Income when producing low-quality is given by

$$(q/\tau)^{\sigma-1} P^{\sigma-1}Y - (a/\tau)\left((q/\tau)^{\sigma} P^{\sigma-1}Y/a + f - \omega\right).$$

The difference between the two expressions equals
$(a/\tau)\left(q^{-1}(q/\tau)^{\sigma} P^{\sigma-1}Y/a - f\right)(q-1)$. The incentive to produce the high quality increases when $\tau^{-\sigma}P^{\sigma-1}Y$ is larger. Gross capital supply equals demand in the capital market equilibrium condition,

$$\begin{aligned} K &= [1 - G(\omega_3)]\left[q^{-1}(q/\tau)^{\sigma} P^{\sigma-1}Y/a + fq\right] \\ &+ \int_{\omega_2}^{\omega_3}\left[\frac{\omega - fq}{1-\lambda\tau^2} + fq\right] dG(\omega) + \int_{\omega_1}^{\omega_2}\left[\frac{\omega - f}{1-\lambda\tau^2/q} + f\right] dG(\omega). \end{aligned}$$

When τ falls, the gross capital demand of entrepreneurs falls and more entrepreneurs become lenders. Hence, whenever some entrepreneurs are credit-constrained, $\tau^{-\sigma}P^{\sigma-1}Y$ must rise such that the capital market equilibrium condition holds. □

While our assumptions on the utility function could be easily relaxed, the assumption of fixed costs, or more generally increasing returns on the technology side, is key. The basic intuition is quite general: entrepreneurs who are most adversely affected by credit constraints are limited in their entrepreneurial actions. The payoff of the latter typically increases in the market size. However, financially constrained firms cannot grow as easily as unconstrained ones due to lack of external funds. Consequently, financially constrained entrepreneurs can take less advantage of market opportunities like trade opening as they make the environment more competitive, hence they are less likely to stay in the market.

As a final point, consider the case where a country pursues a unilateral trade liberalization. To analyze such a policy experiment assume that the South unilaterally lowers trade barriers such that the trade costs from the North to the South equal $\alpha\tau$, with $\alpha < 1$ and the trade costs from South to North still equal τ. Such a move affects only the limit price $p(y, q_j) = \alpha\tau q_j/q$ financially constrained entrepreneurs can charge. The interest rate r, however, is unaffected and equals a/τ because it is determined by the returns to capital for Southern exporters. The critical wealth levels read $\omega_1 = f(1-\lambda)\alpha\tau^2/(\alpha\tau^2 - q)$ and $\omega_2 = f(1 + q - \lambda\alpha\tau^2)$. Qualitatively, a reduction of α has a similar effect to a reduction in τ. Hence, a unilateral trade liberalization has similar effects on the firm structure as discussed in the propositions above.

3.3 Data

3.3.1 Enterprise Surveys and Tariff Data

The World Bank's Enterprise Surveys (WBES) provide firm survey data for more than 135 countries between 2002-2014. We use data from seven Latin American countries – Argentina, Bolivia, Chile, Colombia, Paraguay, Peru, and Uruguay – where firms were interviewed in 2006 and 2010 with standardized questionnaires. The restriction of our sample to Latin America is to ensure a certain homogeneity of countries. Moreover all of these countries are associated with Mercado Comun del Sur (MERCOSUR), a common market in South America promoting free trade. We included all countries for which data was available as a panel for two years with a moderate sample size.[1]

Each firm was surveyed by means of a standardized questionnaire. This questionnaire covers a wide range of topics, including firm characteristics as well as detailed information on the constraints that firms perceived as an obstacle to their business activity. We use this information to construct variables on the prevalence of credit constraints and the intensity of competition, market exit and several measures of productivity-enhancing activities. Moreover we derive a large set of control variables which are used in all of the regressions. The information

[1] Note that we excluded Brazilian firms because they were interviewed in different years, 2003 and 2009.

on financial constraints is derived from a question containing a list of potential obstacles to doing business. In total, the question lists sixteen obstacles, inter alia access to finance, crime, tax administration, tax rates and transportation. Firms are asked which of these obstacles constitute the most binding constraint. This determines our coding of an indicator variable for being financially constrained. The information on firm exit is obtained from the 2010 surveys, which track firms from the first panel round and record information on why firms were unavailable for the second period.

We combine the firm survey data with information on tariff rates from the World Integrated Trade Solution (WITS) database. Using tariff rates from 2006 and 2010 allows us to apply a precise measure of treatment for each individual firm on the four-digit International Standard Industrial Classification (ISIC) level.[1] Hence, we can clearly identify which firms were subject to trade liberalization (i.e., a reduction of tariff protection for their main product) within each country and sector. We do not, however, have information on the provenance and types of inputs that firms use. Thus we cannot compute firm-specific input tariff levels. However, the survey allows us to identify those firms that use imported inputs and also gives information on whether the inputs were imported directly or indirectly. We use this information as a covariate in our estimations.

We drop all observations from retail and services sectors to focus on the manufacturing sector. In total, our data set contains 5,278 observations from seven countries for the years 2006 and 2010. Among these observations, we have 754 firms which were interviewed in both years (balanced panel). The samples were stratified by industry with the main body of observations being from the textiles sector (35%), food sector (30%), and chemicals and paper sector (22%). The majority of panel observations were observed in Argentina (187 firms), Chile (191), Columbia (138), and Peru (123). Fewer observations were sampled in Bolivia (20), Paraguay (25), and Uruguay (70).

[1]For each firm the ISIC code is derived from data given in the firm survey.

3.3.2 Descriptive Statistics

Based on the seven countries and two years of observation we have a data set with a total of 5,278 unit observations. Table 3.1 provides summary statistics for all employed variables in the sample. Further summary statistics are provided in Table 3.7 of the Appendix.

– Table 3.1 about here –

The statistics indicate that there is considerable variation with respect to most variables across the seven countries. With respect to credit constraints the statistics show that about one in eight firms regards access to finance as the biggest obstacle to their business activity. This aggregate share has not changed between the two observed years 2006 and 2010. Table 3.1 also provides information on firm characteristics which we use as control variables in the estimations. We have deflated all monetary values and converted them from local currencies to 2006 US dollars, using exchange rates are taken from the World Bank Development Indicators. Many trade-related variables increase over time. This includes for example the share of firms that are directly importing goods, the share of foreign imported inputs, or the share of exporters. Depending on the country, between 30% and 70% of firms engage in productivity-enhancing activities.

Concerning trade liberalization we observe that only Peru and Uruguay lowered their tariffs on average *over all industries* between 2006 and 2010. In the case of Peru this is due to the United States–Peru Trade Promotion Agreement (PTPA) which was signed on April 12, 2006. According to the Office of the United States Trade Representative, the PTPA provides a secure and predictable legal framework for investors, while strengthening protection for intellectual property, workers, and the environment. Eighty percent of U.S. exports of consumer and industrial products to Peru became duty-free and the remaining tariffs were set to phase out over ten years.[1] For Uruguay, several bilateral trade agreements with the US were also signed between 2006 and 2008. In contrast, Bolivia shows a

[1] More details on the PTPA are provided on the website of the United States Trade Representative (www.ustr.gov/trade-agreements/free-trade-agreements/peru-tpa).

significant increase in average tariff rates which may be due to more protectionist policies following the election of president Evo Morales in January of 2006.

We provide information on the seven countries' main trading partners in the Appendix. For all countries in our dataset we find similar main trading partners. China, the United States, and the European Union account for a large share of both imports and exports. However, there is substantial variation in the trade balance as well as the importance of trade among the seven countries. Bolivia, Chile and Paraguay have a very large share of trade in GDP of more than 70 percent. Argentina shows a large surplus in the trade balance while Colombia and Peru have large deficits.

For the control variables in our estimation, we only use the 2006 information as well as a variable capturing participation of the firm in the follow-up survey in 2010. There are numerous reasons why firms did not participate in the next round. However, only for nine firms the reason was unknown. Thus attrition is not an issue for our approach when exit is the dependent variable because we can use the full first-period sample. When R&D-related activities are used as the outcome (which implies that a firm is still operating in 2010) we restrict our sample to a balanced panel of firms. Table 3.1 provides the respective summary statistics for all relevant variables in this sub-sample.

As noted earlier, follow-up information was not available for many firms for a variety of reasons. This raises the question of how important non-random attrition is in our sample. Overall, the differences with respect to most variables are rather small. Using a simple two-sample t-test, we find a few significant differences at the 5% level. Not surprisingly, the means for credit constraints, competition, market exit, firm size and age are different in the panel and non-panel samples. Some form of non-random attrition can be found in virtually any panel data set and can hardly be corrected for. We use the available information on which firms exited and why (see above reasons) when identifying the effects of trade liberalization and constraints on firms' decisions to engage in R&D-related activities. In a sample in which the most constrained firms are excluded (because of business failure), only considering surviving firms induces a downward bias of our estimates. This does not change the bottom line of our reasoning. Since we expect a negative effect of liberalization on credit-constrained firms' propensity to invest in quality upgrading, the presence of non-random attrition will cause us to *underestimate* the effect. This makes it more difficult to find any significant correlation between credit constraints, tariff reductions, and innovative activity.

An important final concern addresses the issue of multi-product firms. As in the related literature, multi-product firms are larger in terms of number of employees and more likely to engage in exporting compared to single product firms. In order to mitigate the problem of multi-product firms, we drop observations where the share of the main product is less than 30 percent. Moreover we control for the share of the main product in all regressions.

3.4 Empirical Results

As shown in Section 3.2, reducing trade costs makes it more difficult for small and medium-size firms to borrow capital. In addition, equation 3.5 shows that output among credit-constrained firms responds negatively to reduced tariff protection. As a result, we expect to see an increase in the probability of market exit among financially constrained firms in sectors with reduced tariffs. Moreover, investments in R&D-related activities among surviving (financially constrained) firms is expected to decrease.

We use our data set on seven Latin American countries in order to test these predictions. In a first step, we explore empirically whether tariff cuts worsen small firms' access to credit. Furthermore, we test whether annual sales of constrained firms respond negatively to trade liberalization. In the second step, we examine in detail the joint impact of credit constrains and tariff reductions. In particular, we test for the positive effect on market exit as well as the negative impact on R&D efforts.

3.4.1 Trade Openness and Access to Credit

One of the key insights of our theoretical model is that trade openness can worsen access to finance for some firms. Due to data limitations we cannot provide a structural estimation of the model. However, we can use our data set to test whether the model's main predictions find empirical support. In a first step, we estimate the effect of tariff cuts between 2006 and 2010 on the probability of being credit constrained in 2010. The regression is given by

$$\begin{aligned} CC_{j,i,c,t} = \;& \beta_1 CC_{j,i,c,t-1} + \beta_2 \Delta T_{j,i,c,t} + \beta_3 \Delta T^2_{j,i,c,t} \\ & + \gamma \mathbf{X}_{j,i,c,t-1} + \mu_c + \delta_i + \varepsilon_{j,i,c,t} \end{aligned} \tag{3.6}$$

where $CC_{j,i,c,t-1}$ is a dummy variable indicating whether firm j in industry i and country c reports access to credit being a problem in the first period (2006). $\Delta T_{j,i,c,t}$ is the change in tariff rates at the firm level, calculated as the difference between the two periods 2006 and 2010. Note that we add a squared term of $\Delta T_{j,i,c,t}$ to assess a non-linear correlation. Due to sufficient sample size we can add country and industry fixed effects to account for unobserved factors not included in the vector of controls $\mathbf{X}_{j,i,c,t}$. The control variables contain firm size, age, foreign ownership, share of main product, and being an exporter. We report the estimates in Table 3.2.

– Table 3.2 about here –

The results provide evidence of persistence in credit constraints among small and medium-sized firms (columns 1 and 2). Moreover, in this group we observe that large tariff cuts after 2006 are associated with a significantly higher probability of being credit constrained in 2010. This is in line with findings by Föllmi and Oechslin [2014] who use a Difference-in-Difference approach. Small firms' financial health appears to be adversely affected by tariff reductions. Especially large tariff cuts are associated with increased problems in access to finance. This is notably different among large firms where a reduction of tariff rates does not worsen access to finance (columns 3 and 4).

Our model also predicts that among financially constrained firms, output is negatively affected by reductions in trade costs (cf. equation 3.5). In the second part of Table 3.2 we test this prediction. Except for obvious changes, the regression we run is similar to the one shown above. Restricting the sample to only those firms which were financially constrained in 2006, however, implies that we have a relatively small number of observations. Nevertheless, we find significant evidence that tariff cuts are adversely related to output in 2010 (columns 5 and 6). Again in sharp contrast, unconstrained firms show a notably different response to trade liberalization. Large tariff cuts after 2006 are associated with higher annual sales in 2010 (columns 7 and 8).

3.4.2 Market Exit and Quality Upgrading

Based on the model in Section 3.2 we expect firms to respond heterogeneously to tariff reductions. Some firms will continue their business and invest resources into quality-upgrading, some will cut expenses on R&D efforts, and some will exit the market entirely. Therefore our empirical findings on the joint effects of financial constraints and trade liberalization are split into two parts. We start with an evaluation of the idea that a reduction in tariff protection increases credit-constrained firms' propensity to leave the market as suggested in Melitz and Ottaviano [2008] as well as Föllmi and Oechslin [2010]. Second, we estimate the effect of tariff reductions on surviving firms' propensity of developing innovative products and production processes or filing patents.

3.4.2.1 Econometric Approach

Our estimation uses a cross-section of firms where the effect of interest concerns the interaction term on liberalization and credit constraints. The baseline regression is given by

$$\begin{aligned} Y_{j,i,c,t} = \;& \beta_1 CC_{j,i,c,t-1} + \beta_2 \Delta T_{j,i,c,t} + \beta_3 \Delta T^2_{j,i,c,t} \\ & + \beta_4 CC_{j,i,c,t-1} \times \Delta T_{j,i,c,t} \\ & + \beta_5 CC_{j,i,c,t-1} \times \Delta T^2_{j,i,c,t} \\ & + \gamma \mathbf{X}_{j,i,c,t-1} + \mu_c + \varepsilon_{j,i,c,t} \end{aligned} \tag{3.7}$$

where $Y_{j,i,c,t}$ is a dummy variable, capturing the outcome variable obtained from the second period (2010). This can be firm exit, product or process innovation, or filed patents. As before in equation (3.6), $CC_{j,i,c,t-1}$ denotes financial constraints of firm j in industry i and county c in the first period (2006). The change in tariff rates at the firm level, $\Delta T_{j,i,c,t}$, is calculated as the difference between the two periods 2006 and 2010. $\mathbf{X}_{j,i,c,t-1}$ is a vector of control variables including firm size, firm age, foreign ownership, degree of competition, share of main product, foreign input share, being a direct importer, being an exporter, and the share of labor cost.[1] Country-fixed effects are denoted by μ_c while $\varepsilon_{j,i,c,t}$ is the standard error clustered at the country level.[2] The main effect of interest is given by β_4 and β_5 on the interaction terms. These indicate the differential impact of liberalization on firms that were credit-constrained in the initial period, that is before changes in tariff rates. Note that we check for a non-linear effect by adding the squared term of tariff changes. Since all of outcome variables $Y_{j,i,c,t}$ are dummy variables, we use a probit estimator.[3] Estimation of the marginal effects at mean values of covariates takes into account prior work by Ai and Norton [2003] as well as Norton, Wang and Ai [2004].

The standardized questionnaire allows us to draw on two different statistics for market exit as an outcome. First, exit can be defined as business failure which was confirmed in the 2010 survey. Alternatively, following a broader definition, we define exit such that it includes firms which could not be contacted in 2010. This does not contain cases of simple relocation or unwillingness to participate. Instead, this type of market exit may be the result of a dead phone line, a new and unknown postal address, or an unregistered business failure. Note that for our estimations, we entirely use the latter definition. Descriptive statistics indicate that only three percent of firms left the market using the strict definition. In contrast, with the broader definition we have about 14% closing their operations. In order to secure a sufficient number of observations we restrict our estimation

[1]The extent of competition in a given product market has a strong effect on both outcome variables (see, for example, Melitz and Ottaviano [2008]). Using information from the World Bank surveys, we can address this issue. In particular, we define a competition dummy variable that takes a value of one for firms which reported to have five or more competitors in their market.

[2]Note that we add industry fixed effects in the robustness section. Given the small number of observations in regressions with R&D-related outcomes, we prefer the specification without industry fixed effects. However, as shown in Table 3.5, our findings are robust to controlling for industry-specific effects.

[3]When replacing the probit estimator with a logit estimator, the results are very similar.

to the use of this definition of market exit.

The information on financial constraints, $CC_{j,i,c,t-1}$, is taken from the World Bank Enterprise Survey (WBES) questionnaire. If a firm nominates access to finance as the single most important obstacle to their business activity, the financial constraint dummy takes a value of one, otherwise zero.[1] This information is available for both periods 2006 and 2010. In order to identify the joint effect of constraints and liberalization only the 2006 value for the constraint is used in this part of our analysis while the 2010 value is used as an outcome in the previous part.

To identify causal effects of trade liberalization, the change in tariff rates has to be exogenous. Although it is possible that firms influence policies through lobbying it is unlikely that many firms in our sample had sufficient leverage to manipulate national policies. For one thing, the median firm in our data set has only 29 employees. In addition, changes in this policy dimension are often induced by international policies such as free trade agreements or regional organizations. The Peruvian liberalization between 2006 and 2010, for example, was related to the negotiations of the *Peru Trade Promotion Agreement* with the USA, signed in April 2006. We also provide a simple test for endogeneity by regressing tariff changes on a number of firm characteristics. In particular, we fit the linear regression

$$\Delta T_{j,i,c,t} = \beta CC_{j,i,c,t-1} + \gamma \mathbf{X}_{j,i,c,t-1} + \varepsilon_{j,i,c,t} \tag{3.8}$$

where $\Delta T_{j,i,c,t}$ is the tariff change for a specific ISIC code at the four-digit level faced by firm j in industry i and country c between 2006 and 2010. Firm characteristics as of 2006 are summarized as $\mathbf{X}_{j,i,c,t-1}$ and $\varepsilon_{j,i,c,t}$ is the standard error term clustered at the country level. Table 3.3 provides the estimates.

– Table 3.3 about here –

Basically none of the employed firm variables shows a significant effect on tariff changes in the respective ISIC code. In particular, firm size and age in 2006

[1] The question reads: "You have indicated that several obstacles affect the operation of this establishment. Here is a card with the obstacles I mentioned throughout the interview. Please tell me the three that you think are currently the biggest problem, beginning with the worst of all three."

do not appear to explain liberalization patters.[1] Most importantly, firms with financial constraints in 2006 are not more likely to be subject to tariff changes.[2] Nevertheless, we cannot fully rule out that liberalization was —in some unobserved ways— influenced by firms or accompanied by other policies. Hence we interpret our results as robust correlations rather than causal effects. It is, however, noteworthy that any endogeneity would make it more difficult to observe the patterns we see in the data. In particular, policymakers are less likely to lower tariffs in sectors struggling because of financial constraints.

3.4.2.2 Findings for Market Exit

In a first step, we examine the joint effects of trade liberalization and credit constraints on market exit. As explained earlier, we add a measure for the degree of competition to the list of control variables because it is one of the key drivers of market exit. In all regressions we apply the broad definition of market exit which includes firms that could not be contacted for reasons that indicate business failure. The estimation results are shown in the first two columns of Table 3.4.

– Table 3.4 about here –

Our estimates suggest that a tariff reduction among firms facing credit constraints in the initial period led to increased market exit. Reducing tariff protection by one percentage point for these firms is associated with a 0.5 percentage point increase in the probability of leaving the market. Larger tariff cuts are associated with increasing effects as indicated by the significant coefficient on the squared term. This is in line with the idea that firms can absorb small shocks but leave the market in case of large disruptions.

3.4.2.3 Findings for Quality Upgrading

For those firms who do not leave the market, we expect to see a negative effect of liberalization on credit-constrained firms' propensity to develop innovative prod-

[1] The only exception is a weak significant coefficient on firm size in column (3). However, once we add industry-fixed effects, the coefficient loses significance and even turns positive.

[2] With very few exceptions, none of the industry dummy variables shows a significant coefficient. Not surprisingly, however, most of the country dummy variables are significant.

ucts and production processes. We provide the respective regressions in columns 3–8 of Table 3.4. Concerning the impact of tariff reductions itself we find some (weak) evidence of a positive impact on R&D-related activities. This is in line with previous research [Bustos, 2011*b*; Lileeva and Trefler, 2010]. The more important finding for our study, however, is the effect of trade liberalization on credit-constrained firms. Irrespective of which R&D measure we use as outcome variable, we observe a significant negative coefficient on the interaction term of trade liberalization and credit constraints. The significance of the coefficients is also present —and in some cases even larger in magnitude— when adding country-fixed effects to control for unobserved factors. Estimation results suggest that a one percentage point decrease in tariff protection reduces constrained firms' probability of introducing innovative products or processes by about ten percentage points. Moreover, the probability of filing a patent is reduced by three to four percentage points among these firms. Unlike in the case of market exit (columns 1 and 2), there is only mixed evidence of a non-linear effect of larger tariff cuts. This may, however, be due to the limited number of observations in the last four regressions.

Overall, these findings support our theoretical predictions and suggest that limited access to finance not only distorts exporting behavior —as has been shown in previous research— but also correlates with R&D-related activities at the firm level. In Figure 3.2, we illustrate our results.

– Figure 3.2 about here –

The four plots indicate the marginal effects of credit constrains on our outcome variables, depending on the magnitude of tariff change ranging from plus to minus twenty percentage points. Panel (a) shows that firms facing credit constraints are increasingly likely to exit the market when being exposed to tariff reductions. This is not the case with unconstrained enterprises. Panels (b), (c) and (d) illustrate the sharp difference between financially unconstrained and constrained firms with respect to innovative activity. While the former respond positively to tariff cuts, the latter sharply reduce R&D efforts.

We can relate this finding to prior research on trade liberalization and firm-level R&D investments [Atkeson and Burstein, 2010; Bustos, 2011*b*; Lileeva and Trefler, 2010]. The observation that firms facing reductions in tariffs increase their investment in technology upgrading has been established before. The very different response by credit-constrained firms, however, is a novel finding.

3.4.3 Alternative Explanations and Robustness Tests

In order to verify our empirical findings we conduct a number of robustness tests. First, we examine whether the estimates are robust to changes in the econometric approach. Second, we address potential conceptual concerns.

Industry-Specific Differences — Following prior research by Rajan and Zingales [1998], we take into account the idea that different industries rely differently on external finance. As a result, firms in sectors with high dependence on external funding are more vulnerable to capital market frictions. We address this by adding industry-fixed effects to the right-hand side of equation 3.6.

– Table 3.5 about here –

Estimates in Panel A of Table 3.5 show that when adding dummy variables for each industry, we still obtain coefficients on the interaction term of credit constraints and trade liberalization that are similar to those in our main estimation.

Due to the small sample size adding many fixed effects, however, may be problematic. With data from seven countries and more than forty industries, we lose several degrees of freedom in the most demanding regression. Therefore we run a conditional (fixed-effects) logistic regression to test whether our estimates are robust [Chamberlain, 1980]. Results in Panel B indicate that for all four outcome variables we still observe similar coefficients.

Small Firms — It is a well-established fact that a firm's size is strongly related to its probability of being financially constrained. In fact, our model predicts that poorer entrepreneurs run smaller firms and face more difficulties in terms of access to credit. This association is also supported by our data. The status of being financially constrained is more prevalent among small firms.

The numbers indicate that small firms are 36 percent more likely to be credit-constrained than large firms. Moreover there is evidence of serial correlation in the sense that firms reporting credit constraints in 2006 are much more likely to do so in 2010. The fact that firm size and access to credit are correlated may cause a problem for our empirical analysis. It could be argued that our dummy variable for being credit constrained ($CC_{j,i,c,t-1}$) captures other characteristics of small firms. If so, the effects reported above may not (solely) be driven by the importance of financial frictions.

We address this concern by running a 'horse race' between access to credit and being a small firm. In particular, we use a dummy variable for being a small firm in 2006 and use it in the same way as our indicator of being credit constrained. In case it is indeed access to finance which determines firms' behavior, only the interaction term with credit constraints should turn out to be significant. The results in Panel C of Table 3.5 show that the interaction terms with the small-firm dummy are mostly insignificant while the credit constrained dummies remain significant and similar in magnitude to the baseline regression. We take this as evidence that it is indeed access to credit causing differential responses to trade liberalization.

Weighted Average Tariff Rates — Another robustness test concerns the measure of trade liberalization. Throughout our estimations we used the simple average tariff rate at the four digit ISIC code to determine whether firms were subject to a reduction in tariff protection. In Panel D of Table 3.5 we instead use the weighted average tariff rate. Overall, the results are very similar to our baseline estimates.

Heterogeneous Financial Development — The countries in our sample differ substantially with respect to their economic development. We expect to see the impact of credit constraints and tariff reductions to be magnified if capital market imperfections are more severe. In our model this would refer to a decrease in λ, the parameter governing the degree of imperfection of the capital market. To test this we split our sample and run the same regressions as before using only firms from less developed countries. Drawing on data from the International Monetary Fund, the GDP per capita (PPP) as of 2013 differs substantially among the seven countries of our sample: Argentina \$22'300, Bolivia \$5'900, Chile \$22'500, Colombia \$12'800, Paraguay \$8'000, Peru \$11'600, Uruguay \$19'700. We can use this information and restrict the sample to firms in Bolivia, Colombia, Paraguay and Peru.[1] This leaves us with roughly half the number of observations. Note that in our survey data from the World Bank, firms located in Colombia and Paraguay are most likely to report difficulties getting access to credit. This supports the use of income per capita as a proxy for the development of a country's financial sector.

When reducing the sample to the least developed countries and running the same regressions as in Table 3.4, we find support for our hypothesis. The point estimates for the joint impact of credit constraints and tariff cuts are larger in the restricted sample.[2] Significance levels, however, are obviously lower given the substantial reduction in the number of observations.

Financial Crisis — A final concern we address is the impact of the 2007–2008 global financial crisis. Since we employ data from the years 2006 and 2010, the question arises how the crisis affects our empirical findings. In our empirical analysis we explore the differential impact of trade liberalization among firms that faced credit constraints in 2006 and those who did not. Thus we define treatment and control group based on firm characteristics determined before the stock market crash. The shock to the financial system in 2007, however, affected the countries in our sample. We can illustrate this by considering annual growth rates of real GDP. The numbers in Table 3.11 of the appendix show that on average the annual growth rate in 2008 and 2009 was three percentage points lower than in the two years before the financial crisis.

---

[1] Note that although we split the sample by GDP per capita, we could also use various financial indicators such as bank accounts per 1,000 adults, credit to private sector as a share of GDP, financial system deposits as a share of GDP, or the stock market capitalization to GDP. All of these indicators are highly correlated with income per capita. Hence, selecting the sample of less developed countries based on these indicators would leave us with a similar or even identical sample.

[2] We provide all estimation results using the restricted sample in Table 3.6 of the appendix.

The heterogeneity across countries, however, does not affect our empirical findings as we add country fixed effects to the regression. Moreover, if trade policy was affected by the financial crises the most plausible bias would make it more difficult to observe the patterns in the data that we document: Policymakers are less likely to lower tariffs for firms in sectors that are vulnerable to shocks in the capital market.

3.5 Conclusion

In this study we explore heterogeneous responses to trade liberalization at the industry and firm level. As illustrated by a simple theoretical model, we expect financially constrained firms to have a higher probability of leaving the market after liberalization. For surviving firms, the model suggests limited investments in quality-upgrading when firms are financially constrained and face tariff cuts. Using firm level survey data from seven Latin American countries for the period from 2006 to 2010, we assess these predictions empirically. In line with our theory, we find that financially constrained firms being subject to liberalization are associated with more market exits, less product and process innovations, and less filed patents. This impact is shown to be robust to the inclusion of various control variables as well as country- and industry-fixed effects. Moreover, the findings are not driven by any single country or the specification of our regression equation. Instead the impact is generally magnified in less developed countries.

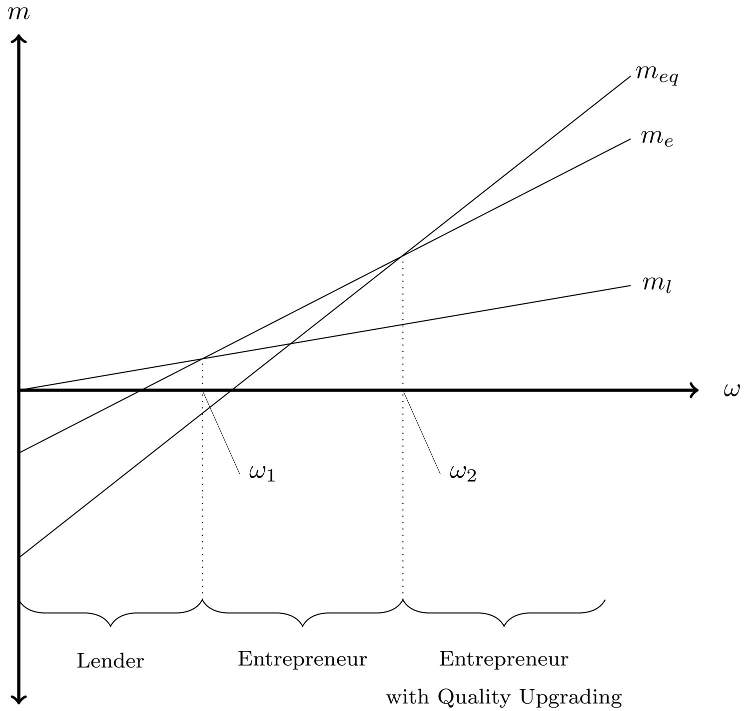
Our results add another dimension to the evidence on how financial constraints affect and distort firm behavior. In the presence of imperfect capital markets, adjustments after trade liberalization are limited at the firm level. As a result, gains from openness can be reduced in developing countries. This adds to the recent literature arguing that low aggregate total factor productivity —especially in developing countries— is the result of a resource misallocation at the firm level [Hsieh and Klenow, 2009; Song, Storesletten and Zilibotti, 2011]. In a broader sense our findings suggest that reductions in the magnitude of one distortion (here: tariffs) do not necessarily lead to a welfare gain if there are other distortions (credit market imperfections) in the economy [Bhagwati, 1971].

In this sense, the findings of our study are linked to policy implications. In the presence of credit constraints, adjustments at the firm level can be impaired. Hence optimal policies must take into account the fact that sectors differ in their reliance on external finance and therefore in their ability to adjust their production to a post-liberalization environment. The importance of credit constraints with respect to international trade has been emphasized by Manova, Wei and Zhang [2015]. Their findings suggest that FDI can reduce liquidity constraints at the firm level. Moreover, firms' credit rating has been shown to affect their propensity to engage in international trade [Muûls, 2015].

There are a number of caveats to our conclusions. First and foremost, tariff reductions across sectors are typically non-random. As a result, identifying causal effects of trade liberalization at the firm level remains challenging. We provide empirical evidence showing that firm characteristics such as size and age in the initial period do not correlate significantly with subsequent tariff changes. In addition, firms with credit constraints before trade policy changes were not more likely to be subject to trade liberalization. Moreover, we add country- and industry-fixed effects to control for unobserved factors. All of this does not affect our empirical findings. Regarding the magnitude of welfare gains from trade, however, Dehejia and Panagariya [2014] suggest that there may be positive spillover effects from liberalization in manufacturing to gross value added, wages, employment, and worker productivity in services. Hence, exploring firms' decision to move from manufacturing to service as a response to trade liberalization appears to be a promising field for future research.

Figures and Tables

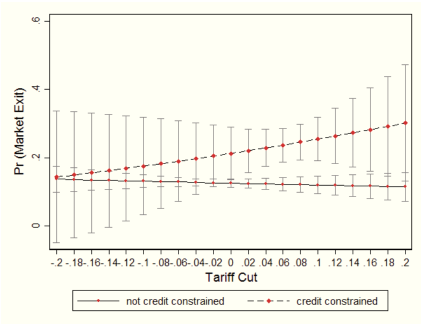
Figure 3.1: Entrepreneurial Decisions by Endowment



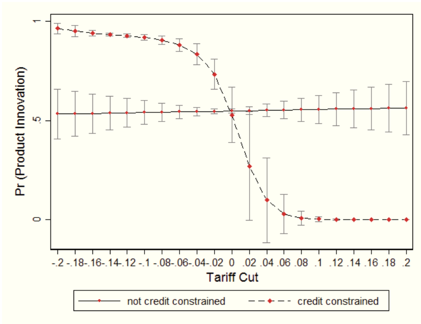
Note: The figure illustrates cutoff values for the initial capital endowment, ω, which determine whether individuals become lenders and receive a nominal income of m_l, become entrepreneurs (m_e) or become entrepreneur and invest in quality upgrading (m_{eq}). Individuals with an initial endowment below ω_1 choose to become lenders while those with a larger endowment become entrepreneurs. If the initial endowment is larger than ω_2, individuals become entrepreneurs and invest in quality upgrading.

Figure 3.2: Marginal Effect of Credit Constraints by Tariff Change

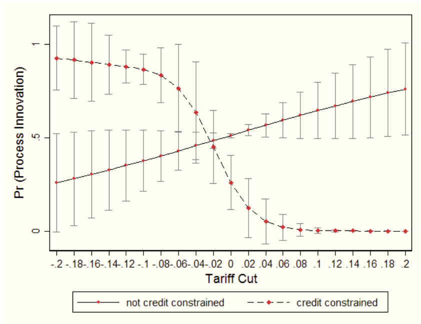
(a) Market Exit



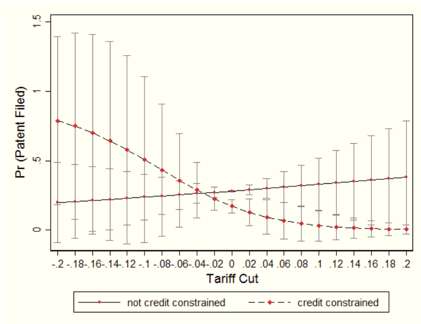
(b) Product Innovation



(c) Process Innovation



(d) Patent Filed



Note: The figures show the marginal effects of credit constraints at different levels of tariff changes, ranging from an increase of 20 percent to a reduction of 20 percent. Tariff Cut is calculated as the four-digit ISIC level tariff in 2006 minus the tariff in 2010. Market exit is defined as business failure including businesses not found in 2010. All three measures of R&D are dummy variables for innovative activity in the three years prior to 2010.

Table 3.1: Summary Statistics for the Panel Sample

	ARG	BOL	CHL	COL	PAR	PER	URY	Total
Tariff 2006	0.12	0.07	0.02	0.15	0.08	0.14	0.11	0.10
Tariff 2010	0.14	0.13	0.05	0.15	0.09	0.08	0.11	0.10
Tariff Cut	-0.02	-0.08	-0.03	0.00	-0.01	0.06	0.00	-0.01
Tariff Cut sq.	0.29	1.46	0.07	0.02	0.02	0.58	0.01	0.23
Credit constrained 2006	0.12	0.10	0.12	0.11	0.16	0.10	0.07	0.11
Credit constrained 2010	0.14	0.10	0.15	0.14	0.08	0.07	0.03	0.12
Market Exit (non-panel)	0.05	0.21	0.08	0.29	0.10	0.14	0.12	0.14
Product Innovation	0.51	0.67	0.56	0.48	0.53	0.65	0.40	0.54
Process Innovation	0.55	0.79	0.59	0.42	0.50	0.55	0.33	0.52
File Patent	0.25	0.25	0.22	0.22	0.40	0.34	0.34	0.26
Small Firm	0.37	0.23	0.31	0.46	0.20	0.33	0.38	0.36
Medium-Size Firm	0.40	0.38	0.49	0.42	0.40	0.44	0.44	0.44
Firm Age	32.62	23.58	31.70	19.84	28.94	23.39	33.58	28.27
Foreign Ownership	0.11	0.25	0.09	0.04	0.16	0.08	0.07	0.09
Share Main Product	0.70	0.73	0.82	0.76	0.76	0.73	0.77	0.76
Share Foreign Inputs	0.27	0.52	0.40	0.24	0.47	0.35	0.50	0.35
Direct Importer	0.55	0.70	0.54	0.32	0.65	0.56	0.66	0.53
Exporter	0.47	0.45	0.31	0.36	0.50	0.49	0.44	0.41
Share of Labor Cost	0.27	0.29	0.31	0.33	0.40	0.26	0.27	0.30
Competition	0.59	0.39	0.57	0.68	0.66	0.60	0.61	0.60
Employees	116.95	92.70	81.47	65.57	92.78	132.59	50.26	93.47
Log Annual Sales	13.98	13.66	14.17	13.24	13.94	13.91	13.78	13.85
# Observations	374	40	382	276	50	246	140	1,508

Note: The table shows descriptive statistics (mean values) for the panel sample of firms which are surveyed in both 2006 and 2010. Tariff Cut is calculated as the four-digit ISIC level tariff in 2006 minus the tariff in 2010. Market exit is defined as business failure including businesses not found in 2010. All three measures of R&D are dummy variables for innovative activity in the three years prior to 2010.

Table 3.2: Tariff Cuts, Credit Constraints, and Annual Sales

	Credit Constrained 2010				Log Annual Sales 2010			
Mean value	0.15		0.09		13.42		14.03	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tariff Cut	-0.224	-0.436	-0.022	0.077	-0.506	-1.208*	2.448**	0.816
	(0.366)	(0.529)	(0.144)	(0.336)	(0.803)	(0.622)	(0.737)	(0.867)
Tariff Cut sq.	0.023*	0.023**	0.009	0.011	-0.117**	-0.108*	0.076**	0.072**
	(0.012)	(0.009)	(0.010)	(0.014)	(0.039)	(0.051)	(0.022)	(0.025)
Credit cons.	0.300***	0.282**	0.087	0.090				
	(0.076)	(0.081)	(0.051)	(0.055)				
Log Sales					0.690***	0.696***	0.861***	0.840***
					(0.078)	(0.077)	(0.056)	(0.049)
Sample	small & medium-size firms		large firms		credit constrained firms in 2006		not credit cons. firms in 2006	
Control Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Country	Country Industry	Country	Country Industry	Country	Country Industry	Country	Country Industry
Observations	1,138	1,138	300	300	359	359	1,079	1,079
R-squared	0.114	0.130	0.067	0.084	0.781	0.793	0.863	0.870

Note: The table shows four separate OLS regressions using different dependent variables as indicated in the top row. In columns (1) and (2) the sample is reduced to firms with less than 100 employees. For columns (3) and (4) the sample includes all firms stating that access to financing (availability and cost) was a 'major' or 'very severe' obstacle in 2006. Tariff Cut is calculated as the four-digit ISIC level tariff in 2006 minus the tariff in 2010. Control variables include firm size, firm age, foreign ownership, competition, and the share of sales of main product. Standard errors are clustered at the country level and shown in parentheses. Significance levels are as follows: *0.10, **0.05, ***0.01.

Table 3.3: Endogeneity of Tariff Changes

	Tariff Cut (in %)			
Mean of Dep. Var.	-0.944			
	(1)	(2)	(3)	(4)
Firm size 2006	0.159	-0.551	-0.373*	0.071
	(0.236)	(0.584)	(0.187)	(0.103)
Firm Age 2006	0.001	-0.003	0.012	-0.001
	(0.015)	(0.014)	(0.009)	(0.007)
Foreign Ownership 2006		-0.136	0.154	-0.295
		(1.115)	(0.586)	(0.502)
Log Annual Sales 2006		0.402	0.387	0.077
		(0.541)	(0.254)	(0.151)
Exporter 2006		-0.153	-0.784	-0.111
		(0.629)	(0.828)	(0.571)
Credit Cons. 2006		-0.266	-0.099	-0.099
		(0.324)	(0.122)	(0.134)
Fixed Effects			Country	Country Industry
Observations	2,976	2,969	2,969	2,969
R-squared	0.001	0.011	0.412	0.568

Note: The table shows four separate OLS regressions using tariff changes as dependent variable. Standard errors are clustered at the country level and shown in parentheses. Significance levels are as follows: *0.10, **0.05, ***0.01.

Table 3.4: Main Results – Exit and Innovation

	Exit		Innovate Product		Innovate Process		File Patent	
Mean value	0.14		0.54		0.56		0.26	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tariff Cut	-0.115	-0.054	0.605	0.094	0.818	1.526	0.854*	0.495
	(-0.233)	(0.103)	(0.496)	(0.373)	(0.586)	(0.966)	(0.514)	(0.870)
Tariff Cut sq.	0.008	0.012***	0.042	-0.027	0.187***	0.133***	0.073	0.070
	(0.011)	(0.004)	(0.060)	(0.038)	(0.054)	(0.029)	(0.048)	(0.048)
Credit Cons.	0.063	0.070	0.167***	0.157**	-0.021	-0.069	-0.041	-0.020
	(0.044)	(0.046)	(0.063)	(0.070)	(0.191)	(0.169)	(0.055)	(0.061)
CC x TC	0.540**	0.490	-9.933***	-10.412**	-11.007*	-10.110**	-2.948*	-4.252**
	(0.279)	(0.345)	(4.343)	(4.581)	(6.513)	(5.453)	(1.873)	(2.020)
CC x TC sq.	0.060***	0.069***	-1.703***	-1.754***	-3.838	-3.139	-1.010	-1.280
	(0.019)	(0.027)	(0.598)	(0.623)	(4.961)	(4.400)	(0.744)	(0.792)
Control Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects		Country		Country		Country		Country
R-squared	0.058	0.104	0.046	0.081	0.046	0.103	0.088	0.094
Observations	1,025	1,025	236	236	197	197	361	361

Note: The table shows eight separate Probit regressions using different dependent variables as indicated in the top row. Coefficients show marginal effects at the means of covariates. Tariff Cut is calculated as the four-digit ISIC level tariff in 2006 minus the tariff in 2010. Control variables include firm size, firm age, foreign ownership, competition, share of sales of main product, exporting status, share of foreign inputs, and share of labor costs. Standard errors are clustered at the country level and shown in parentheses. Significance levels are as follows: *0.10, **0.05, ***0.01.

Table 3.5: Robustness and Specification Tests

	Exit	Innovate Product	Innovate Process	File Patent
Mean of Dependent Variable	0.14	0.54	0.56	0.26
	(1)	(2)	(3)	(4)
A: Country and Industry FE				
Credit Constrained x Tariff Cut	0.567*	-8.942	-10.075*	-5.516**
	(0.339)	(5.978)	(6.955)	(2.758)
Credit Constrained x Tariff Cut sq.	0.070***	-1.547**	-3.679	-1.259
	(0.027)	(0.709)	(4.828)	(0.945)
Observations	1,008	220	193	352
B: Conditional Logit Estimation				
Credit Constrained x Tariff Cut	0.806	-16.139**	-14.408*	-2.390**
	(0.765)	(6.843)	(8.649)	(0.988)
Credit Constrained x Tariff Cut sq.	0.094**	-2.324***	-4.404	-1.013
	(0.039)	(0.802)	(7.391)	(0.810)
Observations	1,025	236	197	361
C: Horse Race vs 'Small Firm'				
Credit Constrained x Tariff Cut	0.409	-10.476**	-10.506*	-4.506**
	(0.522)	(4.545)	(5.739)	(2.112)
Credit Constrained x Tariff Cut sq.	0.065**	-1.785***	-3.317	-1.264
	(0.026)	(0.616)	(4.513)	(0.809)
Small Firm x Tariff Cut	0.210	-0.251	1.968	-1.029*
	(0.525)	(1.105)	(1.902)	(0.597)
Small Firm x Tariff Cut sq.	0.008	-0.097	0.471***	0.132**
	(0.017)	(0.075)	(0.154)	(0.051)
Observations	1,025	236	197	361
D: Weighted Avg. Tariff Rates				
Credit Constrained x Tariff Cut	1.074**	-2.747*	-11.255**	-7.261**
	(0.406)	(1.539)	(5.714)	(2.593)
Credit Constrained x Tariff Cut sq.	0.114***	-0.280	-1.521***	-0.431**
	(0.044)	(0.195)	(0.565)	(0.181)
Observations	1,008	220	193	352

Note: The table shows sixteen separate regressions using different dependent variables as indicated in the top row. Coefficients show marginal effects at the means of covariates. Except for Part B, we always apply a Probit estimator. Tariff Cut is calculated as the four-digit ISIC level tariff in 2006 minus the tariff in 2010. In Part D, tariff cut is based on weighted (instead of simple) tariff rates. All estimations include country fixed effects as well as control variables as in the baseline regression (c.f. Table 3.4). Standard errors are clustered at the country level and shown in parentheses. Significance levels are as follows: *0.10, **0.05, ***0.01.

Appendix

I. Further Tables

Table 3.6: Effect on Exit and Innovation in Less Developed Counties

	Exit		Inno. Product		Inno. Process		File Patent	
Mean value	0.21		0.56		0.57		0.27	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tariff Cut	-0.602***	-0.098	0.896	1.097	-1.663	-2.924*	-0.560	-1.962
	(0.150)	(0.135)	(1.509)	(1.597)	(1.617)	(1.541)	(0.693)	(1.427)
Tariff Cut sq.	-0.004	0.015***	0.030	-0.043	0.400**	0.416***	0.140***	0.184***
	(0.012)	(0.002)	(0.089)	(0.082)	(0.161)	(0.111)	(0.054)	(0.063)
Credit Cons.	0.082	0.092	0.097	0.041	0.162	0.102	-0.063	0.019
	(0.067)	(0.344)	(0.148)	(0.130)	(0.272)	(0.232)	(0.135)	(0.162)
CC x TC	0.626	0.831***	-38.176	-40.124*	-21.302	-25.888	-4.257	-5.708
	(0.450)	(0.375)	(15.872)	(14.769)	(17.863)	(13.247)	(4.819)	(4.920)
CC x TC sq.	0.043***	0.046***	26.894	34.323	-27.864**	-27.804***	-0.149	-0.814
	(0.008)	(0.008)	(13.384)	(12.186)	(15.606)	(7.870)	(3.735)	(3.957)
Control Var.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Eff.		Country		Country		Country		Country
R-squ.	0.071	0.093	0.097	0.142	0.105	0.152	0.152	0.172
Obs.	546	546	116	116	111	111	180	180

Note: The table shows eight separate Probit regressions using data from less developed countries (Bolivia, Colombia, Paraguay and Peru). Dependent variables are indicated in the top row. Coefficients show marginal effects at the means of covariates. Tariff Cut is calculated as the four-digit ISIC level tariff in 2006 minus the tariff in 2010. Standard errors are clustered at the country level and shown in parentheses. Significance levels are as follows: *0.10, **0.05, ***0.01.

Table 3.7: Summary Statistics in Detail: Full Sample

Variable	Mean	Std. Dev.	Min.	Max.	N
Tariff 2006	0.10	0.06	0	0.31	5,050
Tariff 2010	0.10	0.07	0	0.3	5,029
Tariff Change	-0.01	0.05	-0.2	0.17	2,976
Tariff Change sq.	0.26	0.76	0	4.03	2,976
Credit cons. 2006	0.13	0.33	0	1	3,250
Credit cons. 2010	0.13	0.34	0	1	3,674
Market Exit	0.14	0.35	0	1	2,174
Product Innovation	0.54	0.50	0	1	1,821
Process Innovation	0.56	0.50	0	1	1,599
File Patent	0.26	0.44	0	1	2,897
Small Firm	0.37	0.48	0	1	5,278
Medium-size Firm	0.40	0.49	0	1	5,278
Firm Age	26.94	21.16	1	146	5,278
Foreign Ownership	0.10	0.30	0	1	5,274
Share Main Product	0.75	0.23	0.3	1	5,236
Foreign Input Share	0.37	0.34	0	1	5,245
Direct Importer	0.54	0.50	0	1	4,123
Exporter	0.42	0.49	0	1	5,275
Share Labor Cost	0.30	0.18	0	0.97	1,748
Competition	0.59	0.49	0	1	2,858
Employees	97.71	216.49	1	3200	5,278
Log Annual Sales	13.85	1.99	5.75	21.15	5,278
Panel Observation	0.29	0.45	0	1	5,278

Note: The table shows detailed descriptive statistics on all variables used in the empirical section. The sample includes all data.

Table 3.8: Summary Statistics in Detail: Panel Sample

Variable	Mean	Std. Dev.	Min.	Max.	N
Tariff 2006	0.10	0.06	0	0.27	1,448
Tariff 2010	0.10	0.07	0	0.3	1,443
Tariff Change	-0.01	0.05	-0.2	0.17	1,442
Tariff Change sq.	0.23	0.65	0	4.03	1,442
Credit cons. 2006	0.11	0.31	0	1	1,508
Credit cons. 2010	0.12	0.32	0	1	1,508
Market Exit	0	0	0	0	750
Product Innovation	0.54	0.50	0	1	473
Process Innovation	0.52	0.50	0	1	400
File Patent	0.26	0.44	0	1	751
Small Firm	0.36	0.48	0	1	1,508
Medium-size Firm	0.44	0.50	0	1	1,508
Firm Age	28.27	20.98	1	146	1,508
Foreign Ownership	0.09	0.28	0	1	1,505
Share Main Product	0.76	0.23	0.3	1	1,499
Foreign Input Share	0.35	0.32	0	1	1,498
Direct Importer	0.53	0.50	0	1	1,168
Exporter	0.41	0.49	0	1	1,508
Share Labor Cost	0.30	0.18	0	0.88	571
Competition	0.58	0.49	0	1	1,314
Employees	93.47	223.13	2	3200	1,508
Log Annual Sales	13.85	1.8	7.92	20.11	1,508
Panel Observation	1	0	1	1	1,508

Note: The table shows detailed descriptive statistics on all variables used in the empirical section. The sample is restricted to firms observed in 2006 and 2010.

II. Definition and Source of Variables

All firm-level information is taken from the World Bank Enterprise Survey (WBES). This includes variables on firm characteristics as well as the outcome variables (market exit and innovative activity). Using 4-Digit ISIC codes, we complement the firm data with information on tariff rates from the World Integrated Trade Solution (WITS) database. In Table 3.9 we provide details on how we constructed the variables for the empirical analysis.

Table 3.9: Variable Definitions and Data Sources

Variable	Definition	Source
Tariff	simple average level of nominal tariff protection	WITS
Tariff cut	tariff rate in 2006 minus tariff rate 2010	WITS
Tariff cut sq.	squared value of tariff cut	WITS
Credit constrained	= 1 if access to finance single most serious obstacle	WBES
Market Exit	= 1 if firm could not be surveyed in 2010	WBES
Product Innovation	= 1 if firm introduced new product in 2007–2010	WBES
Process Innovation	= 1 if firm introduced new production process 2007–2010	WBES
File Patent	= 1 if firm applied or filed for any patent in 2007–2010	WBES
Small firm	= 1 if 19 or less full-time employees	WBES
Medium-size firm	= 1 if 20–99 full-time employoyess	WBES
Firm age	years since operation began	WBES
Foreign owner	= 1 if positive share owned by foreign individuals / company	WBES
Share foreign inputs	share of total material inputs of foreign origin	WBES
Direct importer	= 1 if any material inputs are imported directly	WBES
Share of labor cost	wages, bonuses and social payments as share of total costs	WBES
Competition	= 1 if more than 5 competitors in market of main product	WBES
Employees	permanent full-time employees	WBES
Exporter	= 1 if any revenue from direct or indirect exports	WBES
Log annual sales	log total annual sales in 2006 USD	WBES

Note: When asked about the most serious obstacle, each firm was given a card with 16 different obstacles, including among others corruption, regulation, electricity, instability, taxes, or transportation. If the firm chose access to credit as the most serious obstacle, we code credit constrained as one.

III. Background Information on Countries in the Data

1) Trade Information

Our empirical analysis uses data from seven Latin American countries over the period 2006–2010. By means of Table 3.10, we provide background information on the countries' trading partners, trade balance, and trade openness.

Table 3.10: Trade Statistics for Countries in Our Dataset

Country	Main Importers	Main Exporters	Balance	Trade/GDP
Argentina	BRA (28), EU (17), USA (15), CHN(12)	BRA (21), EU (12), CHN (7), USA (6)	1,864	40.1
Bolivia	BRA (17), CHN (14), USA (13), EU (13)	BRA (34), ARG (20), USA (10), EU (7)	237	74.6
Chile	USA (20), CHN (20), EU (16), BRA (7)	CHN (25), EU (15), USA (13), JPN (10)	329	69.2
Colombia	USA (28), CHN (17), EU (13), MEX (9)	USA (32), EU (16), CHN (9), PAN (6)	-1,806	35.3
Paraguay	CHN (29), BRA (27), ARG (14), EU (8)	BRA (30), EU (14), RUS (10), ARG (9)	-627	100.2
Peru	USA (25), CHN (15), EU (11), BRA (5)	USA (18), CHN (18), EU (16), CAN (7)	-1,777	49.2
Uruguay	CHN (17), EU (15), BRA (15), ARG (14)	CHN (22), BRA (17) EU (15), ARG (5)	373	54.7

Note: Information on trade partners and balance is as of 2013 and taken from the European Commission. Numbers in parentheses express shares in percent. The trade balance is expressed in million Euro. Data on the share of trade in GDP is as of 2012 and reported by the WTO.

2) Financial Crisis

The time window for our empirical analysis is from 2006 to 2010. Hence, the 2007/08 financial crisis falls into this window. In Table 3.11 we show —for all seven countries in our data set— how real GDP grew during the time of our analysis. The impact of the financial crisis is visible in all countries with a steep drop in growth rates in 2009.

Table 3.11: Financial Crisis in the Countries in Our Dataset

Country	**Real GDP Growth Rate**				
	2006	**2007**	**2008**	**2009**	**2010**
Argentina	8.4	8.0	3.1	0.1	9.1
Bolivia	4.8	4.6	6.1	3.4	4.1
Chile	4.4	5.2	3.3	-1.0	5.8
Colombia	6.7	6.9	3.5	1.7	4.0
Paraguay	4.8	5.4	6.4	-4.0	13.1
Peru	7.5	8.5	9.1	1.0	8.5
Uruguay	4.1	6.5	7.2	2.4	8.4

Note: Information on GDP growth is taken from the World Bank database. Numbers indicate the annual change in percent based on 2005 U.S. dollars.

Chapter 4

Trading off Welfare and Immigration in Europe

This chapter is based on joint work with Ole-Petter Moe Hansen from the Norwegian School of Economics (NHH).

4.1 Introduction

Debates about immigration and the welfare state are growing in importance across Europe. The 2015 Eurobarometer, for example, shows that immigration, economic conditions, and unemployment are ranked as the main concerns among European citizens. For the first time in the survey's 42-year history, Eurobarometer finds immigration to be at the top of voters' concerns.[1] As a result of both the financial and the Euro crisis, many European countries have experienced particularly high unemployment rates and stagnating economies. Not surprisingly, this has spurred discussions about welfare benefits. Weakening economic conditions, however, have also affected individual views on immigration. And while

[1] For the U.S., the Polling Report shows that in December 2014, economic concerns and immigration ranked first and second in the list of most important issues that Congress should be dealing with in 2015. Recent press coverage by *The Economist* in an article entitled 'Looking for a home' (published August 29, 2015) emphasizes the importance of the topic.

discussions about the welfare system appear to often find a consensus, immigration issues are fiercely debated as seen in the case of the ongoing wave of Syrian refugees arriving in Europe. Established 'populist' right-wing parties such as *Front National* or *UKIP* achieved notable successes in recent elections while new political movements such as the *Alternative for Germany* have emerged. These parties differ from 'traditional' right-wing movements in their support for the welfare state. Instead of opposing any government intervention, new populist parties want to keep the welfare system but restrict access to natives.[1] However, their remarks on immigration have faced severe criticism from other established parties.

In this paper, we argue that these trends in European politics are the result of altered political preferences among voters. But how can we explain these shifts in preferences? Are individual views on immigration and the welfare state determined by ideology or can they be explained by economic factors? In general, when considering welfare and immigration from an economic point of view, it is important to note that the two policy dimensions are intertwined. A generous welfare state is irreconcilable with open immigration because supply of low-skilled immigrants would become infinite. And since low-educated natives are more likely to benefit from social expenditures, there is ample evidence showing that this group supports welfare spending but opposes immigration (e.g. Hainmueller and Hiscox, 2007).

Using data from the European Social Survey (ESS) on sixteen countries from 2002–2012, we first document that voter preferences shifted in favor of redistribution but polarized over low-skill immigration. There is a notable increase in the share of individuals supporting the welfare state but heavily opposing immigration. In the second step, we provide empirical evidence that individual characteristics as well as macroeconomic conditions are correlated with time trends in policy preferences. Our findings largely support prior results in the literature on redistribution and immigration. However, much of the literature assumes some

[1] This is illustrated by an article in *The Guardian*, entitled 'Marine Le Pen emerges from father's shadow', published March 21, 2011: "Her father led a movement of anti-system, extremist outsiders, [...] who railed against the state and loathed the public sector. But [Marine] Le Pen now styles herself as a defender of the republic, its benefits and welfare state, "the state as protector", she calls it. But not benefits for all. French people must come first."

form of xenophobia to explain why some natives reject immigration while the support for immigration is usually explained by labor market effects.

In our paper, we argue that neither xenophobia nor labor market effects from immigration are necessary to obtain a polarization in preferences over immigration. Moreover, in contrast to the literature asking the question why we see *opposition* to immigration in the absence of labor market effects, we ask why anyone *supports* immigration in the absence of wage and employment effects. To explain this, our theoretical model allows individuals to support low-skill immigration as well as transfers to the native poor potentially out of altruistic reasons.[1] For the simulation of our model, we use the ESS data to directly estimate the parameters of the utility function. One notable finding in this exercise is that low-educated natives generally like low-skill immigration. Using the estimated preference parameters, we find that a higher share of foreigners in the country shifts native preferences towards preferring less redistribution. A higher unemployment rate, in contrast, increases support for income redistribution among natives. The effect on preferences for immigration, however, depends on educational status. While the high-educated support immigration amid high unemployment rates, the low-educated natives show increased opposition to immigration. Finally, a more educated population shifts aggregate preferences towards less redistribution and more immigration. In total, our model replicates qualitatively the overall changes in policy preferences observed in the sixteen European countries between 2002 and 2012.

The estimation of utility preferences depends on strong assumptions, particularly on the approximation to the utility function and behavior of the government. Hence, it is reassuring to find that a flexible multinomial logit model—controlling for country- and year-fixed effects and other characteristics—yields similar results as in the structural model. A larger foreign-born population lowers support for redistribution while high unemployment is associated with more favorable views towards redistribution and a polarization over immigration.

Our study provides several novelties. First, we document important trends

[1]Note that our model allows for a different interpretation as well: Individuals could simply value diversity in the composition of the population. This is, however, not the focus of our study.

in policy preferences among European citizens. Among others, this adds to prior work by Mudde [2007] who discusses populist right-wing parties in Europe. By means of a novel theoretical model, we then illustrate the importance of considering preferences along both policy dimensions simultaneously. Only a few prior papers take into account both policies [Facchini and Mayda, 2008, 2009; Sánchez-Pagés and García, 2015]. Furthermore, our results support the idea that economic motives have in fact strong explanatory power for voter preferences. Finally, our paper indicates why support for 'populist' right-wing parties can arise in times of high unemployment rates. If individuals perceive a trade-off between free immigration and generous welfare spending, low-income voters tend to support anti-immigration but pro-redistribution movements. We argue that this trade-off, at times of high unemployment and increasing shares of foreign-born populations throughout Europe, is crucial for explaining observed changes in policy preferences over redistribution and immigration in recent years.

The paper proceeds as follows. In Section 4.2, we describe the data sources and provide summary statistics. Next, Section 4.3 documents that policy preferences over redistribution and immigration changed between 2002 and 2012. We investigate which macroeconomic variables account for such a change. Our theoretical model and its implications are shown in Section 4.4. We simulate voter preferences using the model's structure and parameters estimated based on the ESS data. Next, in Section 4.5 we discuss alternative explanations for the political trends in Europe. Finally, Section 4.6 concludes.

4.2 Data

For the empirical analysis we employ data from the European Social Survey (ESS). We use survey responses from all six biannual waves between 2002 and 2012. Table 4.1 shows all sixteen countries which participated in each wave of the ESS.[1]

[1] The set of countries includes Belgium, Denmark, Estonia, Finland, France, Germany, Hungary, Ireland, Netherlands, Norway, Poland, Portugal, Sweden, Slovenia, Switzerland, United Kingdom. The following countries did not participate in at least one ESS wave and are excluded from the analysis: Albania, Austria, Bulgaria, Croatia, Cyprus, Czech Republic, Estonia, Greece, Iceland, Israel, Italy, Kosovo, Latvia, Lithuania, Luxembourg, Romania, Russia, Slovakia, Turkey and Ukraine.

For each country and year, we indicate the number of observations.

— Table 4.1 about here —

Political preferences are measured along two dimensions. First, survey participants are asked about their views on immigration. We are particularly interested in the question focusing on migrants from poorer countries outside Europe. This is chosen to reflect low-skill immigration, including refugees and asylum seekers. The second question concerns government redistribution. Each survey participant is asked whether the government should reduce differences in income levels. For both questions, the participants are provided a list of possible answers. These are shown in Table 4.2. In addition, we indicate the total number of respondents choosing each possible answer.

— Table 4.2 about here —

In order to illustrate the distribution of policy preferences, we plot the total number of observations for each cell in Table 4.2. These statistics are based on the pooled data set for all periods from 2002 to 2012. Note that we define both immigration (L) and redistribution (z) as ranging from 1 to 4, with lower values indicating support and high values rejection. For example, the cell ($z = 1$ and $L = 1$) represents individuals who agree strongly with redistribution and would like to allow many poor immigrants to enter the country.

We observe that about two thirds of survey respondents agree or agree strongly with income redistribution (44.43 + 26.52 percent). In contrast, less than fifteen percent disagree with equalizing the income distribution. With respect to the question, how many poor non-European immigrants should be allowed to enter the country, we observe that two in three survey respondents want 'some' or 'few' immigrants. About thirteen percent are completely hostile towards such immigration.

In the econometric specification we can exploit the fact that each survey participant is asked a large array of additional questions covering education, work, income, and many further individual characteristics. In Table 4.3, we provide summary statistics for each variable employed in the empirical analysis.

— Table 4.3 about here —

The table shows that, on average, survey participants are 47 years old.[1] Roughly half of them are male, 26 percent have a higher education, six percent are on welfare benefits, 26 percent are retired, and 56 percent earn a positive wage. We enrich the ESS data by several macroeconomic variables, taken from the OECD, World Bank, and UNU WIDER. In the sixteen countries of our sample, during the time period 2002–2012, the unemployment rate was on average about eight percent. The share of foreign-born population was seven percent and the Gini coefficient was 40 percent.[2] Trends in these macroeconomic variables are discussed in the next section.

4.3 Explaining Trends in Policy Preferences

In this section, we use our data set to show that, in fact, political preferences with respect to redistribution and immigration changed between 2002 and 2012. Moreover, we discuss which macroeconomic variables could explain such a change. Several econometric exercises explore whether education, unemployment, and the share of foreigners can account for altered political preferences.

4.3.1 Time Trends in Policy Preferences

Using biannual data from the European Social Survey (ESS) on sixteen countries over the period 2002–2012, we investigate voter preferences over redistribution and low-skill immigration. Figure 4.1 shows aggregate trends for both policy dimensions. The plot is based on data for all sixteen countries that participated in every wave. We compute the difference between the share of individuals choosing a particular answer —either on immigration or redistribution— in 2012 and 2002.[3]

[1] We deleted one observation where the recorded age was 123 years. Our findings are not sensitive to this deletion.

[2] The fact that there are some missing values in our data set results from lack of available data. This is also the reason why we use the total foreign-born population instead of relying on even less available data on foreign-born populations by country of origin.

[3] In Figures 4.17 and 4.18 in the Appendix we plot the share of survey participants selecting each possible answer for both questions in each year. Moreover, in Figures 4.19 and 4.20 in

— Figure 4.1 about here —

We find a general trend towards more demand for redistribution. In particular, the share of survey participants who agree strongly with government redistribution increased by 6.7 percentage points (or 29.3 percent). This is notably different from trends observed in the United States [Kuziemko et al., 2015]. Over the same time period, however, we find a polarization in attitudes towards low-skill immigration with a significant increase in the share of individuals heavily opposed to immigration. A striking finding in the ESS data is the large increase (+4.9 p.p. or 50.6%) in the share of individuals who answer 'none' when asked how many poor people from outside should be allowed to enter their country. The observed polarization supports the hypothesis that after an economic crisis —in our case the 2008 financial crisis followed by the Euro crisis— ideological preferences become more polarized [Mian, Sufi and Trebbi, 2014].

4.3.2 Macroeconomic Factors and Policy Preferences

Having documented the shift in political preferences, we now investigate why these shifts occurred. In particular, we examine whether the observed shifts in policy preferences can be reconciled with aggregate trends in macroeconomic variables.[1] To identify crucial macroeconomic factors, we take into account findings of prior work in the literature on the determinants of political preferences. A large number of studies has investigated why people support or oppose redistribution and immigration. Several papers conclude that economic motives influence policy preferences with respect to income redistribution. However, non-economic factors have also been shown to play a significant role [Corneo and Grüner, 2002].

Only a few papers consider both immigration and redistribution simultaneously. Facchini and Mayda [2008] document that in most countries only a small minority favors open immigration. The authors investigate which factors shape

the Appendix we plot the share of survey participants selecting each possible answer to the immigration and redistribution question for each country separately.

[1]We are not the first to link time trends in public opinion to macroeconomic conditions. Wilkes and Corrigall-Brown [2011] investigate twenty years of data from Canada and find that individuals change their mind in response to altered economic conditions.

voters preferences and find that economic concerns such as labor market competition play a key role.[1] In a subsequent paper, Facchini and Mayda [2009] address the question how the welfare state affects attitudes towards immigration. Using data on 18 high-income countries from 1995, the authors find that in the presence of low-skill immigration, income is negatively correlated with support for immigration while skill is positively correlated. In countries with skilled immigration, the relationships are reversed.

Education — A general finding in the literature is that low-educated individuals are more likely to oppose low-skill immigration [Hainmueller and Hiscox, 2007, 2010; Hatton, 2014]. Usually this is driven by fears of competition in the labor market. This public fear has not been reduced by research showing that migrants typically are a weak substitute for natives [Card, 2009]. However, low-educated natives could also oppose low-skill immigration due to the expected fiscal costs [Scheve and Slaughter, 2001]. Another explanation for why anti-immigration sentiments are more prevalent among low-income natives is that poor immigrants usually reside in the neighborhood of low-educated natives [Halla, Wagner and Zweimüller, 2015]. Finally, several prior studies suggest that higher education affects values and thus the way natives view immigrants [Hainmueller and Hiscox, 2007].[2]

Economically, highly educated natives may benefit from low-skill immigration since their skills are complementary. Moreover, they have more financial means to support poor immigrants. However, they also pay the lion share of taxes and low-skill immigration comes at a fiscal cost. In terms of social values, it is often hypothesized that education positively affects individuals attitudes towards foreigners. This may explain why even in the absence of wage or employment effects —as suggested by Hatton [2014]— highly educated natives favor low-skill

[1] They also examine how attitudes translate into policy outcomes. Given the tiny fraction of supporters, the authors argue it is puzzling to find any immigration at all. Facchini and Mayda use a model with interest groups to explain this observation.

[2] In addition, Bechtel, Hainmueller and Margalit [2014] use survey data from German voters to explore determinants of support for *international* redistribution. In the wake of the European bailout program, the authors find that individuals' own economic status provides limited explanatory power in comparison with social attitudes such as altruism and cosmopolitanism.

immigration. In our model, we argue that altruism could provide an explanation for observed policy preferences among individuals with high incomes. Supporting poor immigrants might be motivated by a range of social considerations, including social pressure, guilt, sympathy, or a simple warm glow [Andreoni, 1989, 1990]. However, doing so comes at a twofold cost: the government faces additional expenses and either has to increase taxes (on the rich) or cut welfare spending for natives.

Unemployment Rate — The time window of our analysis, from 2002 to 2012, covers both the financial and the Euro crisis. Hence, the rise in unemployment rates is a natural candidate to explain changes in political preferences, especially for the increasing demand for redistribution [Cusack, Iversen and Rehm, 2006]. Most of the literature concludes that support for redistribution is decreasing with individual income. However, there are also studies which explain why wealthy individuals actually support the welfare state. Piven and Cloward [1971], for instance, argue that support for redistribution may arise from the idea that it prevents crime and other forms of social unrest.[1]

Share of Foreign-Born Population — Several studies suggest that a high share of foreign-born population affects natives' attitudes towards both immigration and redistribution. Razin, Sadka and Swagel [2002*a*] discuss the explanatory power of the standard theory on taxation and redistribution. This theory predicts a positive correlation between pre-tax income inequality and the amount of redistribution. They add to this model a fiscal leakage from native-born individuals to immigrants. With this modification, low-skill immigration can lead to less redistribution even though migrants would join the pro-tax coalition. The reason for this is that immigration might shift the general attitude of natives against taxation because a larger fraction of transfers ends up in the pockets of immigrants. This shift can be larger than the effect of migrants voting for high taxes. Razin, Sadka and Swagel use data from eleven European countries from 1974-1992 to

[1]Furthermore, religious beliefs and social norms have also been found to explain voting behavior. In our theoretical model, we follow this idea and assume income redistribution to have a general benefit similar to public goods.

support their model's predictions.

Similarly, Luttmer [2001] argues that attitudes toward welfare spending are driven by interpersonal preferences. In particular, he argues that there is a negative exposure effect: individuals lower their support for redistribution if the welfare recipiency rate in their community increases. Moreover, there is evidence of a racial group identity: support for redistribution increases in the share of local recipients from their own racial group. This result is supported empirically by Dahlberg, Edmark and Lundqvist [2012]. The authors find that increased low-skill immigration reduced support for redistribution in Sweden, especially among high-income natives.

Low-skill immigration, however, has not only been found to affect attitudes towards redistribution but also towards migrants themselves. According to a study by Halla, Wagner and Zweimüller [2015], the residential proximity to poor immigrants increases the support for right-wing parties which reject immigration. Conditional on education (which raises pro-immigration attitudes), income has been found to reduce support for immigration. This might reflect concerns about high tax rates as a result of immigrants' welfare dependency. Hatton [2014] also points out that anti-immigration sentiments are often diffuse while support is usually concentrated.

Macroeconomic Trends in Europe — Based on the discussion of the literature, in our main analysis we focus on three explanatory variables: (i) higher education, (ii) unemployment rate, and (iii) share of foreign population. For each variable, we show time trends between 2002 and 2012 in Figure 4.2.

— Figure 4.2 about here —

The data is based on the sixteen countries which participate in each wave of the ESS between 2002 and 2012. We observe a monotone, positive trend in both the share of individuals with higher education as well as the fraction of foreign population. The numbers increase from 21.3 to 32.6% and from 5.9 to 8.2%, respectively. For the unemployment rate, we observe a decrease prior to the 2007/08 financial crisis but a steep increase afterwards. In 2012, the average unemployment rate was 9.6 percent compared to its 2002 level of 7.0 percent. To link these macroeconomic trends to political views, we present a theoretical model in Section 4.4 which shows how individual characteristics and macroeconomic variables affect voter preferences over redistribution and immigration.

4.3.3 Empirical Analysis

Before turning to our theoretical model, we first investigate whether observed time trends in education, unemployment rates, and the share of foreigners are empirically associated with trends in policy preferences. For the sake of brevity our analysis here is focused on two observations from Figure 4.1. First, we want to explain the increased opposition to immigration. This is measured by the fraction of survey participants choosing 'none' when being asked how many poor non-European people should be allowed to immigrate. Second, we test whether rising support for income redistribution is a result of trends in education, unemployment rates, and the share of foreigners.

4.3.3.1 Econometric Approaches

We use several econometric approaches to investigate the relationship between macroeconomic variables and individual political preferences. First, a pooled regression model sheds some light on correlations in the data. Second, we explore whether a country's time trends in macroeconomic variables correlates with its trend in policy preferences. Third, we use the full set of answers to the survey questions and run a multinomial logit regression.

Pooled Regressions — In order to explore the determinants of policy preferences, we first run several pooled regressions. For each individual i located in country c who participated in the ESS survey at time t, we know whether he or she revealed a certain policy preferences $\text{POLPREF}_{i,c,t}$. We fit the linear model

$$\begin{aligned}\text{POLPREF}_{i,c,t} =& \alpha_c + \gamma_1 I(HighEdu_{i,c,t}) + \gamma_2 I(Unemployed_{i,c,t}) \\ &+ \delta \mathbf{X}_{i,c,t} + \mu_1 \mathrm{u}_{c,t} + \mu_2 f_{c,t} + \varepsilon_{i,c,t}\end{aligned} \tag{4.1}$$

where $I(HighEdu_{i,c,t})$ and $I(Unemployed_{i,c,t})$ indicate whether individual i has a higher education and is unemployed, respectively. Other individual characteristics such as age, gender, or being retired are summarized by $\mathbf{X}_i$. Finally, $u_{c,t}$ and $f_{c,t}$ denote country c's unemployment rate and share of foreigners, respectively. Note that we add country-fixed effects because we are interested in within-country time trends, not cross-country differences. The standard error $\varepsilon_{i,c,t}$ is clustered at the country level.

Trend Correlations — The two most striking findings in Figure 4.1 are the surge in support for redistribution and in heavy opposition to immigration. For both policies, we estimate the time trend in each country:

$$\mathrm{SP}_{c,t} = \alpha_c + \beta_c \mathrm{YEAR}_t + \varepsilon_{c,t} \tag{4.2}$$

where $\mathrm{SP}_{c,t}$ refers to the share of people in country c at time t who strongly oppose immigration (or strongly agree with redistribution). In each country, we also obtain a time trend for both macroeconomic variables by fitting the linear regression

$$\mathrm{MACROVAR}_{c,t} = \gamma_c + \delta_c \mathrm{YEAR}_t + \mu_{c,t} \tag{4.3}$$

where $\mathrm{MACROVAR}_{c,t}$ is either the unemployment rate or the stock of foreigners. In the Figures shown below, we then plot δ_c against β_c. This serves as suggestive evidence for the explanatory power of each macro variable with respect to observed time trends in policy preferences.

Flexible Estimation — In addition to the aforementioned analyses, we estimate a flexible model using a multinomial logit estimator. We use as a dependent variable all of the sixteen possible answer combinations in the redistribution–immigration–space. Omitting the subscripts emphasizing that it is unique to each policy mix, we fit the following equation:[1]

$$\begin{aligned} V = & I(HighEdu)\left[\beta_{H,t}t + \beta_{H,f}f + \beta_{H,u}u\right] + \\ & I(LowEdu)\left[\beta_{L,t}t + \beta_{L,f}f + \beta_{L,u}u\right] + \\ & I(Retired)\left[\beta_{R,t}t + \beta_{R,f}f + \beta_{R,u}u\right] + \\ & \beta_{Minority}I(Minority) + \sum \beta_{C=c}I(C=c) + e \end{aligned}$$

where $I(\cdot)$ is an indicator function taking the value one if the survey participant has a high education ($HighEdu$), low education ($LowEdu$), is retired ($Retired$), or belongs to an ethnic minority ($Minority$). Years are denoted by t and $I(C = c)$ are country-fixed effects. We estimate separate coefficients for each outcome using a total of 153,128 observations. Standard errors are clustered at the country level.

4.3.3.2 Findings

In a first step, we run pooled OLS regressions described in equation (4.1). The results shown in Table 4.4 suggest that higher educated people are less likely to support redistribution, more likely to support low-skill immigration and less likely to heavily oppose immigration. This confirms prior research in the literature [Hainmueller and Hiscox, 2007].

— Table 4.4 about here —

We also find that older survey participants support redistribution but oppose low-skill immigration, confirming prior results in the literature [Dotti, 2016; O'Rourke and Sinnott, 2006]. Being a welfare recipient is associated with similar political preferences. Finally, at the country level we find that support for redistribution increases if the unemployment rate is higher [Cusack, Iversen and Rehm, 2006]. Overall, we do not interpret these findings as evidence of causal relationships but as correlations which motivate further research. In addition, it strengthens our confidence in the survey data from the ESS as the results confirm

[1] The error term e is assumed to satisfy the usual requirements of multinomial logit model, i.e. the independence of irrelevant alternatives and having a Gumbel distribution. Furthermore, one of the equations must be normalized to zero.

several patterns documented in the related literature. We continue by investigating in more detail how education, unemployment, and the share of foreigners affect policy preferences.

Education and Policy Preferences — We investigate the impact of education on how individuals answer the two questions on redistribution and immigration in the ESS survey. Figure 4.3 shows how clearly education affects preferences along the two policy dimensions.[1] Among those survey participants that disagree with redistribution but want to allow many immigrants, almost fifty percent have a high education. In contrast, 85 percent of people opposed to any immigration but strongly in favor of redistribution have a low education.

— Figure 4.3 about here —

The pattern revealed in Figure 4.3 is remarkably stable over time. In Figure 4.4, we separate the 2002 and 2012 wave. The plots suggests a substantial stability of the education-preferences nexus. We also find the same pattern when using only data from a single country.

— Figure 4.4 about here —

This implies that education has a lot of predictive individual policy policy preferences. The share of highly educated is highest in the top-left corner, opposing redistribution but supporting low-skill immigration. In contrast, very few people with high education choose a 'populist' right-wing policy mix in favor of redistribution but against immigration. The fact that highly educated are less likely to support redistribution is not surprising and in line with previous cross-country evidence provided by Guillaud [2013].

Unemployment and Policy Preferences — One of the major macroeconomic trends in Europe between 2002 and 2012 was the surge in unemployment rates as a consequence of both the financial and the Euro crisis. Using our data

[1] For a comparison, we show the same figures for individuals with low and high *incomes* in Figure 4.16 in the Appendix.

set, we test whether a positive trend in unemployment rates is associated with (i) increasing opposition of immigration and (ii) support for redistribution.

— Figure 4.5 about here —

We observe a positive gradient when plotting a country's time trend in unemployment against its trend in the share of people opposed to immigration. This suggests that increasing unemployment rates in Europe could be one reason for the recent anti-immigration movements. With respect to support for redistribution (Figure 4.5, right-hand side), we see a positive correlation between trends in unemployment and support for redistribution. Using the detailed survey responses from the ESS, we explore these correlations using a multinomial logit regression. In order to illustrate the results, in Figure 4.6 we plot marginal effects at mean values of all other variables. We do this separately for high- and low-educated individuals.

— Figure 4.6 about here —

The results indicate that both educational groups are more in favor of redistributing income if the unemployment rate increases. Moreover, there is some shift in preferences over immigration. When unemployment rates increase, opposition to low-skill immigration becomes more prevalent. Overall, less people favor a 'moderate' political position (some immigration, agree with redistribution). In line with findings by Mian, Sufi and Trebbi [2014], we observe that political preferences become more diverse and extreme.

Share of Foreigners and Policy Preferences — In the next step, we test whether an increasing share of foreigners in a country's population is correlated with (i) increasing opposition to immigration and (ii) support for redistribution. Figure 4.7 suggests that anti-immigration sentiments increased in those countries that have had a positive trend in the share of foreign population (left-hand side).

— Figure 4.7 about here —

With respect to redistribution, we do not observe a clear correlation (right-hand side). The small gradient is largely driven by some outliers, including Ireland

(IE). Except for these, if anything, we see a weak positive correlation. This would imply that a larger fraction of foreigners in the population is associated with less support for redistribution. Concerns about 'fiscal leakage' could be one explanation behind this correlation [Facchini and Mayda, 2009].

As before, we can also use the detailed survey data provided by the ESS and estimate the impact of increasing the share of foreign-born individuals in the population on preferred policy combinations. To illustrate the multitude of regression estimates, we plot marginal effects at mean values of covariates in Figure 4.8.

— Figure 4.8 about here —

In line with Luttmer [2001], we find that both high- and low-educated individuals become more hostile towards redistribution when the stock of foreigners increases. Furthermore, opposition to immigration increases among low-educated individuals which is similar to findings by Halla, Wagner and Zweimüller [2015].

Concluding the Empirical Findings — Overall, we find evidence that individual characteristics as well as macroeconomic variables are in fact correlated with political preferences. We can confirm three central patterns and mechanisms suggested by the literature. First, education has strong predictive power for preferences over redistribution and immigration. The more educated an individual, the more likely he is to support immigration but oppose redistribution. Second, we find that rising rates of unemployment are associated with more opposition to low-skill immigration and more support for redistribution. Our third finding is that increasing the share of foreign-born population leads highly educated individuals to reject the welfare state and low-educated individuals to oppose immigration.

While these findings are not new, they raise several questions. If labor market effects of immigration on natives' wages and employment are negligible (e.g., Ottaviano and Peri, 2012), why does anyone support low-skill immigration which comes at a significant fiscal cost? And why would rational voters reject immigration despite this empirical evidence? Much of the literature assumes some form of xenophobia in order to explain why a fraction of the native population rejects immigration. However, Dancygier and Donnelley [2014] find that xenophobia has not risen since 2002 in Europe. Hence, could it not just be economic constraints that explain the surge in anti-immigration attitudes? The purpose of our theoretical model is to answer these questions, testing how much of the trends in policy preferences observed in Europe we can explain using solely economic rationales.

4.4 Theoretical Model

In order to understand how unemployment rates, shares of foreign-born population and the average education level affect preferences over immigration and redistribution, we develop a new model featuring both policy dimensions. The setup of our model is intended to follow the related literature. We model preferences split up by educational status and focus primarily on how the government budget ties together the different policies. The government redistributes income from all employed individuals, both high and low skilled, to the unemployed as well as a group of permanently poor. Consistent with prior empirical research on the fiscal effects of low-skill immigration, we assume poor immigrants to be recipients of transfers. We also take into account a large body of empirical research on the labor market effects of immigration [Hatton, 2014; Manacorda, Manning and Wadsworth, 2012; Ottaviano and Peri, 2012]. The general conclusion from this literature is that labor market effects on natives are small and largely negligible. In our model, we thus abstract from any impact on wages or employment of natives.[1]

The absence of labor market effects, however, necessitates a new explanation for why some people support low-skill immigration. If highly-educated individuals do not benefit from migrants providing labor that is complementary to their labor, why would they support such immigration? In addition, if low-skill natives are not xenophobic and do not face disadvantages in the labor market resulting from immigration, why do some of them reject immigration? To answer these

[1] A detailed discussion of how such price effects would alter our results is presented in Section 4.5.

questions, it is important that our model features both policy dimensions, immigration and income redistribution. In a first step, we assume that an individual's utility is increasing in his own consumption. Given an exogenous risk of being unemployed, the individual will then prefer at least some redistribution as a hedge against losses to personal income. In a second step, we allow individuals to derive utility directly from immigration and from income redistribution. Hence, we allow voters to be motivated partly by altruism towards poor natives as well as (potential) immigrants. This is based on the idea that income redistribution through the government can be considered a public good from which all individuals benefit. One motivation for this concept is a reduction in crime.[1] We then argue that consumption is weakly decreasing in immigration through the negative fiscal effect. Any given level of transfers to the poor foreigners requires a higher tax rate in the model.

In what follows, we first explain the setup of the model in more detail. Thereafter, we estimate the preference parameters in Section 4.4.2. We use these estimates and simulate the model to show static preferences in Section 4.4.3. Then, in Section 4.4.4 we illustrate how changing macroeconomic variables alters political preferences in our model. Finally, Section 4.4.5 provides an out-of-sample prediction for aggregate preferences in the hypothetical case where Europe returns to a low unemployment rate but experiences continuing increases in education and foreign-born population.

4.4.1 Setup

Consider an economy populated by three groups: high- and low-skilled wage earners as well as welfare recipients (henceforth referred to as poor). The poor earn a small non-taxed income of $\phi^2 + z$ each period, where ϕ^2 is income with $\phi \in (0, 1)$ and z is a non-negative transfer from the government. There are two types of poor: Some individuals are permanently poor and will always remain in this group. The second group are those high- and low-skilled individuals who

[1] It is worth noting that we do not model the utility function such that individual i's happiness increases simply with other individuals' income. This would potentially violate empirical findings by Luttmer [2005].

experienced a negative employment shock. Following Alesina and La Ferrara [2005], wage earners face a probability $u > 0$ of being unemployed and joining the group of welfare recipients. In contrast, with probability $(1-u)$ this negative shock does not occur and they earn an income of $1-\tau$ and $\phi-\tau$, respectively. We use τ to denote a lump sum tax imposed by the government. The government does nothing but redistribute income from wage earners to the poor. It must balance its budget each period. The budget constraint is given by

$$(1-u)(N_h+N_l)\tau = z\left(u(N_h+N_l)+N_p\right) \tag{4.4}$$

where N_h, N_l, and N_p denote the number of high- and low-skilled natives as well as the permanently poor in the economy, respectively. For simplicity, we assume that there are no natives which permanently receiving welfare benefits. Hence, let the number of permanently poor be given by the sum of foreign-born immigrants from poor countries $N_p = L + F$ where L is the number of admitted immigrants and F is the number of already migrated foreign-born individuals.[1] It is crucial to emphasize that we focus exclusively on the immigration of poor people. First and foremost this is to match the respective question from the European Social Survey. Participants are asked how many *poor* people from outside Europe should be allowed to enter their country. Implicitly, the focus on poor immigrants implies that each migrant comes at a fiscal cost. The assumption that there is a net government loss is supported by prior research.[2] Letting lower case letters denote shares of the total population, we can re-write equation (4.4) to get

$$\tau = z\frac{u(1-l-f)+l+f}{(1-u)(1-l-f)} \tag{4.5}$$

where $[1-l-f]$ is the fraction of natives in the population. Note that equa-

[1]Note that while native voters cannot determine F, the number of additional poor immigrants, L, is a choice variable.

[2]Storesletten [2003] finds that in a European welfare state, an average immigrant represents a total net loss to the government of about 20,000 US dollar. In contrast, the estimates are generally positive for the United States where there is a much smaller welfare state [Storesletten, 2000]. Recent work by Bratsberg, Raaum and Røed [2014] supports the assumption of poor labor market outcomes among migrants from low-income countries.

tion (4.5) pins down the relationship between taxes, transfers, and population shares. To satisfy the incentive constraint, it must further hold that transfers are such that disposable income for low skilled is at least as high as income of the poor:[1]

$$z \leq \phi(1-\phi)(1-u)(1-l-f). \tag{4.6}$$

Following empirical studies by Ottaviano and Peri [2012], Dustmann, Frattini and Preston [2013], Docquier, Ozden and Peri [2013], or Foged and Peri [2016] we assume that labor market effects on employment or wages of natives are small enough that we can neglect them. This raises the question why anyone in the model would support low-skill immigration. Following Olsen [1965], Becker [1974] and Andreoni [1990], we assume that individual i's utility does not only depend on his own consumption. Instead, each individual also cares about the well-being of others. This altruism, however, is not limited to natives but extends to foreigners as well. The way we model altruism—like prior work by DellaVigna, List and Malmendier [2012]—allows for pure altruism (i.e., caring about the total amount of charity), impure altruism (i.e., warm-glow motives), and prestige [Harbaugh, 1998]. In principle, we also allow for spiteful behavior or negative social preferences [Levine, 1998]. Moreover, irrespective of whether an individual receives welfare benefits himself, we assume him to benefit from the presence of a welfare state. This is motivated by prior research by, for example, Piven and Cloward [1971]. In particular, support for redistribution may arise from the idea that it prevents crime and other forms of social unrest.

We focus our analysis on the political preferences of natives. Let utility be given by $V(c, z, l)$ where c denotes private consumption and we allow z and l to enter as separate arguments. We will not take a stand on the exact functional form of the utility function, but rely instead on Taylor approximations. A first-order approximation to utility around the policy with no redistribution nor immigration

[1]As after tax income of the low-skilled must be at least as high as the after-transfer income of the poor, we have $\phi - \tau \geq \phi^2 + z$. Using (4.4) to substitute for τ and solving for z gives (4.6).

(i.e., $z = l = 0$) is given by

$$V(c, z, l) = V(w, 0, 0) + V_c(c - w) + V_z z + V_l l + e \tag{4.7}$$

where w is disposable income before any transfers or taxes and e captures higher-order terms as well as country differences in marginal utility.[1] Note that we follow this approach of using a first-order approximation around the $(z = l = 0)$ policy combination to match the answer in the ESS questionnaire.[2] Using this first order approximation, expected utility for an individual of type $i \in \{h, l\}$ is thus

$$\begin{aligned} EV_i = (1+u)\,&[V(w_i,0,0) - V_c(w_i,0,0)\tau + V_z(w_i,0,0)z + V_l(w_i,0,0)l] \\ &+ u\,[V(w_p,0,0) + V_c(w_p,0,0)z + V_z(w_p,0,0)z + V_l(w_p,0,0)l] \\ &+ (1+u)e_{i,1-u} + ue_{i,u}. \end{aligned}$$

In the equation above, w_p is disposable income for the poor, and $(1+u)e_{i,1-u} + ue_{i,u}$ is a linear combination of the higher order terms of the first order approximations to the utility functions in the employed and unemployed states, respectively. Let $\beta_{k,j}$ denote marginal utility from k (i.e., immigration) in state j relative to marginal utility of consumption in the employed state, $V_c(w_i, 0, 0)$. We can then transform the equation above, by dividing by $V_c(w_i, 0, 0)$ and subtracting the constant terms:

$$\tilde{EV} = (1+u)\,[\tau + \beta_{z,1-u}z + \beta_{l,1-u}l] + u\,[(\beta_{c,u} + \beta_{z,u})\,z + \beta_{l,u}l] + \tilde{e} \tag{4.8}$$

where $\tilde{e}$ is a positive linear transformation of the higher-order terms e_u and e_{1-u}. We then obtain the convenient property that expected utility in the given

[1] In a study by Eugster et al. [2011], the authors discuss differences in demand for social insurance across countries in Europe.

[2] Survey participants in the European Social Survey (ESS) can respond 'none' and 'disagree' when asked how many poor immigrants should be allowed to enter their country and how they think about income redistribution, respectively.

no-redistribution and and no-immigration state is zero:

$$\tilde{EV}(z = 0, L = 0) = 0 \tag{4.9}$$

Explaining the Model Dynamics — Before we simulate the model, we first discuss its mechanics in more detail. First of all, it is important to note that the tax rate required to finance a given level of transfers (z) is increasing in the level of immigration (l), the share of foreigners (f), and the unemployment rate (u):

$$\frac{\partial \tau}{\partial u} = \frac{z}{1-u}\left(1 + \frac{u(1-l-f)+l+f}{(1-u)(1-l-f)}\right) > 0 \tag{4.10}$$

$$\frac{\partial \tau}{\partial l} = \frac{\partial \tau}{\partial f} = \frac{z}{1-l-f}\left(1 + \frac{u(1-l-f)+l+f}{(1-u)(1-l-f)}\right) > 0 \tag{4.11}$$

Second, we see that the maximum feasible transfer, defined as the transfer that makes the incentive constraint (equation 4.6) hold with equality, is decreasing in both the unemployment rate and the share of foreigners:

$$\frac{\partial z^{max}}{\partial u} = -\phi(1-\phi)(1-l-f) < 0 \tag{4.12}$$

$$\frac{\partial z^{max}}{\partial f} = \frac{\partial z^{max}}{\partial l} = -\phi(1-\phi)(1-u) < 0. \tag{4.13}$$

In other words, we see that both unemployment and a high share of foreign-born population have two effects on fiscal policy. First, the tax rate required to finance a given level of transfers increases. This makes income redistribution more costly for tax payers in the economy. Second, the policy space shrinks in the sense that the maximum feasible transfer decreases.

4.4.2 Simulation of the Model

The model can be simulated by transforming it into a standard discrete choice framework. In a first step, let utility be given by the $\tilde{EV}_i(z, L)$, as defined by equation (4.8), and a stochastic term ε:

$$U_i(z, L) = \tilde{EV}_i(z, L) + \varepsilon \tag{4.14}$$

where ε has a Gumbel distribution. This implies that the probability of an individual preferring a particular policy is given by

$$P_i(z = z^*, L = L^*) = \frac{exp(\tilde{EV}_i(z = z^*, L = L^*))}{\sum\limits_{z,L} exp(\tilde{EV}_i(z, L))} \tag{4.15}$$

where the numerator is the exponentiated, deterministic part of utility for a given outcome, and the denominator is the sum of these exponentiated deterministic utilities over all possible policy combinations. Following Berry [1994], equation (4.15) can be linearized by taking logs:

$$\ln P_i = \tilde{EV}_i(z = z^*, l = L^*) - ln \sum exp(EV_i). \tag{4.16}$$

As the deterministic part of utility in the $(z = 0, L = 0)$ policy case is zero, we use $\ln P(0, 0) \equiv -ln \sum exp(\tilde{EV})$ as the base outcome. We can then replace the last term on the right hand side of equation (4.16) by the log of the choice probability of policy $(z = 0, L = 0)$:

$$\ln P_i = \tilde{EV}_i(z = Z, L = l) + \ln P(z = 0, L = 0). \tag{4.17}$$

The log of the choice probability for a given outcome is the deterministic part of utility added to the log of the choice probability in the $(0, 0)$ policy outcome. The left hand side of (4.17) is now the log of the share of respondents reporting that they prefer policy mix $(z = Z, l = L)$. We must take into account the restrictions implied by this setup, in particular that the log of the choice probability in the $(z = 0, L = 0)$ as well as marginal utility of consumption enters with coefficients equal to unity. To take these restrictions into account, we define a transformed dependent variable as the log of the choice probability, subtracted log of the choice probability in the $(z = 0, L = 0)$ state, and subtracted the difference in consumption relative to the base outcome:

$$\Psi \equiv \ln P - lnP(0,0) + \tau(1-u) \tag{4.18}$$

Substituting equation (4.8) into equation (4.16) and collecting terms equal to the right hand side of equation (4.18), we then have a linear expression that can be estimated:

$$\begin{aligned} \Psi_{i,c,t} =& \beta_{z,1-u_{c,t}}(1+u)z_{i,c,t} + \beta_{l,1-u}(1+u_{c,t})l_i \\ &+ \left[\beta_{c,u} + \beta_{z,u}\right] u_{c,t} z_{i,c,t} + \beta_{l,u} u_{c,t} l_i + \tilde{e}_{i,y,t} \end{aligned} \tag{4.19}$$

where i denotes each of the sixteen possible policy combinations over redistribution and immigration, c denotes the country and t is the year. Recall that $\tilde{e}$ is a positive transformation of the higher order terms of the utility function. We define the possible policy space as $l = \{0, 0.05/3, 0.05*2/3, 0.05\}$, which is the same for all countries and years.[1] The redistributive policy space is similarly defined over four points with the extremes $z_{min} = 0$ and $z_{max} = \phi(1-\phi)(1-u)(1-l-f)$.[2] We estimate this for all ESS waves for different educational levels where we have complete saturation (i.e., where all observed probabilities are strictly greater than zero). The results are provided in Table 4.5.

— Table 4.5 about here —

These parameter estimates show the marginal utility from higher z and L relative to c in the two states. The coefficients vary significantly between high- and low-skilled individuals. In particular, high-skilled people receive far less utility from z relative to low skilled types, and similarly, high skilled types receive greater utility from L relative to the low skilled and those on benefits.[3] A notable finding in Table 4.5 is that low-educated natives generally like low-skill immigration.

[1] Defining the upper limit of the policy space with respect to immigration is necessarily somewhat arbitrary. We decided to set 5% of the current population as reflecting *many* immigrants.

[2] For our specification, we assume $\phi = 0.5$. Using varying parameter values based on inequality data from UNU Wider yields similar results.

[3] This latter result is consistent with findings by Card, Dustmann and Preston [2012] who

4.4.3 Static Preferences

In the first part of our simulation of the model, we illustrate policy preferences over redistribution and immigration in the static case. For this, we assume the share of immigrants and the unemployment rate to be at their average level between 2002 and 2012.[1] We use this information and plot the probability distribution over the policy space.

— Figure 4.9 about here —

Part (a) of Figure 4.9 shows the preferences of those with high education. We observe a large mass at the bottom left corner. This implies full support for income redistribution and immigration. However, there is also a 'tail' towards less redistribution. In Part (b) of Figure 4.9, we show the distribution of probabilities for the low-skilled natives. These individuals have a strong preference for redistribution but appear more indifferent with respect to immigration. Finally, in Part (c) we show the probability ratio between high- and low-skilled individuals at each policy point.

Notably, we obtain a pattern that is very similar to Figure 4.4 which was based on ESS survey results. Among those individuals that prefer no redistribution but open immigration, the share of highly educated peaks. In contrast, primarily low-educated individuals choose a policy combination of heavy redistribution and no immigration.

4.4.4 Effects of Macroeconomic Trends

We now investigate how altered macroeconomic circumstances like higher shares of foreign-born people or unemployment affect political preferences in our model. Moreover, we investigate the impact of a compositional change in the population by increasing the share of natives with high education.

document that higher levels of immigration reduce welfare from compositional amenities to which low skilled types are more exposed.

[1] For the sixteen countries in the European Social Survey, these numbers are given by 7.2% unemployment and 7.6% foreign-born population.

Higher Share of Foreign-Born Population — Increasing the share of foreign-born citizens in our model is equivalent to increasing the share of permanently poor. This has the consequence of lowering the maximum feasible transfer (z). At the same time, it implies that the government, to balance its budget, must increase the tax required to finance a given given level of transfers.

— Figure 4.10 about here —

A simulation of our model is shown in Figure 4.10. Quantitatively, we follow real-world changes and increase the share of foreigners from 5.7% to 8.2%, as it happened in the sixteen European countries of the ESS. We observe that both high- and low-educated individuals become more opposed to income redistribution if the share of foreigners increases. However, it is important to understand that their motivations differ. In the presence of a large poor foreign population, the highly educated individuals decide to maximize their utility (equation 4.7) from charitable activity through immigration instead of redistribution. Essentially, they support free immigration while reducing the welfare state to a minimum. For the low educated, the situation is different. They also turn against the welfare state if the share of foreigners is high. This is again because transfers to the poor (z) are increasingly expensive. However, their preferences over immigration do not change.

Higher Unemployment Rate — For the second macroeconomic variable, the unemployment rate, we can also simulate the effects on voter preferences. Again, we mimic the true time trends in European countries and change the unemployment rate from 6.9% to 9.5%. This sharp increase reflects the impact of the financial and the Euro crisis.

— Figure 4.11 about here —

Figure 4.11 reveals that under these circumstances all individuals are more in favor of income redistribution. If the unemployment rate in the economy increases, both high- and low-skilled native voters become more favorable to income redistribution. However, while the low-skilled reduce their support for immigration, the high-skilled increasingly support immigration. In order to understand this difference, we can refer to equation (4.8). At times of high unemployment, both groups of natives face an increased risk of losing their job and join the group of welfare recipients. Hence, they put more weight on this outcome and show increased support for income redistribution.[1] At the same time, preferences over immigration are also affected by the economy's unemployment rate. Highly educated natives do not significantly alter their support support for immigration while low educated natives become more likely to choose a 'populist' right-wing policy combination of opposing immigration but supporting the welfare state. The latter occurs because the low-educated are facing both higher tax rates to maintain a given transfer, thus reducing consumption in the employed state, at the same time as they face a higher probability of becoming unemployed. Therefore, stuck between a rock and a hard place, the low educated choose to reduce their support for immigration.

Higher Educational Attainment — The final dynamic we consider is a compositional change in the population. The fraction of respondents in the ESS sample who possess a high level of education increased from 20% in 2002 to 28.5% in 2012.

— Figure 4.12 about here —

Not surprisingly, higher educational attainment shifts preferences towards more immigration and somewhat less redistribution. As shown in Figure 4.12, all policy combination that support immigration become more favorable. Notably, the 'populist' right-wing policy combination loses most in support.

Total Changes from 2002 to 2012 — After we simulated separately the effects of rising shares of foreigners, increasing rates of unemployment, and a larger share of highly educated individuals, we can now simulate the combined effect.

[1] Our finding is in line with a study by Alesina and La Ferrara [2005] who argue that individuals' preferences for redistribution are driven by expectations about future incomes.

In particular, we can simulate how policy preferences for native voters change if we alter several parameters at the same time. For the simulation, we again match true changes in macroeconomic variables between 2002 and 2012. Figure 4.13 indicates how the distribution of preferences changes.

— Figure 4.13 about here —

We observe that changing the macroeconomic parameters of the economy like it happened between 2002 to 2012 leads to more individuals supporting income redistribution. Overall, there is a total increase in those who strongly agree that the government should redistribute income by 2.9 percentage points. Qualitatively, the model also captures the trend of increased polarization along the immigration dimension. Both the share of respondents supporting high and very low immigration levels increase. There are less individuals with 'moderate' preferences over immigration. Underestimating quantitatively the changes is not surprising given the linear approximation to utility. In Section 4.6 of the appendix, we show that using a quadratic utility function gives similar qualitative predictions but increases the quantitative predictive power of the model.

— Figure 4.14 about here —

In Figure 4.14, we provide a comparison of the quantitative results of the simulation to observed changes in the ESS data. The latter were also shown in Figure 4.1. For the preferences over income redistribution our model yields somewhat smaller changes (2.9 percentage points) than observed in the ESS (6.7 percentage points). The predicted polarization in preferences over immigration is also smaller than what we observe in the ESS data.

4.4.5 Prediction for the Year 2024

A final question that we address in the simulation is how aggregate preferences might change in the future. This prediction is based on the assumption that the share of foreign-born population continues to increase linearly. Similarly, the average educational level continues to grow at the same rate as between 2002 and 2012. However, we assume that the unemployment rate reverts to a low level.

— Figure 4.15 about here —

Figure 4.15 shows the estimated change in preferences from 2012 to 2024 when unemployment drops from the mean 2012 level of 9.4% to 4%, and the share of highly educated as well as the share of foreign-born individuals increase to a level of 37% and 10.7%, respectively. Based on these figures, our prediction is that immigration will be a less controversial issue in the future. It will be more common to have a favorable attitude towards immigration. However, due to the reduced unemployment rate there is also a significant reduction in support for redistribution. About 11 percent of voters move towards indifference or opposition to redistribution.

In a broader picture, our model sketches the key trade-off Europe faces in the current migration debate. According to Kagan [2003], Europeans have long followed Immanuel Kant's idea of a 'perpetual peace' which is achieved through consensually agreed rules, transnational negotiation, cultural conventions and a large redistributive welfare state. In contrast, the United States is built on the concept of Thomas Hobbes with security and a liberal order depending on the possession and use of military force. If the current immigration flows from poor countries continue, European governments will have to significantly shrink welfare benefits. This is predicted by our model. Alternatively, European countries must find a way—including foreign aid or the use military power—to limit the migration flows.[1]

4.5 Alternative Explanations

In this section, we address several concerns and discuss additional findings. First, we explore differences in policy preferences across the sixteen European countries. Then we discuss the role of income inequality. Third, we provide a discussion of how labor market effects from immigration or xenophobia would alter our findings. Finally, we shed some light on how support for low-skill immigration reflects charitable activity.

[1]The trade-off is discussed in more detail in the article 'Farewell to the Era of No Fences' by Bret Stephens, published on September 7, 2015 in *The Wall Street Journal*.

4.5.1 Differences Across Countries

For the empirical analysis in this paper, we use data from sixteen European countries. An interesting question is how policy trends differ across countries or whether trends are largely similar. In Figures 4.19 and 4.20 in the appendix, we show the share of survey participants supporting any of the four answers to the question on immigration and redistribution, respectively. We plot separately the share for each year and country. The key observation is that—despite large differences in levels—the support for redistribution increased in almost every country. Simultaneously, the share of survey participants hostile towards immigration increased substantially in many countries, including Great Britain, Ireland, Portugal, Spain, or Hungary. At the country-level aggregate, however, the increase in opposition to immigration is not a universal phenomenon. Notably, we observe that in Germany opposition to poor immigration declined. In a study by Lubbers, Gijsberts and Scheepers [2002], the authors find that there are several explanations for why extreme right-wing parties gain significant support only in some European countries. The single most important factor is the right-wing parties themselves, their specific policies and leaders. In addition, public opinion on immigrants and democracy plays a relevant role as well.

4.5.2 Income Inequality

One of the most significant concerns among European voters in the past decade has been income inequality. Hence, we discuss how it affects our findings both theoretically and empirically. In the framework of our model, increasing income inequality can be reflected by lowering ϕ, which determines how much the low-skilled earn relative to the high-skilled natives. It is straightforward to see how this would affect policy preferences. With the gap between high- and low-skilled natives widening, the difference shown in Figure 4.9 would become more pronounced.

Empirically, we can also investigate whether trends in a country's income distribution are correlated with (i) increasing opposition of immigration and (ii) support for redistribution. Following a simple median voter model (e.g., Meltzer and Richard, 1981), widening income gaps could increase the support for redistribution.[1] The analysis, however, provides little evidence of a correlation between the two. Using the Q4/Q5 ratio or other measures of income inequality—including the Gini coefficient, the P90/P50, P90/P10, or P50/P10 income ratio—we find no significant correlation with the two policy outcomes. One reason for this could be the rather small variation in the inequality measures. Given that we consider a time period of only ten years and a selected set of European countries, this might not be surprising. Furthermore, experimental evidence by Kuziemko et al. [2015] shows that individuals do not in general increase their support for redistribution if they are provided with information about rising income inequality.

4.5.3 Labor Market and Price Effects

Throughout our analysis we assumed away any impact of immigration on the labor market. In particular, native workers' wages and probability of unemployment are not altered by the magnitude of immigration. Moreover, in our model high-skilled natives do not benefit from low-skill immigration through other price effects. This may include immigrants to provide household services at low cost.

We justify the absence of such effects by the fact that we focus on poor immigrants from outside Europe who often come as refugees or asylum seekers. As such they usually lack a work permit. Even if they are allowed, their labor market participation rates are fairly low. Moreover, empirical research documented the negligible wage and employment effects of immigration [Ottaviano and Peri, 2012]. One could, however, use our framework and allow low-skill immigration to have wage effects. Assuming poor immigrants to compete with low-skill natives, we could allow their wages to be depressed by competition in the labor market: $\partial\phi/\partial l < 0$. Fears of such effects would obviously reduce support for immigration among low educated natives. However, the purpose of our model is to show opposition to (and support for) immigration arising in the absence of labor market effects.[2]

[1] Theoretically this prediction becomes less clear if one takes into consideration beliefs on the role of effort and luck in shaping the income distribution [Alesina and Angeletos, 2005; Fong, 2001] or individual expectations about future incomes [Benabou and Ok, 2001].

[2] It could also be that despite the absence of labor market effects from immigration, many

Concerning price effects, one can think of low-skill immigrants reducing the costs of domestic services like gardening. In addition, foreigners might offer new varieties of non-tradeable goods and services. Our model could be extended to allow for such ramifications. However, adjusting the utility function to include diversity would be somewhat adhoc and also significantly complicate the analysis.

4.5.4 Xenophobia

In order to discuss the impact of xenophobia, we can follow prior work by Luttmer [2001] and extend our model to feature two groups of natives. First, individuals with the type of utility function we so far assumed for everyone. Second, a group of individuals who are (latent) xenophobic. These people receive no utility from immigration.[1] Moreover, the utility they receive from redistribution (as public good) is multiplied with a factor φ which is defined as

$$\varphi = \frac{\text{Natives on welfare}}{\text{Foreigners on welfare}}. \tag{4.20}$$

Thus they support the welfare state but less so if the welfare benefits are increasingly received by foreigners. In other words, with an increasing share of foreign-born population, xenophobic individuals show reduced support for redistribution. It is, however, important to note that our model can already capture the key idea by Luttmer [2001]. We do not need to assume anyone to be xenophobic in order to have some individuals turn against redistribution if the share of foreigners is high.

An alternative way to model xenophobia would be to include costs from immigration that are heterogeneous across natives. Following Card, Dustmann and Preston [2012], immigration can change the composition of the local population, reducing the utility natives receive from compositional amenities like neighborhoods, schools, or workplaces. If poor immigrants primarily move to areas pop-

natives fear such effects. However, Hainmueller, Hiscox and Margalit [2015] find no evidence that such fears affect attitudes toward immigration.

[1] It makes little sense to assume they receive utility from L and are xenophobic with respect to redistribution. An alternative setup following Benabou and Tirole [2006] would have several groups of natives with different utility from contributions to social goods.

ulated by poor natives, an increase in anti-immigration preferences among low-skilled natives is expected. Empirically, this hypothesis is supported by Halla, Wagner and Zweimüller [2015].

4.5.5 Support for Immigration as a Form of Charity

In our theoretical model, we consider each individual to be potentially altruistic. Moreover, we assume there are no price, wage, or employment effects from immigration. Hence, supporting low-skill immigration can only be the result of humanitarian considerations. Allowing poor people to enter the country and receive welfare benefits increases the utility of natives only insofar as they are altruistic. In our results, we find that being poorer —or expecting to be poorer or unemployed in the future— lowers the support for low-skill immigration. This finding suggests that altruism is increasing in income, confirming prior research by Hoffman [2011].

— Table 4.6 about here —

We can use the ESS data to further investigate whether other forms of altruistic behavior are also more prevalent among wealthier individuals. In Table 4.6, we use charitable work and money donations to humanitarian organizations as proxy variables for altruism. For both variables and varying years, we find that income is positively correlated with altruistic behavior. Furthermore, those individuals that support (reject) low-skill immigration show the same positive (negative) attitude towards charitable activities that might primarily benefit natives. Unfortunately, the questions on charitable activity vary from one ESS round to the next. Hence we cannot provide evidence of time trends in charity. Nevertheless, the estimation results shown in Table 4.6 favor our idea of considering support for low-skill immigration as a form of charitably behavior.[1] Those natives that support immigration of poor people also contribute to humanitarian organizations through charitable work and donations.

[1]Notably, these findings as well as our interpretation are supported by Poutvaara and Steinhardt [2015] who find that bitterness in life is correlated with opposition to immigration.

4.6 Conclusion

This paper addresses recent trends in European politics. In particular, we provide an economic explanation for why immigration has become a dividing concern among voters and why populist right-wing parties—supporting redistribution but heavily opposing immigration—have surged in polls. In a first step, we document two important shifts in voter preferences. Based on data from the European Social Survey on sixteen countries for the period 2002–2012, we show that there is a general trend towards more demand for income redistribution. At the same time, we observe a polarization with respect to preferences over low-skill immigration. In particular, there is a growing share of the native population that is strongly opposed to any immigration while showing support for the welfare state.

Having documented these trends, we examine potential economic explanations for these shifts. Guided by a theoretical model, we investigate how individual education as well as macroeconomic trends such as unemployment rates and the share of foreign-born population affect voter preferences. With respect to education, our model predicts that highly educated individuals are more likely to support free immigration but reject income redistribution. In contrast, low-educated individuals oppose immigration but strongly support welfare spending. These predictions find strong support in the data. In fact, for all sixteen countries and in virtually all years we find a clear pattern in the data which resembles simulated outcomes of our model.

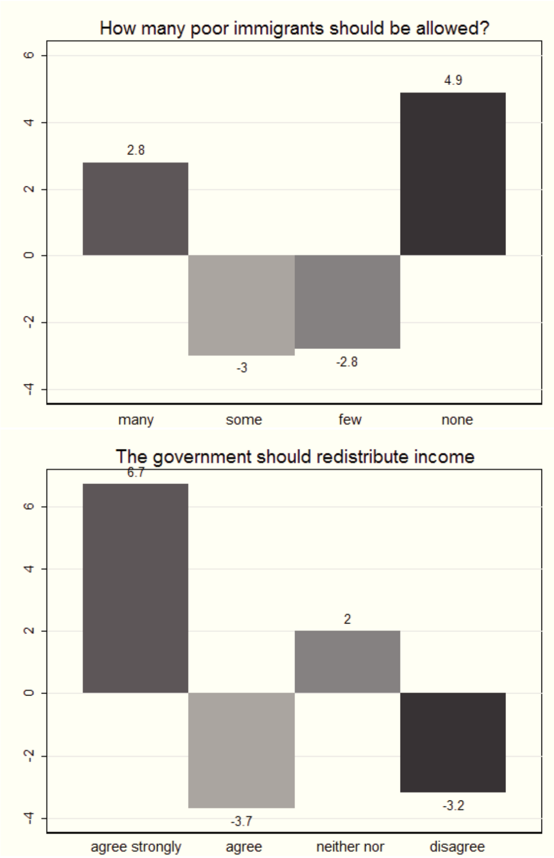
Over the period from 2002 to 2012, European countries experienced a sharp increase in the share of foreign-born population. In addition, unemployment rates rose significantly as a result of the financial crisis. We find that these macroeconomic trends are correlated with observed shifts in policy preferences. In line with our model's predictions, an increase in unemployment leads to higher demand for income redistribution. Furthermore, opposition to immigration grows among low-educated natives. An increase in the share of foreign-born population is associated with less support for redistribution as the mass of individuals eligible for transfers increase and hence taxes required to finance a given transfer will be higher. Both high- and low-skilled natives become less in favor of income redistribution. However, we simultaneously observe a polarization in their preferences over immigration, with low-skilled increasingly hostile to immigrants.

The main purpose of our study is to investigate the explanatory power of economic forces in shaping political trends observed in Europe in the past decade. Hence we abstract from any form of xenophobia and also take into account the empirical evidence that labor market effects on wages and employment of natives are negligible. Our model's assumption that support for low-skill immigration is motivated by altruism—and therefore is akin to a 'public good'—to some extent legitimizes that individuals with lower and uncertain income prioritize own needs rather than potential immigrants. In the trade literature, one can devise transfers across groups to achieve Pareto improving policies. Such policies are less feasible here, as not only natives but also immigrants will be eligible for said transfers. Hence, treating all poor in a country equally has the consequence that more poor immigrants warps the policy space. Therefore, we expect all feasible policies to involve difficult trade-offs between the well-being of different groups.

While we can confirm many patterns established in the political economy literature, we offer new insights why some natives support and some reject low-skill immigration. Our model abstracts from both xenophobia and price effects which so far have been predominantly used to explain voter preferences over immigration. It is important to note that we do not rule out the existence of xenophobic views or labor market effects from immigration. But we argue that voters do not have to be xenophobic to vote for a populist right-wing party. Xenophobia helps but it is not necessary to reject immigration. Quantitatively at least, we find that economic motives alone do not explain the full surge of anti-immigration attitudes. Hence more research should be done to uncover the additional causes of political trends in Europe.

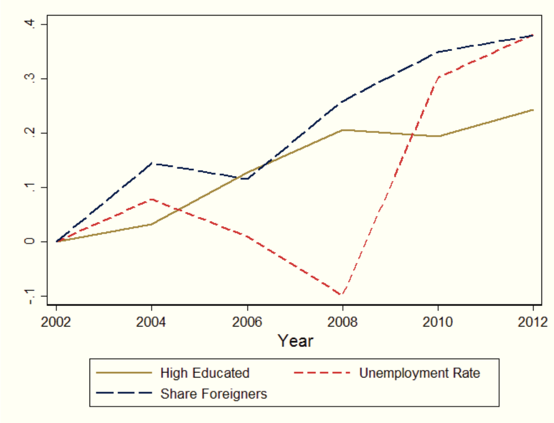
Figures and Tables

Figure 4.1: Changes in Policy Preferences between 2002 and 2012



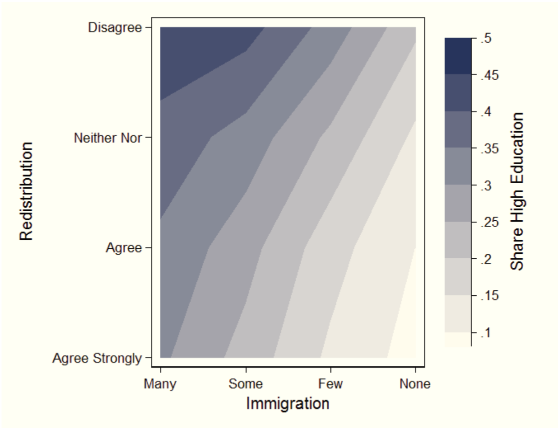
Note: The figure shows changes in the share of survey participants selecting each possible answer to (a) the immigration and (b) the redistribution question in the ESS. The scale is such that, for example, the share of individuals who 'agree strongly' with redistribution rose by 6.7 percentage points. The sample is restricted to those sixteen countries which participated in each wave. We do not report changes in respondents who responded "Do not know" or "Refuse to answer".

Figure 4.2: Macroeconomic Trends between 2002 and 2012



Note: The figure shows accumulated (percent) changes in the share of highly educated individuals, unemployment rate, and the share of foreigners. The sample is restricted to those sixteen European countries which participated in each wave of the ESS.

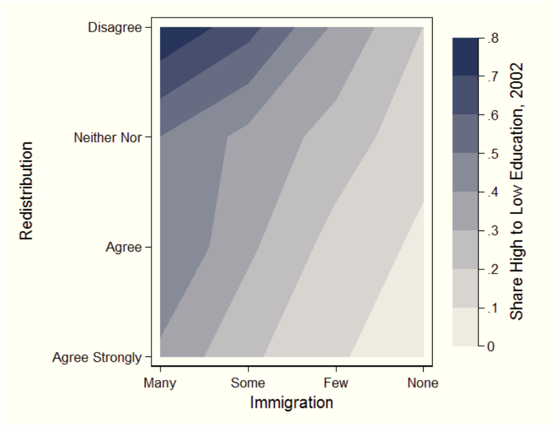
Figure 4.3: Policy Preferences by Education



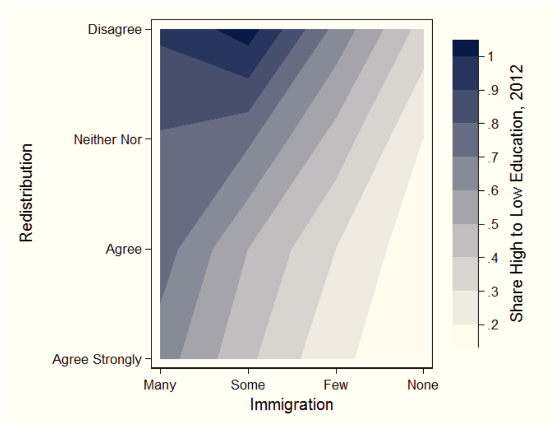
Note: The figure shows the fraction of individuals with high education for each combination of an answer to the question on immigration (x axis) and income redistribution (y axis). We use the ESS data from those sixteen countries that participated in each wave between 2002–2012 and pool the data.

Figure 4.4: High-to-Low Education Ratio Over Time

(a) 2002



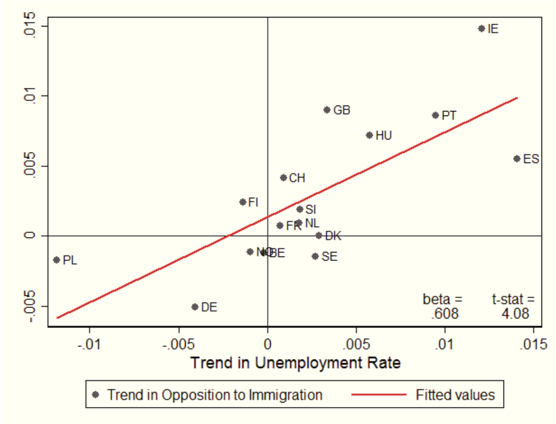
(b) 2012



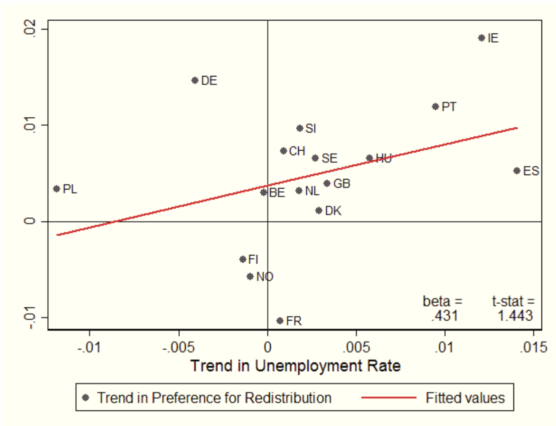
Note: The figures show the ratio of the share of highly educated individuals relative to the share of low-educated individuals in the two-dimensional map. While preferences over immigration are on the horizontal axis, views on redistribution are shown on the vertical axis. In plot (a), ESS data from 2002 are shown and in plot (b) the 2012 sample is applied.

Figure 4.5: Trends in Unemployment and Policy Preferences

(a) Opposition to Immigration



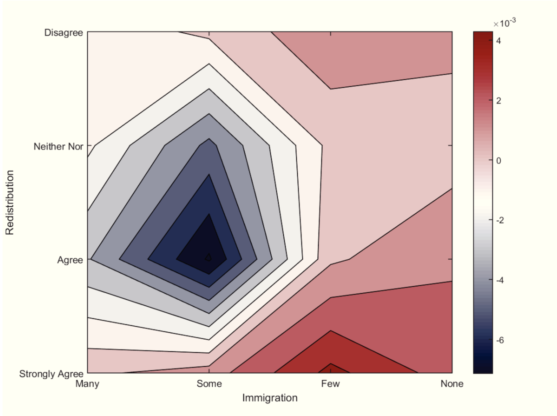
(b) Support for Redistribution



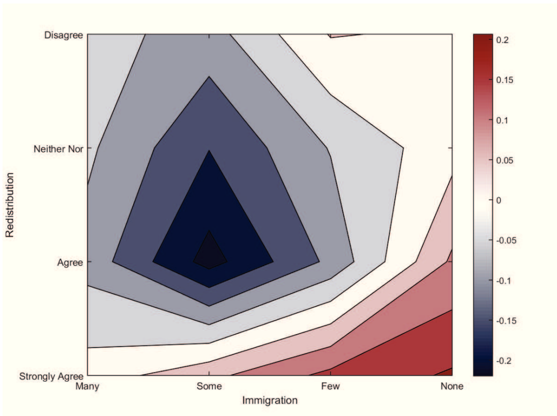
Note: The figure shows the correlation between each country's trend in unemployment (horizontal axis) and the share of its survey responds who oppose immigration (top figure) or support redistribution (lower figure).

Figure 4.6: Marginal Effects of Higher Unemployment Rates

(a) High Education



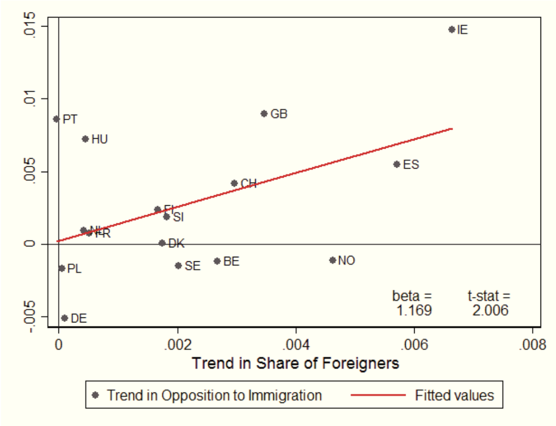
(b) Low Education



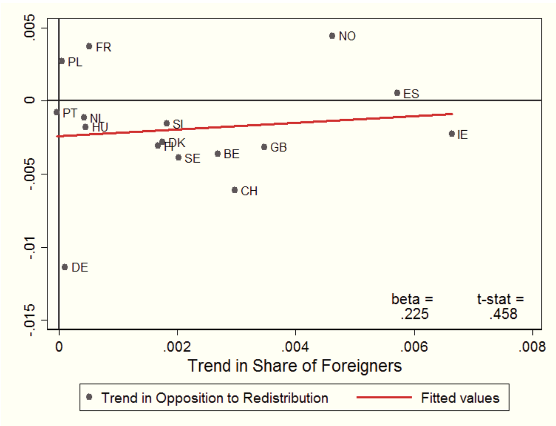
Note: The figures show the estimated marginal effects of a higher unemployment rate from a multinomial logit regression, using as a dependent variable the probability of a survey participant choosing one of the sixteen policy combinations. On the horizontal (vertical) axis, there are four possible answers to the question on immigration (redistribution). In plot (a), results for highly educated individuals are shown and in plot (b) the results for low-educated are plotted.

Figure 4.7: Trends in Foreign Population and Policy Preferences

(a) Opposition to Immigration



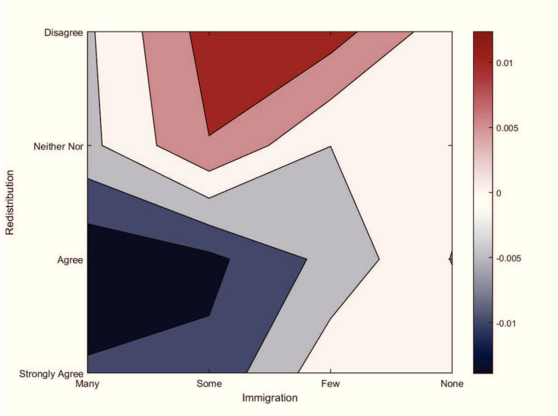
(b) Opposition to Redistribution



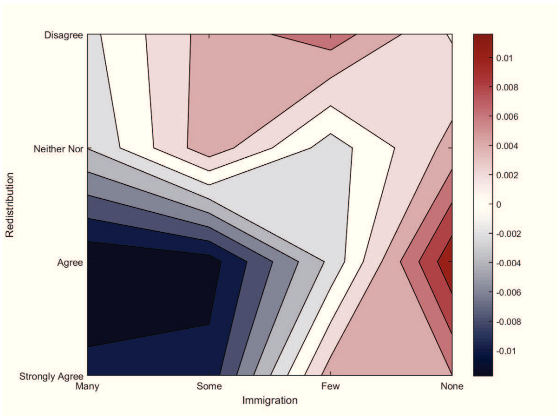
Note: The figure shows the correlation between each country's trend in the stock of foreigners (horizontal axis) and the share of its survey responds who oppose immigration (top figure) or support redistribution (lower figure).

Figure 4.8: Marginal Effects of Higher Foreign-Born Population

(a) High Education



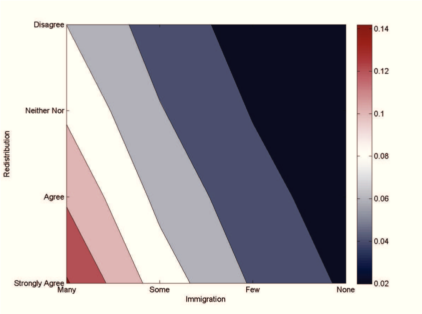
(b) Low Education



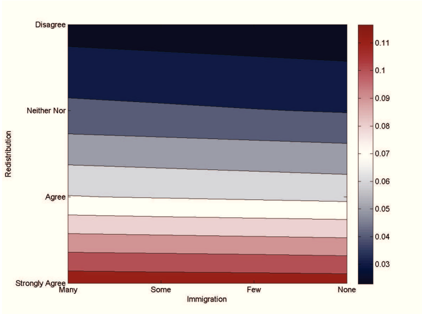
Note: The figures show the estimated marginal effects of a higher share of foreign-born population from a multinomial logit regression, using as a dependent variable the probability of a survey participant choosing one of the sixteen policy combinations. On the horizontal (vertical) axis, there are four possible answers to the question on immigration (redistribution). In plot (a), results for highly educated individuals are shown and in plot (a) the results for low-educated are plotted.

Figure 4.9: Predicted Policy Preferences by Education

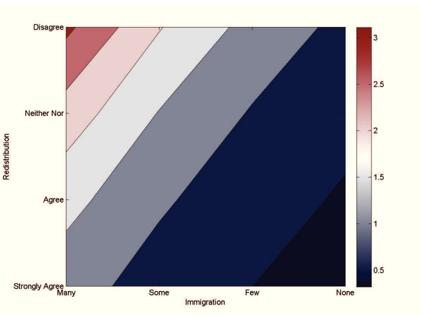
(a) Highly Educated



(b) Low Educated



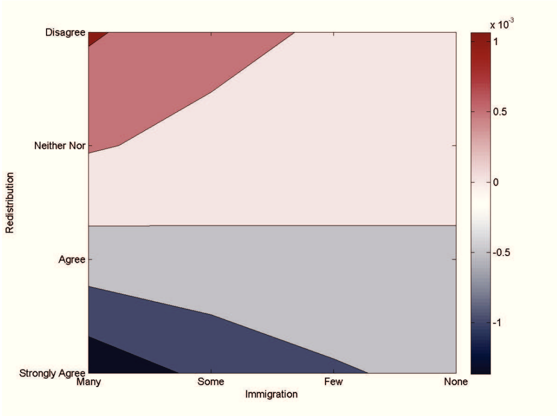
(c) Ratio High-to-Low



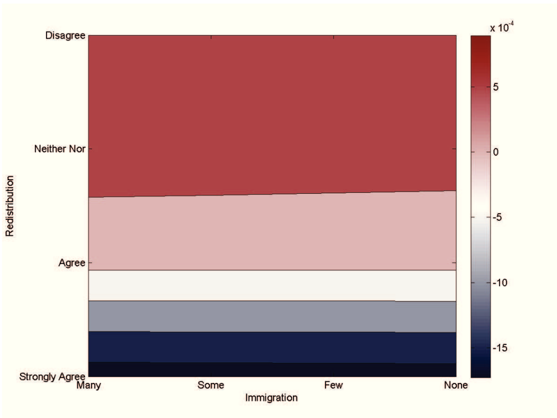
Note: The figures show the model's predicted probability of choosing the different outcomes for those with a high education (Panel a) and low education (Panel b). We also plot the probability ratio of high-to-low educated (Panel c).

Figure 4.10: Changes in Preferences with High Stock of Foreigners

(a) Highly Educated



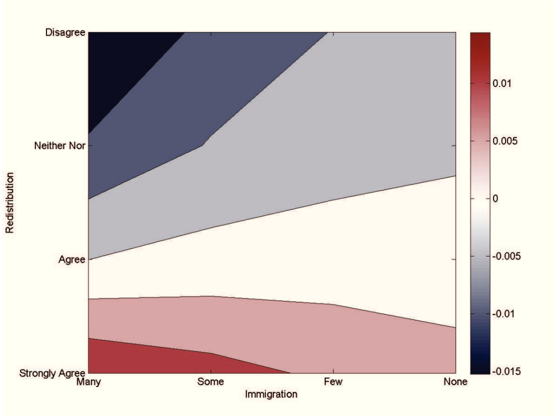
(b) Low Educated



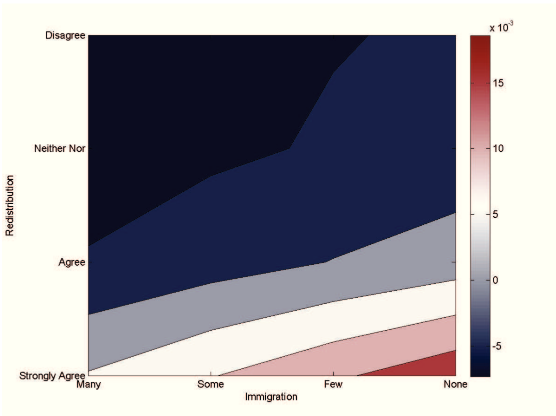
Note: The figures show the model's predicted changes in policy preference when the share of foreign-born citizens increases. The simulated increase mimics the empirically observed change from 5.68% to 8.19%. We show the effect on highly educated individuals in Panel (a) and on low-educated in Panel (b).

Figure 4.11: Changes in Preferences with High Unemployment

(a) Highly Educated

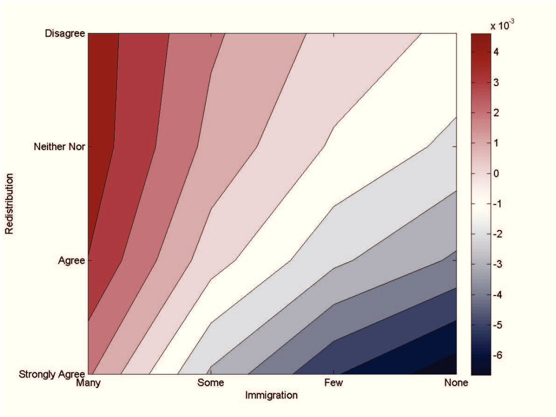


(b) Low Educated



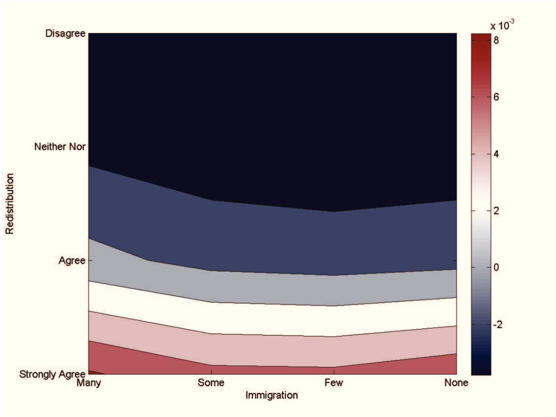
Note: The figures show the model's predicted changes in policy preference when the unemployment rate increases. The simulated increase mimics the empirically observed change from 6.94% to 9.5%. We show the effect on highly educated individuals in Panel (a) and on low-educated in Panel (b).

Figure 4.12: Total Change in Preference by Higher Education



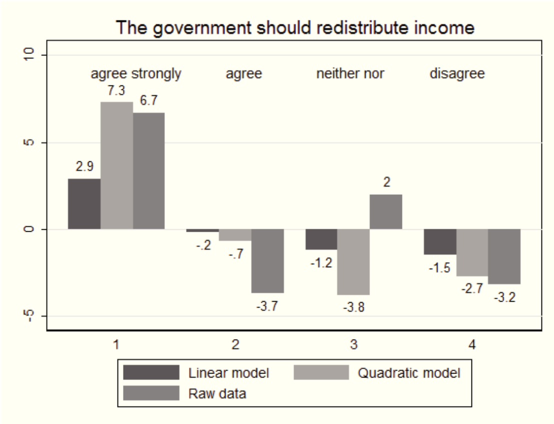
Note: The figure shows the model's predicted changes in policy preference when the share of highly-educated individuals increases. The simulated increase mimics the empirically observed change from 20% to 28.5%.

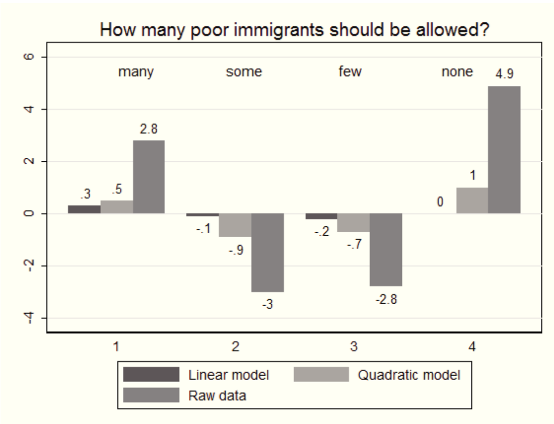
Figure 4.13: Simulated Change in Preferences 2002–2012



Note: The figure shows the model's predicted changes in policy preference when the share of foreign-born citizens increases from 5.85% to 8.19%, unemployment increases from 6.94% to 9.5% and the share of highly educated individuals increases from 21.38% to 32.60%.

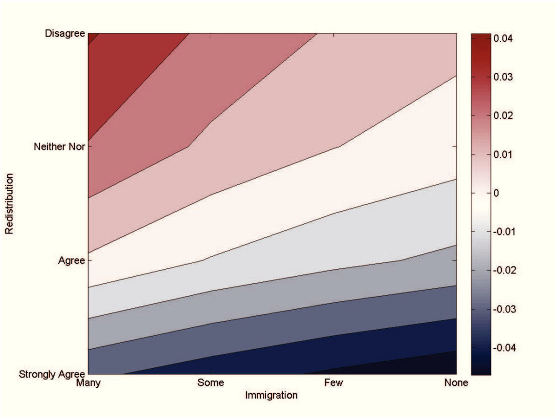
Figure 4.14: Changes in Policy Preferences: Model versus Data





Note: The figures show changes in the shares of voters who support one of the four possible answers to the question on redistribution (Panel a) and immigration (Panel b). The first bar shows the results from simulating the linear model. The second bar is based on the simulation of a quadratic model. And the third bar shows the observed trends in the ESS data which are also shown in Figure 4.1.

Figure 4.15: Simulation of Future Policy Preferences



Note: The figure shows a prediction for the change in total probabilities of choosing policies over immigration and redistribution. For the simulation, we assume average education to follow a linear trend and increase from 28.5% to 37%; the share of foreign born citizens to increase from 8.19% to 10.7%, and the unemployment rate to drop from 9.4% to 4%.

Table 4.1: Countries and Number of Observations

Country	Total Observations	Observations in Year					
		2002	2004	2006	2008	2010	2012
Belgium	10,808	1,899	1,778	1,798	1,760	1,704	1,869
Denmark	9,334	1,506	1,487	1,505	1,610	1,576	1,650
Finland	12,188	2,000	2,022	1,896	2,195	1,878	2,197
France	11,064	1,503	1,806	1,986	2,073	1,728	1,968
Germany	17,445	2,919	2,870	2,916	2,751	3,031	2,958
Great Britain	13,402	2,052	1,897	2,394	2,351	2,422	2,286
Hungary	9,820	1,685	1,498	1,518	1,544	1,561	2,014
Ireland	13,100	2,046	2,286	1,800	1,764	2,576	2,628
Netherlands	11,586	2,364	1,881	1,889	1,778	1,829	1,845
Norway	10,267	2,036	1,760	1,750	1,549	1,548	1,624
Poland	10,815	2,110	1,716	1,721	1,619	1,751	1,898
Portugal	12,453	1,511	2,052	2,222	2,367	2,150	2,151
Slovenia	8,383	1,519	1,442	1,476	1,286	1,403	1,257
Spain	11,618	1,729	1,663	1,876	2,576	1,885	1,889
Sweden	11,048	1,999	1,948	1,927	1,830	1,497	1,847
Switzerland	10,803	2,040	2,141	1,804	1,819	1,506	1,493

Note: The table shows the number of observations for each country and year. The selection of countries shown here is restricted to those that participated in each of the biannual ESS waves between 2002 and 2012.

Table 4.2: Distribution of Policy Preferences

Government should redistribute income	**How many poor immigrants** many	some	few	none	$\sum$
disagree	2,961	10,408	9,327	3,071	25,767
in %	1.68	5.91	5.29	1.74	14.62
neither nor	2,899	10,911	8,961	2,662	25,433
in %	1.64	6.19	5.08	1.51	14.42
agree	9,113	33,126	2,6536	9,530	78,305
in %	5.17	18.79	15.06	5.41	44.43
agree strongly	5,874	16,040	15,947	8,890	46,751
in %	3.33	9.1	9.05	5.04	26.52
$\sum$	20,847	70,485	60,771	24,153	176,256
in %	11.83	39.99	34.48	13.70	100.00

Note: The table shows how many survey participants preferred each of the sixteen possible policy combinations over immigration ('How many poor people from outside Europe should be allowed to enter the country?') and redistribution ('Do you agree with the statement: The government should reduce differences in income levels?'). In addition to the total number of observations, we also show the share of people choosing a policy combination. We combine the answers 'disagree' and 'disagree strongly' on the redistribution question. The sample covers all biannual surveys from 2002 to 2012.

Table 4.3: Descriptive Statistics

Variable	Mean	SD	Min	Max	Obs.
Age	47.75	18.56	14	105	183,476
Male	0.47	0.50	0	1	183,981
High Education	0.26	0.44	0	1	183,106
Retired	0.26	0.44	0	1	179,286
Wage Earner	0.59	0.49	0	1	179,286
Benefits Recipient	0.06	0.24	0	1	179,286
Share of Foreigners	0.07	0.05	0.00	0.23	85
Unemployment Rate	0.08	0.04	0.03	0.25	96
Gini Coefficient	0.40	0.10	0.22	0.54	72
Income Ratio P90/P50	2.30	0.33	1.83	3.20	63
Income Ratio P10/P50	0.39	0.06	0.17	0.50	63

Note: The table shows summary statistics for each variable we use in the empirical part. The top part shows data taken from all ESS biannual surveys between 2002 to 2012. In the lower part, we show macroeconomic variables based on data from OECD, World Bank, and UNU WIDER.

Table 4.4: Determinants of Policy Preferences in ESS Data

	Income Redistribution				How Many Immigrants			
	Agree Strongly		Disagree		Many		None	
Mean value	(0.265)		(0.710)		(0.118)		(0.137)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Male	-0.024***	-0.022***	0.055***	0.054***	-0.005	-0.006	-0.003	-0.002
	(0.006)	(0.006)	(0.007)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)
Age	0.001***	0.002***	0.000	-0.000	-0.002***	-0.002***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
High Edu.	-0.068***	-0.054***	0.086***	0.080***	0.071***	0.069***	-0.086***	-0.079***
	(0.011)	(0.009)	(0.006)	(0.007)	(0.012)	(0.010)	(0.012)	(0.008)
Benefits R.	0.080***	0.098***	-0.040***	-0.055***	0.013	0.024***	0.054***	0.047***
	(0.013)	(0.010)	(0.012)	(0.012)	(0.008)	(0.006)	(0.013)	(0.011)
Retired	0.017	0.002	-0.030***	-0.024***	-0.003	0.002	0.019**	0.011**
	(0.012)	(0.006)	(0.007)	(0.005)	(0.007)	(0.004)	(0.008)	(0.005)
U. Rate	1.337***	0.490*	-0.928***	0.048	0.217	-0.075	0.487	0.457*
	(0.417)	(0.238)	(0.295)	(0.105)	(0.156)	(0.284)	(0.317)	(0.240)
Share F.	-0.535	0.312	0.136	-0.346	0.170	-0.001	-0.525	0.558*
	(0.454)	(0.408)	(0.237)	(0.293)	(0.188)	(0.343)	(0.371)	(0.303)
Country FE	-	Yes	-	Yes	-	Yes	-	Yes
Obs.	153,744	153,744	153,744	153,744	153,744	153,744	153,744	153,744
R-squ.	0.034	0.082	0.033	0.075	0.022	0.060	0.040	0.091

Note: The table shows the results of eight separate OLS regressions. The dependent variable is a dummy variable taking the value one according to the survey answer indicated in the second row of the table. In columns (3) and (4) the dependent variable takes the value one for all individuals who either ‘agree strongly’ or ‘agree’ with redistribution. Unemployment rate and share of foreigners are measured at the country level, all other variables at the individual level. We use sampling weights based on year and country’s population size. Standard errors (in parentheses) are clustered at the country level. Significance at the 10% level is indicated by *, at the 5% level by **, and at the 1% level by ***.

Table 4.5: Estimation of the Model

	β / (T-stat)
Low Education × z × (1 - Unemployment Rate)	1.947**
	(2.75)
High Education × z × (1 - Unemployment Rate)	-1.679*
	(-2.02)
Low Education × z × Unemployment Rate	76.97***
	(9.45)
High Education × z × Unemployment Rate	60.61***
	(6.53)
Low Education × L × (1 - Unemployment Rate)	7.563**
	(2.67)
High Education × L × (1 - Unemployment Rate)	31.41***
	(9.55)
Low Education × L × Unemployment Rate	-63.36*
	(-2.05)
High Education × L × Unemployment Rate	-16.53
	(-0.48)
Observations	2,336

Note: The table shows the coefficients from an estimation of equation (4.19). We indicate t-statistics in parentheses and significance at the 10% level are denoted by *, at the 5% level by **, and at the 1% level by ***.

Table 4.6: Determinants of Charity in the ESS Data

	Doing Charitable Work				Donating Money	
Mean of Dep.Var.	(0.139)		(0.130)		(0.113)	
	(1)	(2)	(3)	(4)	(5)	(6)
Income Decile	0.005***	0.005***	0.004***	0.004***	0.010***	0.010***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Log Age	0.039***	0.043***	0.028**	0.034***	0.007	0.013**
	(0.010)	(0.011)	(0.011)	(0.011)	(0.005)	(0.006)
Male	0.013**	0.011*	0.002	0.002	-0.031***	-0.032***
	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)
High Education	0.062***	0.058***	0.053***	0.051***	0.103***	0.096***
	(0.006)	(0.006)	(0.007)	(0.006)	(0.011)	(0.010)
Retired	0.039*	0.036*	0.023***	0.023***	0.003	0.004
	(0.020)	(0.019)	(0.007)	(0.008)	(0.006)	(0.007)
Wage Earner	-0.011	-0.012	-0.016*	-0.015*	-0.002	-0.002
	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.009)
Benefits Recipient	-0.004	-0.002	-0.020**	-0.017*	-0.010	-0.008
	(0.011)	(0.011)	(0.008)	(0.009)	(0.011)	(0.011)
Support Immigration		0.028**		0.028***		0.069***
		(0.010)		(0.007)		(0.017)
Reject Immigration		-0.024***		-0.014**		-0.038***
		(0.007)		(0.006)		(0.012)
Year of Data	2006	2006	2012	2012	2002	2002
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,711	26,494	43,066	40,774	25,739	24,339
R-squared	0.072	0.071	0.077	0.077	0.079	0.083

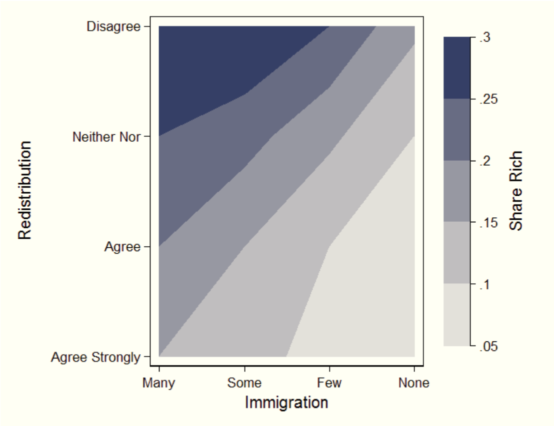
Note: The table shows the results of six separate OLS regressions. The dependent variable is a dummy variable taking the value one if the survey participant does charitable work at least once a month (columns 1-4), or donates money to a humanitarian organization (columns 5-6). Support (reject) immigration is a dummy variable taking the value one if the survey participant chooses many (none) when asked how many poor immigrants should be allowed to enter the country. We use sampling weights based on year and country's population size. Standard errors (in parentheses) are clustered at the country level. Significance at the 10% level is indicated by *, at the 5% level by **, and at the 1% level by ***.

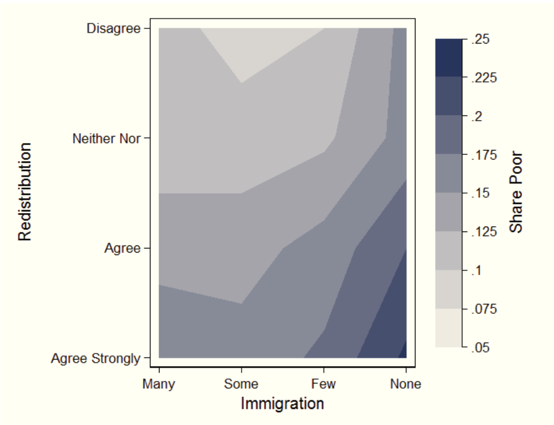
Appendix

Policy Preferences by Income

We show policy preferences in the two-dimensional space by education in Figure 4.3. For each combination of answers to the question on redistribution and immigration, we calculate the share of high-to-low educated individuals. In the figure below, we plot again the shares but now for individuals with high or low *income*. As before, we show the share choosing a particular combination of preferences over immigration and redistribution.

Figure 4.16: Policy Preferences by Income





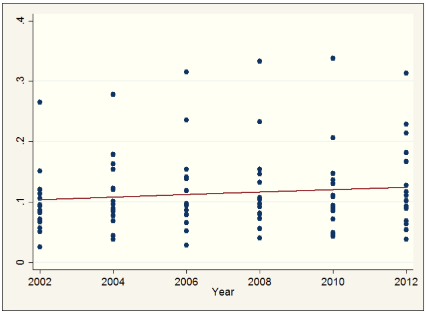
Note: The figures show the ratio of the share of highly educated individuals relative to the share of low-educated individuals in the two-dimensional map. While preferences over immigration are on the horizontal axis, views on redistribution are shown on the vertical axis. In the top figure, the frequency of individuals with high income, in the lower figure individuals with low income are shown.

Trends in Policy Preferences

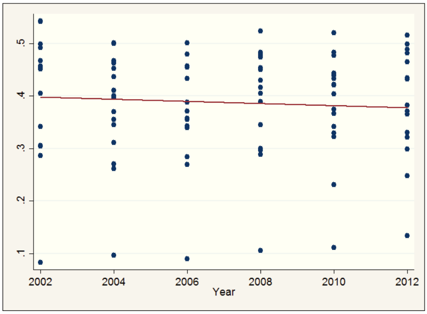
In Figure 4.2, we show how policy preferences changed in the sixteen countries that participated in each of the ESS waves between 2002 and 2012. This provides information on the trend aggregated over all countries and the entire time period. In Figures 4.17, 4.18, 4.19 and 4.20 below, we show how policy preferences changed year-by-year and in each country.

Figure 4.17: Trends in Policy Preferences over Immigration by Year

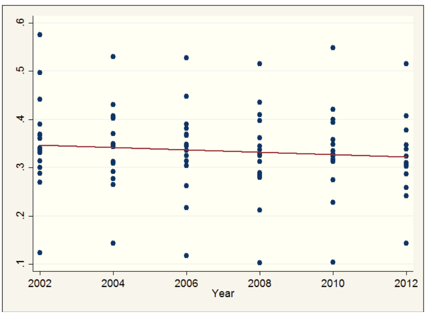
(a) Many



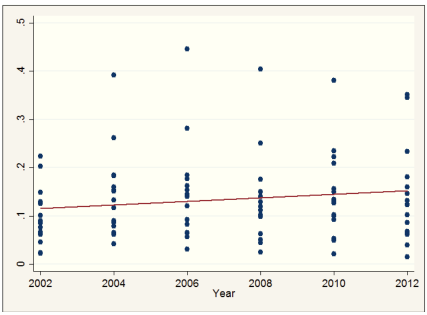
(b) Some



(c) Few



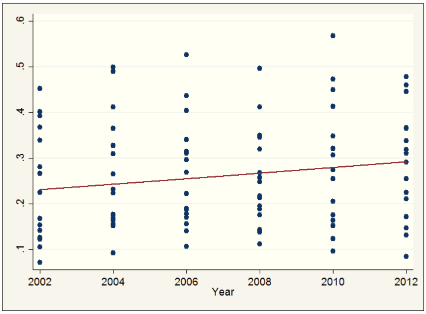
(d) None



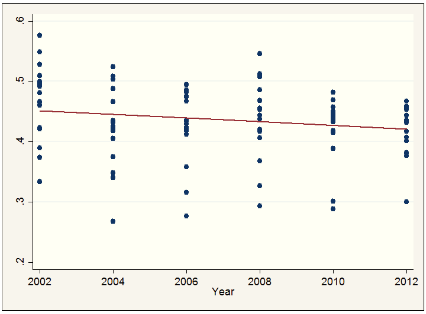
Note: The figure shows the share of survey participants selecting a given answer to the question on immigration. Each dot refers to one country-year observation.

Figure 4.18: Policy Preferences over Redistribution by Year

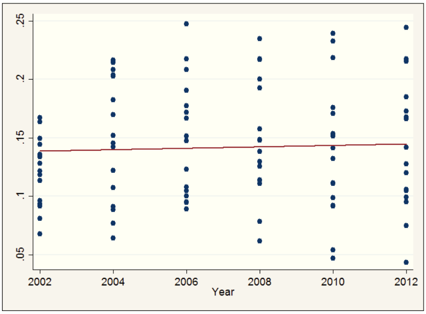
(a) Agree Strongly



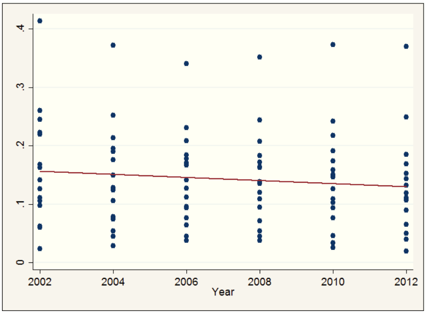
(b) Agree



(c) Neither Nor



(d) Disagree

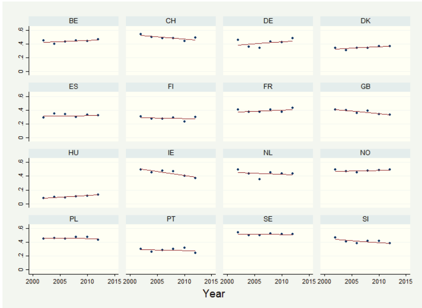


Note: The figure shows the share of survey participants selecting a given answer to the question on income redistribution. Each dot refers to one country-year observation.

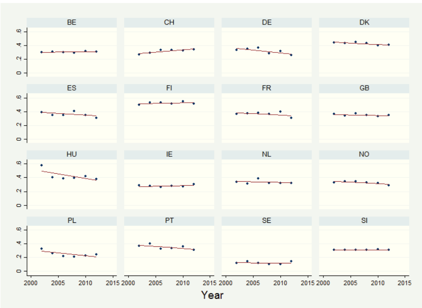
Figure 4.19: Preferences over Immigration by Year and Country

(a) Many

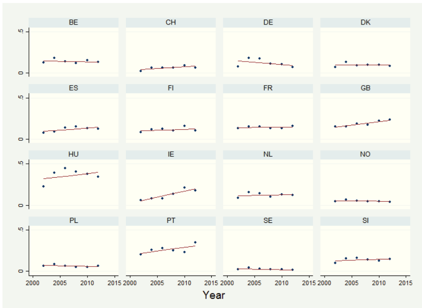
(b) Some



(c) Few



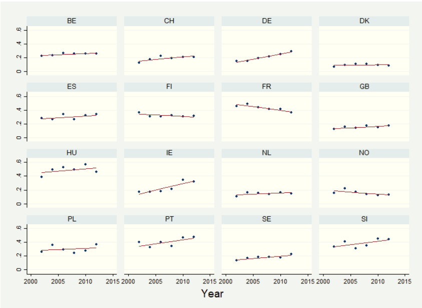
(d) None



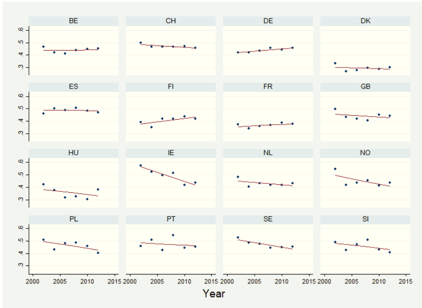
Note: The figure shows the share of survey participants selecting a given answer to the question on immigration.

Figure 4.20: Preferences over Redistribution by Year and Country

(a) Strongly Agree



(b) Agree



(c) Neither Nor

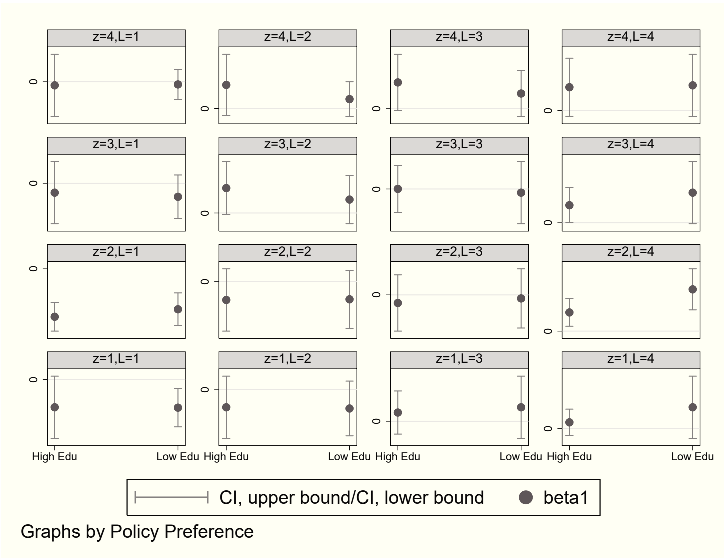
(d) Disagreee

Note: The figure shows the share of survey participants selecting a given answer to the question on income redistribution.

Regression Output from the Multinomial Logit

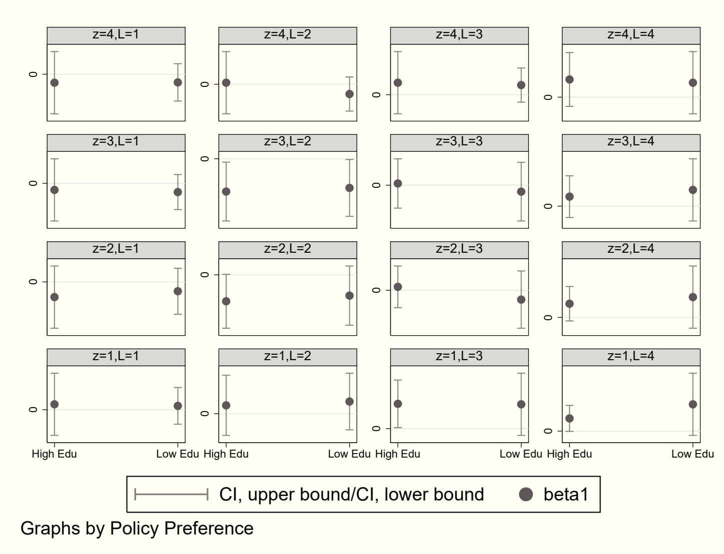
Figures 4.8 and 4.6 illustrate the estimated marginal effects of a higher share of foreign-born population and a higher rate of unemployment from multinomial logit regressions, using as a dependent variable the probability of a survey participant choosing one of the sixteen policy combinations. In order to provide the point estimates as well as the confidence intervals, we show the regression output in Figures 4.22 and 4.22 below.

Figure 4.21: Marginal Effects of Higher Share of Foreigners



Note: The figure shows the estimated marginal effects (at mean values of covariates) from the multinomial Logit regressions. On the horizontal (vertical) axis, there are four possible answers to the question on immigration (redistribution). The vertical bars indicate 95% confidence intervals.

Figure 4.22: Marginal Effect of Higher Unemployment Rate



Note: The figure shows the estimated marginal effects (at mean values of covariates) from the multinomial Logit regressions. On the horizontal (vertical) axis, there are four possible answers to the question on immigration (redistribution). The vertical bars indicate 95% confidence intervals.

Quadratic Utility Function

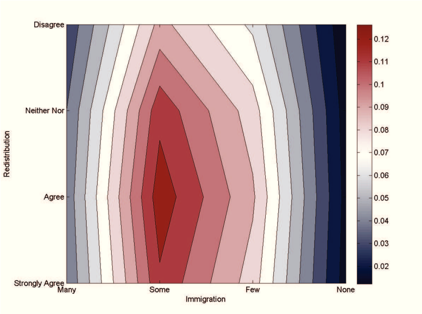
The model in section 4.4 use a linear approximation to utility, which might be seen as a too simple approach. As a robustness check, the following sections redo the analysis using a quadratic approximation of utility. In general, the quadratic gives a quantitatively better fit, but the qualitative results are similar. In total, the quadratic model predicts an increase in L4 by 1.05 percentage points, and an increase in L1 by .55 percentage points. Finally, the quadratic version of the model predicts a 7.27 percentage point increase in Z1, i.e. the most favorable preference towards redistribution.

4.6.1 Static Preferences

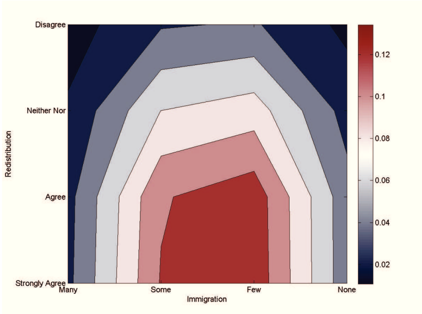
Similar to Figures 4.9 which is based on the linear model, we show static preferences over immigration and redistribution in the figure below.

Figure 4.23: Predicted Policy Preferences by Education

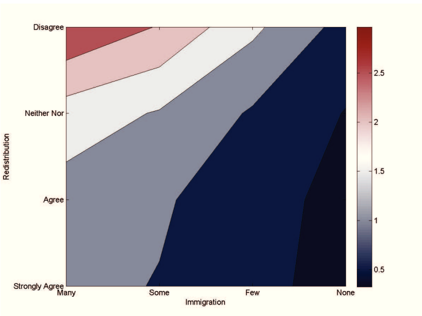
(a) High Education



(b) Low Education



(c) Ratio High-to-Low



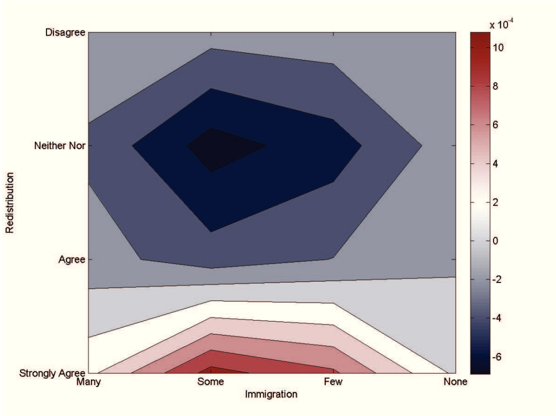
Note: The figures show the model's predicted probability of choosing the different outcomes for those with a high education (a) and low education (b). We also plot the probability ratio of high-to-low educated (c). The simulation is based on a quadratic version of our model.

Effects of Macroeconomic Trends

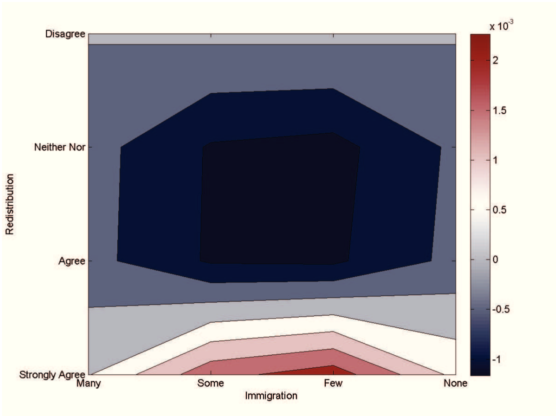
Adding to the simulation of the linear model show in Figures 4.10, 4.11, and 4.12, below we show how policy preferences change in the quadratic model if the share of foreign-born population, the unemployment rate, or the share of highly educated increases. In addition, following Figure 4.13 which is based on the linear model, we show the total effect of all macroeconomic variables using a quadratic model in the figures below.

Figure 4.24: Changes in Preferences with High Stock of Foreigners

(a) High Education



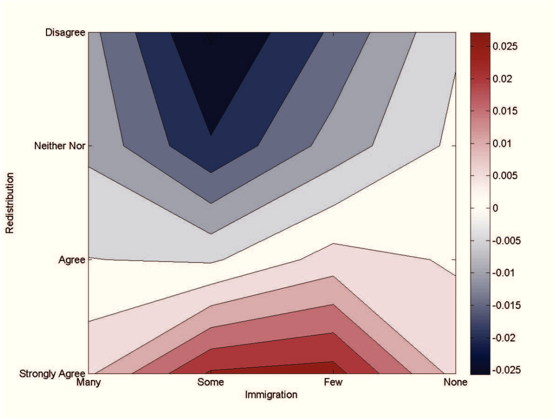
(b) Low Education



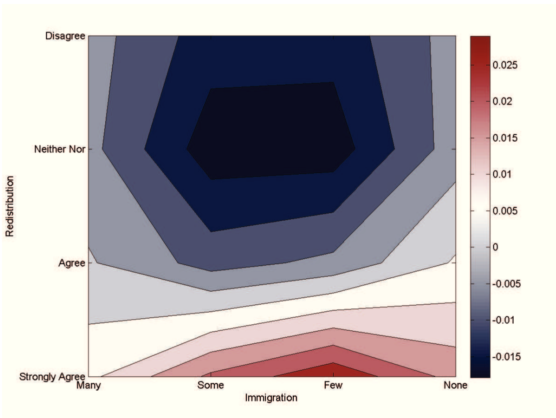
Note: The figures show the model's predicted changes in preference probabilities when the stock of foreign born citizens increases from 5.68% to 8.19%. High education on left panel and low education on right panel. The simulation is based on a quadratic version of our model.

Figure 4.25: Changes in Preferences with High Unemployment

(a) High Education

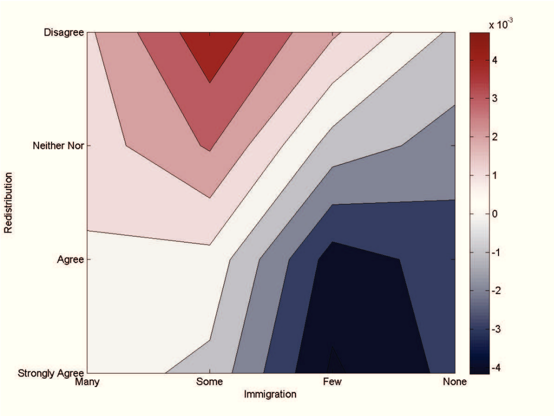


(b) Low Education



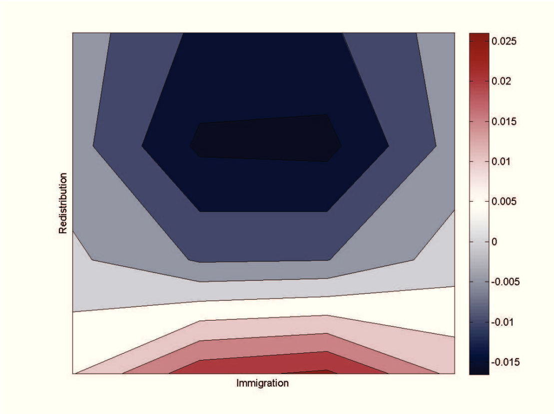
Note: The figures show the model's predicted changes in preference probabilities when the unemployment rate increases from 6.94% to 9.5%. High education on left panel and low education on right panel. The simulation is based on a quadratic version of our model.

Figure 4.26: Change in Preferences with Increasing Education



Note: The figure shows the change in total probabilities when average education increases from 20% to 28.5%. The simulation is based on a quadratic version of our model.

Figure 4.27: Simulated Change in Preferences 2002–2012



Note: The left panel show the change in total probabilities when stock of foreign born citizens increases from 5.85% to 8.19%, unemployment increases from 6.94% to 9.5% and average education increases from 21.38% to 32.60%. The simulation is based on a quadratic version of our model.

Chapter 5

Media Attention and Betting Markets

This chapter is based on joint work with Lukas Schmid from the University of St.Gallen and published as “Media Attention and Betting Markets” in the *European Economic Review* (2016) Vol. 87, p.304–333.

5.1 Introduction

There is a large and growing body of literature documenting the media’s impact on economic, political, and social outcomes [DellaVigna and La Ferrara, 2015; Strömberg, 2015]. In particular, prior research has pointed out that media outlets play a key role in reducing complex information for consumers with limited time resources [Falkinger, 2007]. Hence the media focus on certain events, groups, and individuals which in turn has been shown to affect decisions and behavior in various domains, namely which political candidates are elected [Epstein and Robertson, 2015], how politicians react to disasters [Eisensee and Strömberg, 2007], which issues voters care about [Mastrorocco and Minale, 2016], what family size people prefer [Jensen and Oster, 2009; La Ferrara, Chong and Duryea, 2012], or what consumption bundles consumers demand [Bursztyn and Cantoni, 2016].

While this evidence suggests that the media can influence what individuals think of the past and the present, we know surprisingly little about its impact on what people think about the future. In order to study individual perceptions about future events, economists typically use betting markets because they provide an incentive-compatible way for individuals to truthfully reveal their expectations [Wolfers and Zitzewitz, 2006]. However, several empirical studies document that behavioral biases or risk-loving preferences might lead to a situation in which prediction markets do not reveal true probabilities [Friedman and Savage, 1948; Kahneman and Tversky, 1979; Snowberg and Wolfers, 2010]. Another important and unexplored explanation for why betting odds deviate from true probabilities, however, is that bettors make decisions using information they receive from the media. If media outlets allocate significantly more attention to some individuals than to others, expectations about future events related to these individuals might change. For example, bettors might overestimate future prospects of politicians who receive the lion's share of media coverage after a debate. Similarly, athletes who succeed in one tournament and are covered extensively by media outlets may be perceived more likely to succeed in subsequent tournaments. The methodological challenge researchers face when estimating such an effect of media attention on expectations is that the allocation of attention across individuals is not randomized but correlated with unobserved factors like skill or talent.

This paper offers a novel identification strategy to estimate the causal effects of media attention on betting markets by analyzing close competitions in World Cup alpine skiing from 1992–2014. For several reasons this setting provides a unique real-world natural experiment. First, there are considerable differences in how much media attention athletes receive. Those individuals at the top of the classification typically receive the bulk of attention. Second, unlike in most other settings with ranking schemes, our data set enables us to observe the same individual numerous times. Hence we can compare the amount of attention athletes receive when they achieve or miss a top ranking position by including athlete-fixed effects. Third, in alpine skiing small time differences resulting from random shocks, like weather and snow conditions, can manipulate ranking positions. Hence luck is more prevalent in skiing than in other sports such as, for example, golf [Connolly and Rendleman, 2008]. In close races it is often a tiny margin —a few hundredths of a second— that determines whether an athlete finishes first or second, third or fourth, or even sixth or tenth.[1] Assuming such small time differences to reflect random noise allows us to argue that those athletes who achieved a higher rank in a close race quasi-randomly received more media attention. Hence we can explore the causal effects of media coverage on betting market outcomes for these athletes.

Our findings suggest that top-ranked individuals receive 39.5% more media attention compared to athletes with an arguably similar performance who barely missed the highest ranking positions. We also document that betting odds *generally* decrease with both ranking positions and media attention. However, when focusing on *close races* for which ranking positions and thus media attention are allocated randomly, we find no discontinuity in betting market outcomes. This suggests that media attention does not cause betting market participants to change their beliefs about the future prospects of athletes. In contrast, we find that betting markets are efficient in the sense that observed odds match estimated true probabilities. Only when extending the sample to include athletes relatively far away from top ranking positions, we find a significant difference in both true probabilities and betting odds between top ranked and not top ranked athletes.

We conduct a series of robustness checks to assess the validity of our findings. First, we show that there are no systematic differences in pre-determined covariates between those who achieve, for instance, a podium finish in a close race and those who barely miss it. This supports our empirical strategy that uses random top ranks as exogenous manipulation of media attention. Second, our results are shown to be robust to the choice of bandwidth for the definition of random top ranks. Third, we document that higher ranking positions do not causally change an athlete's risk-taking behavior or performance in subsequent races. Fourth, we point out that our main results from reduced form estimations

[1] Didier Cuche's last seconds in the 2010 Vancouver Olympics downhill race exemplify that small time differences can result in large differences in ranking positions. Lagging only 0.06 seconds behind leader Didier Défago, Cuche was second at the beginning of the last ten seconds. But Cuche did not optimally pass the last gate and finished sixth, trailing the winner Défago by 0.36 seconds.

do not change if we estimate instrumental variables regressions. Finally, we show that our core findings on media attention are not confined to the Swiss newspaper archive ‘Swissdox’ but generalize to the American-based ‘NewsLibrary’.

Our paper provides several novelties. First, to our knowledge, we are the first to provide causal evidence that rankings generate significant differences in media attention. After a close race, newspapers focus on those athletes who achieved the top positions in the ranking even if performance differences were tiny. A central explanation for why media outlets focus on individuals who achieved a top rank is that they compete for consumers in an information-rich economy [Falkinger, 2007, 2008]. Recent advances in information technology have dramatically increased the supply of information. Newspapers and TV programs respond to this by focusing particularly on the most relevant and successful individuals. This form of biased media attention is relevant because previous research documents that it can affect election outcomes [DellaVigna and Kaplan, 2007; Epstein and Robertson, 2015] as well as consumer preferences [Gentzkow and Shapiro, 2010].

Second, we show that media attention has no effect on betting behavior. This finding adds to prior research on how individuals deal with rankings and ratings. Salganik, Dodds and Watts [2006] find that the availability of music ratings increases downloads of already successful songs. Similarly, Feenberg et al. [2015] show that NBER working papers which are listed first in the weekly newsletter receive substantially more views and citations. Furthermore, the use of explicit ranking schemes has been found to affect how consumers choose restaurants [Luca, 2016].[1] Our findings, however, suggest that the allocation of attention does not affect betting market outcomes such as odds or the number of bets placed. We argue that the incentive structure and efficiency of the market provide an explanation for this observation.

Third, our results show that professional sport athletes exhibit substantial serial correlation in their performance but do not change their risk behavior or performance following a top ranking position in one tournament. Previous research has pointed out that lagging behind in a ranking increases risk-taking and

[1] Further empirical evidence on the powerful impact of rankings in the field of M&A markets has been provided by Derrien and Dessaint [2015].

lowers final performance [Genakos and Pagliero, 2012].[1] Furthermore, our findings contribute to a sizable literature on the so-called hot-hands effect [Gilovich, Vallone and Tversky, 1985; Green and Zwiebel, 2015; Miller and Sanjurjo, 2015]. In particular, we find that one-time successes do not have a causal positive effect on subsequent performance. Finally, several studies show that the mere provision of a relative ranking position can affect student performance [Kuhnen and Tymula, 2012; Tran and Zeckhauser, 2012] as well as employee satisfaction [Card et al., 2012]. The key difference between our work and these earlier studies is that we analyze the effect of rankings using a setting in which ranking positions are arguably randomized in close competitions. In addition, having more than twenty years of data allows us to observe the same individual multiple times in both the treatment and control group.

The paper proceeds as follows. Section 5.2 presents general information on World Cup alpine skiing as well as descriptive statistics on our dataset. This comprises a description of the data on media attention as well as the betting market. Section 5.3 discusses the problem of identifying the causal effect of media attention on betting market outcomes. We explain the concept of the quasi-random allocation of top ranks and how this can be used to overcome the identification problem. Section 5.4 presents our empirical findings on how media attention affects betting market outcomes. This includes a series of robustness checks. Finally, Section 5.5 concludes.

5.2 Data

Our data set provides an opportunity to study the behavior of World Cup athletes, media outlets, and bettors in an environment with large stakes and fierce competition.[2] In this section, we first provide background information on World Cup skiing competitions. Moreover, we describe our data set which includes in-

[1] In a laboratory setting, Gill et al. [2016] document a U-shaped relationship between ranking position and effort, reflecting both 'first-place loving' and 'last-place loathing'.

[2] Klaassen and Magnus [2009] discuss the usefulness of sports data to examine behavioral questions. DellaVigna [2009] provides a summary of research documenting that behavioral biases may disappear among experienced individuals.

formation on race results, media attention and betting market outcomes for all tournaments.[1]

5.2.1 World Cup Alpine Ski Tournaments

The origins of alpine skiing competitions go back to the 1930s when European ski clubs, most prominently in Switzerland, Austria, and Germany, decided to organize races. In 1967, the *Fédération Internationale de Ski* (FIS) decided to bring these separate events together and launched the FIS World Cup. Today, alpine skiing competitions enjoy great popularity, particularly in Europe. The downhill race in Wengen (Switzerland), for instance, was followed by a TV audience of over one million viewers in Switzerland (one eighth of the country's population) for each of the races between 2007 and 2012 [Ski World Cup Wengen, 2012]. A similar appeal comes from the downhill and slalom races in Kitzbühel, each of which is watched by more than 1.3 million Austrians. In addition to large audiences, sizable prize money is also par for the course. Among all top-ten athletes in the season of 2012/2013, the prize money sums up to $4.4 million for men and $4.2 million for women (FIS 2013). However, the distribution of income in prize money is highly skewed. The highest income among male athletes was $589,009 and among females it was $771,289. Number ten of the prize money ranking earned only $109,010 and $126,858, respectively. A considerable fraction of 76% of male and 80% of female athletes earned less than $50,000.[2]

The goal of alpine skiing is to slide down a race course in the fastest overall time. Each course consists of a series of gates. All of them have to be passed correctly, so that all athletes run the same course. The five disciplines differ in terms of the vertical and horizontal distance between the gates as well as the horizontal distance between start and finish. The average speed of a downhill

[1] While our data set on World Cup tournaments and media attention includes all races from 1992–2014, the analysis of betting markets is restricted to the period 2006–2014 because we only have the respective data for this period.

[2] Note that these prizes are large enough to incentivize athletes to exert high effort but not too large to cause one-time winners to reduce their subsequent efforts. Besides the prize money, top ranks in World Cup races can also lead to better sponsorship contracts. While there is no reliable data on sponsorship incomes, insiders estimate that in the case of top athletes, this source of income makes up three to four times the amount of prize money.

racer is about 100 km/h, while in slalom races the athletes usually achieve about 40 km/h.

5.2.2 Data on World Cup Skiing

We use a panel data set on 473 male and 428 female athletes in all 1,587 World Cup ski races for the period of 1992–2014. The data set includes information on whether an athlete finished the race, the exact result in hundredths of a second, as well as gender, age, and the discipline of competition.[1] The panel structure allows us to measure each athlete's performance in subsequent races. In total, our data set contains 23,761 observations when the unit of observation is an athlete in a specific race.

— Table 5.1 about here —

Table 5.1 reports descriptive statistics for the data we use in the empirical analysis. In part (I), we show information on all athletes competing in World Cup tournaments for the period of 1992–2014. Since we focus on top ranks, the sample is restricted to all athletes in the top fifteen. The share of athletes on the podium is 20.3% of the total number of observations. Observations for today's race are more numerous than observations for past and future performance because we only use outcomes within season, which results in missing values for the first and last race of the season for each race discipline. Furthermore, we cannot use observations at the end of the season because we would lack future performance and betting odds. Taken together, these account for around a fourth of the total observations. Finally, the number of observations is reduced because only 82% of those athletes who compete in the next race finish the race and get a positive race time. It is important to note that competition in alpine skiing is fierce. Only few junior athletes make it to the World Cup team and, among them, only a small group is successful. From our total sample, only about seven percent of athletes ever won a race during their entire career. This suggests that only a small set of competitors is very successful over a lifespan which is in line with

[1] Race times are actually measured more precisely than stated in official reports. For any time in hundredths of a second, the measurement was accurate at the millisecond level.

empirical evidence concerning the presence of superstars in music, entertainment, and academia [Hamlen, 1991; Rosen, 1981].[1]

5.2.3 Data on Media Attention

We complement our data set of individual World Cup tournament results with information on media attention and betting odds. For the former, we scraped data from the Swiss newspaper database "Swissdox" for various time horizons before and after the race. Overall, the Swissdox database covers more than 200 newspapers, almost all of which for the entire time period of 1992–2014. Our search queries included an athlete's name and the time horizon of the search. We measure media attention by the number of articles mentioning the athlete's name. Part (II) of Table 5.1 shows descriptive statistics for the number of articles published at various points in time. Not surprisingly, there are more articles about a particular athlete the more we extend the time window.

In Figure 5.1, we show the distribution of media attention across ranking positions. Panel (a) depicts the average number of newspaper articles that mention an athlete's name on the day after the competition. While winners are mentioned in 17.8 articles on average, athletes ranked second or third get an average media presence of 13.0 and 11.2 articles, respectively. The average number of articles is considerably lower for other athletes, namely 5.4 for athletes on positions four to ten and 3.7 for athletes on position eleven to fifteen. The pattern in media attention is notably similar when focusing on media attention during the week or month following the race.[2]

— Figure 5.1 about here —

Because the source of our media data, Swissdox, may be biased towards Swiss and German-speaking athletes, we repeat the scraping procedure for "NewsLibrary", a US-based online news database that includes more than 4,000 outlets.

[1] The most successful athletes in the history of alpine skiing are Ingemar Stenmark (Sweden, 1973–1989) with 86 victories and Lindsey Vonn (USA, 2002-) with 76 victories as of June 2016.

[2] Note that the distribution of media attention we obtain using our scraping algorithm yields very similar results for other fields of sports, such as Formula 1 competitions as illustrated by Figure 5.2 in the Appendix.

In the robustness section we use this data to show that we obtain similar regression results for the media attention, irrespective of which source of media data we use. In addition, Figure 5.1 in the Appendix illustrates that the distribution of media attention across ranking positions is almost identical among newspapers covered by NewsLibrary when compared to newspapers covered by Swissdox in Panel (a) of Figure 5.1. The reason why we use Swissdox data for our main analysis is that the NewsLibrary search often finds mainly articles with ranking lists. In contrast, using newspapers in Swissdox, we find mostly specific reports on World Cup skiing. Since mere ranking lists are not the kind of media attention that we are primarily interested in, we employ Swissdox data for our main analysis. Nevertheless, it is important to stress that we obtain similar distributions of media attention from both sources.

5.2.4 Data on Betting Behavior

In order to obtain information on betting market outcomes, we collected data from the world's largest internet betting exchange "Betfair". We focus exclusively on bets for a specific athlete to win the next race. This corresponds to about three quarters of all bets.[1] This set of data includes a total of 77,202 individual bet observations and is available for the period 2006-2014. Note that each individual observation corresponds to a bet offered for a specific event (e.g., athlete A to win the next tournament) with a specific odd. For our analysis, we use three different odds for each athlete-event combination.

— Figure 5.2 about here —

As illustrated in Figure 5.2, the betting market for the subsequent race opens after a given race t. At the beginning, a bet for each athlete who is likely to compete in race $t + 1$ is offered at an *initial odd*. Over time, this odd changes due to new information (e.g., a competitor being injured) or because of altered

[1] In our sample, 75.4% of all bets and 72.0% of the total volume are placed on who will win a race (the group we focus on in our study). In contrast, only 2.4% of all bets and 1.8% of the total volume are placed on the winner of the overall World Cup. The rest of bets is placed on a variety of different outcomes (for example, whether athlete A performs better than athlete B).

demand for bets. If, for example, the demand for a bet on athlete A surges, the odd offered by the betting agency for this athlete will go down. At some point prior to the race $t+1$, the betting market closes and we record the *final odd* for each athlete. Hence, we have an initial and a final odd for each athlete and event. In addition, we aggregate all individual odds for a specific event and athlete to obtain an *average odd*. In the robustness section, we use all three odds to investigate how media attention affects the betting market at different points in time.

Panel (b) of Figure 5.1 illustrates the distribution of the inverse average betting odds across ranking positions. Betting odds are what a bettor gets paid for a bet of one unit if the specific event of the bet is realized. For example, a betting odd of 1.2 on athlete A means that the bettor collects \$1.2 for a one-dollar-bet if athlete A actually wins the competition. This translates into an inverse odd of 1/1.2=0.833 that reflects the implied probability that athlete A wins the race. Because these implied probabilities are directly comparable to the true probability that athlete A wins the race, we will henceforth use the reciprocal odd as our main measure of betting odds.

Similar to the distribution of media attention, we note substantial differences in average betting odds across ranks. While winners of the current race have an average implied probability of 0.25 in the subsequent race (which is equivalent to paying off \$4 for each dollar invested), a bet on one of the other athletes on the podium has an implied probability of about 0.14. Placing a bet on an athlete on position four to fifteen yields an average implied probability that is substantially smaller (0.08). As expected, those finishing on the top ranks in a tournament receive significantly more bets—both in terms of the number of bets and total money volume—in the subsequent tournament.

— Figure 5.3 about here —

It is important to understand who actually participates in the alpine skiing betting market. The total betting volume in our data set is \$1.7 million. Figure 5.3 depicts the distribution of the betting volume for all bets up to 200 dollars (equivalent to 97.4% of the total betting volume). The figure reveals that more than 37% of all bets are below \$20 while the average individual bet is about \$35. There are more than 660 bets placed per race. This indicates that the pool of bettors is composed of many individuals who bet relatively small, yet not insignificant, amounts. Since these small bets are associated with considerable transaction costs, we conclude that a large majority of bettors in the alpine skiing market are non-professional individuals whose betting strategy might be influenced by changes in media coverage. This interpretation is in line with previous studies of sports market betting who have pointed out that most bettors participate for recreational fun [Lee et al., 2013].

Finally, it is worth mentioning that the average inverse odd for an athlete in our estimation sample is 0.18 and 0.12 for close victories and podium finishes, respectively. In both cases this probability is higher than the true average probability of an athlete winning the subsequent race (0.15 and 0.09, respectively). The difference illustrates the markup charged by the betting agency. The fact that betting odds do not sum up to one reflects that our data is in line with the literature on betting markets [Levitt, 2004].[1]

5.3 The Identification Problem

5.3.1 Selection on Observables

The central research question of this paper is whether media attention affects betting market outcomes. A naive way to test this hypothesis is to run an ordinary least squares regression, assuming selection on observables. If we possessed all variables that affect outcomes, we could simply use our data set and fit the

[1] In a standard betting market model for a podium finish, the agency faces a classical monopolist problem with its markup given by $m = p_3 + p_4 - 1$ with p_3 and p_4 being the inverse odds for the 3rd and 4th athlete. Simulating the decision, we find that the agency optimally sets a mark-up larger than zero. Empirically, we can standardize the inverse odds for each race such that the sum of all probabilities is equal to one. The results we obtain are virtually identical to the estimates we show in Table 5.4. The point estimates (standard errors) for the estimation using average odds as the dependent variable are given by -0.026 (0.028) for the podium specification and by -0.011 (0.031) for the victory specification.

empirical model:

$$Y_{i,t+1} = \phi_i + \tau \text{TOP\_RANK}_{i,t} + \mathbf{X}_{i,t}\,\phi + \varepsilon_{i,t} \tag{5.1}$$

where $Y_{i,t+1}$ denotes the outcome variable which can be media attention, performance, or betting market outcomes of athlete i in race $t+1$. Note that media attention is measured by the number of articles on the day after race t. The coefficient of interest, τ, indicates the impact of a top rank, either a victory or podium finish. Finally, $\mathbf{X}_{i,t}$ denotes a vector of athlete i's observed characteristics (gender, experience, prior successes, competitors), ϕ_i is an athlete-fixed effect, and $\varepsilon_{i,t}$ denotes the standard error clustered at the athlete level.

— Table 5.2 about here —

The results of estimating this model are shown in Table 5.2. We observe that top ranks (i.e., podium and victory) are positively correlated with media attention on the day after a competition. Moreover, the estimates show a negative relationship between high ranking positions and average odds. For the risk and performance measures, we find that a victory is associated with a higher probability to win the subsequent race. All these correlations are robust to the inclusion of several control variables as well as athlete-fixed effects. However, these results rely on the selection on observables assumption and should thus not be interpreted as causal effects.

5.3.2 Identification Strategy

The fundamental problem with estimating equation (5.1) is that ranking positions—including victories and podium finishes—are not randomly assigned across athletes. Although we can use a wealth of information to proxy for unobserved variables, including fixed effects to net out skill differences, we do not have information on, for example, injuries prior to the competition. In this section, we suggest a novel identification strategy to overcome this problem. The key idea of our approach is that in close races it is often a tiny margin that determines athletes' positions. If, for example, the time difference between two athletes ranked third and fourth is only a tenth of a second, we document that this can be attributed to random weather shocks. Hence, we can use those athletes that achieve a top rank (i.e., victory or podium finish) in a close race to estimate the causal effect on various outcomes. Throughout our analysis, we restrict the sample to races within season and discipline. This is necessary because times vary substantially across disciplines and seasons are separated by more than half a year.[1] When limiting our sample to close races in which top ranks are arguably randomly assigned, we have to exclude most combined competitions and focus on slalom, giant-slalom, super-G, and downhill races.[2]

The focus of our analysis is on top ranking positions, namely victories and podium finishes. We motivate this by the fact that individual tournaments reward mainly athletes on the podium, and more specifically the winner of a race. This is reflected by FIS World Cup points, substantially higher prize money and increased media attention. In comparison with World Cup victories, however, podium finishes have a couple of advantages. First, we can draw on significantly more observations. Furthermore, when estimating the effect of a quasi-random victory, implicitly all observations in the control group finished on the podium. Hence, those athletes are also treated with a top rank, although the "treatment dose" is arguably lower. Thus, when using a victory as treatment, we only estimate the additional effect compared to a podium finish.

5.3.2.1 Random Top Ranks in Alpine Skiing

In contrast to other fields of sport, World Cup alpine skiing offers a unique feature that allows us to determine quasi-random top ranks. We illustrate this by a simple thought experiment. Assume there are only three variables that determine athlete i's final race time, $T_{i,t}$, in a given race t: the time-invariant skill of athlete i, denoted by θ_i, her training or fitness level, denoted by $\lambda_{i,t}$, and a noise parameter $n_{i,t}$ that captures all kinds of random shocks such as weather or snow conditions

[1] Typically, the last race of a World Cup season is in March, while the first race of the next season takes place in October.

[2] In combined races, time differences are larger and there is a higher variance over time. This is mainly because combined races tend to be longer and have a smaller group of starters, which makes competition less fierce. Furthermore, there are only about five combined races per year as opposed to the other disciplines with about eleven.

that can be heterogeneous or homogeneous across athletes.[1] This setting allows us to write the time of athlete i in race j as a function of her skill and training levels as well as some random noise:

$$T_{i,t} = f(\theta_i, \lambda_{i,t}, n_{i,t}). \tag{5.2}$$

Moreover, her position in the final ranking, $P_{i,t}$, is a function of her own time as well as her competitors' times (all athletes k except i):

$$P_{i,t} = g(T_{i,t}, T_{k,t}) = g(\theta_i, \lambda_{i,t}, n_{i,t}, \theta_k, \lambda_{k,t}, n_{k,t}) \qquad \forall k \neq i \tag{5.3}$$

By means of this equation, we can illustrate why quasi-random top ranks are possible. Usually, skill differences explain most of the variation in ranking positions. This does not, however, imply that ranking positions are entirely driven by skill levels. Figure 5.3 in the Appendix depicts a histogram of winners' and third-ranked athletes' ranking positions in the previous race. The fact that 40.3% of current winners and 22.6% of current third-ranked athletes achieved a podium in their past race documents positive serial correlation of our success measures. Yet the spread of the distribution reveals substantial variation in ranking positions. This challenges the idea that skill differences entirely determine ranking positions. In particular, if two athletes have almost identical race times, random fluctuations in the noise term become critical. Variations in $n_{i,t}$ can reduce athlete i's race time sufficiently to overcome skill and training deficits. In this way, a less skilled athlete can be lucky and draw a very low $n_{i,t}$ which enables her to achieve a better race time than a more skilled competitor.

For our estimation, the key identifying assumption is that the noise parameter $n_{i,t}$ has sufficiently large effects on individual race times in order to randomly assign relative ranking positions in close races. In skiing, the individual noise term, $n_{i,t}$, is comprised of several components. First, alpine skiing is an outdoor event and thus wind and weather conditions vary significantly over the course of a single race. Most notably changes in snow, wind, and sight alter individual prospects

[1] We assume skill to be persistent. However, we take into account that injuries may affect athletes' ability to exploit their skills by including the time-variant fitness parameter.

of success and can also lead to cancellation if race conditions are considered to be a serious risk for the athletes.[1] Yet, the mere presence of unstable external conditions does not lead to cancellation and is broadly accepted as a natural source of variation among competitors. The impact of random wind, weather, and snow conditions is amplified by the fact that individual race times critically depend on the performance in key sections of the course. An error in these sections not only leads to an immediate time loss but also affects speed, and thus time, in the following sections.

It is crucial for our analysis to test whether there is any bunching of data around the thresholds which determine who wins a race or finishes on the podium. Following McCrary [2008] there should be a smooth distribution of observations around the cutoff. Otherwise there might be a distorting factor we need to address. Figure 5.4 in the Appendix illustrates that the number of observations is in fact smooth at the cutoff for both treatments, victory and podium. This supports the assumption that athletes are not systematically located around the threshold.

5.3.2.2 Bandwidth Choice

What is a close race and what time difference can be considered random? An important identifying assumption of our research design is that treated and non-treated athletes are not systematically different with respect to pre-determined covariates. If finishing on the podium in a close race is driven by skills instead of luck, our approach would not allow us to assess the causal effects of quasi-random top ranks. To address this concern, we compare the characteristics of treated and non-treated athletes. We do this in two steps. First, by means of Figure 5.5 in the Appendix, we show that athletes who win or finish on the podium are *in general* different from less successful athletes when comparing prior success. However, we find that there are no significant differences with respect to prior success when considering *close races*. This indicates that those who win or make it to the

[1] The following excerpt about the performance of U.S. competitor Bode Miller in 2009 illustrates the impact of wind. “Miller, a two-time overall World Cup winner, finished ninth Saturday as the Saslong downhill in Val Gardena, Italy, marked its 40th year. His performance was affected by a strong headwind that whipped up just as he and the other top contenders took the course”, *New York Times*, December 20, 2009.

podium in a close race are not systematically more skilled. Once we restrict the sample to tournaments in which the time difference between successful and non-successful athletes, the running variable, is less than 0.15 seconds, the differences in pre-determined characteristics are insignificant. Note that we calculate the running variable as the time difference to the third-placed athlete for all athletes who did not finish on the podium and, equivalently, the time difference to the athlete on the fourth position for all athletes on the podium.[1] This leaves us with a negative running variable for athletes on the podium and a positive running variable for athletes who are not on the podium.

— Table 5.3 about here —

In Table 5.3, we compare athletes on the podium with those who missed it by up to fifteen hundredths of a second. There is no significant difference in any of the observable athlete characteristics. The top-ranked athletes in close races are not more experienced, successful, or risk-loving than their contestants in the control group. Moreover, we find no difference in their competition, media attention prior to the race, or betting market outcomes for the race which determines who is in the treatment and control group.[2] Importantly, treated and non-treated athletes are also not different in terms of the probability of competing in the following races, which rules out the possibility that lower ranked athletes are discouraged from participating in subsequent races. However, it is important to note that treated and non-treated athletes obviously become systematically different if we extend the bandwidth. If we include athletes trailing the podium by a large time difference it is no longer plausible to consider the podium finish a result of 'luck'. Hence, we have to restrict the sample to observations with sufficiently small time differences in order to exploit 'random top ranks'. The decision to choose 0.15 seconds as the bandwidth for our estimations is the result of a trade-off: We can use more observations with a larger bandwidth but the allocation of ranking positions (and thus media attention) is only plausibly random for small

[1]For the victory treatment, the running variable is the time difference to the winner for athletes who did not win and the time difference to the second-ranked athlete for all winners.

[2]Note that the balance tests remain unchanged when we include the running variable in the regression as documented in Table 5.1 in the Appendix.

bandwidths. As we show in Section 5.4.3, our specific choice of bandwidth does not affect the empirical findings.

5.3.2.3 Econometric Specification

For our estimation we assume that, except for the treatment, there is no reason why subsequent outcomes ($Y_{i,t+1}$) like media attention, performance, or betting odds should be a discontinuous function of the race time. We support this assumption using a large set of balance tests (cf. Tables 5.3, 5.1, and Figure 5.5). Hence, any discontinuity in the outcome variable at the cutoff level c_t is identified as the causal effect of the treatment. We estimate the treatment effect τ by fitting the linear regression

$$Y_{i,t+1} = \phi_i + \tau D_{i,t} + \beta(T_{i,t} - c_t) + \gamma[D \times (T_{i,t} - c_t)] + \mathbf{X}_{i,t}\,\delta + \varepsilon_{i,t} \tag{5.4}$$

where ϕ_i is an athlete-fixed effect, $D_{i,t}$ indicates treatment (victory or podium), $[D \times (T_{i,t} - c_t)]$ allows for different slopes on each side of the cutoff, $\mathbf{X}_{i,t}$ is a vector of control variables, and $\varepsilon_{i,t}$ is the standard error term which we cluster at the athlete level.[1] For the outcome variable, $Y_{i,t+1}$, we use media attention after race t, performance and risk-taking behavior in race $t+1$ as well as both betting odds and the number of bets in race $t+1$. Note that the inclusion of covariates could in principle improve the precision of the estimation [Frölich, 2007]. We explored this possibility but did not find notable differences in the estimates. Moreover, squared and cubic terms of $(T_{i,t} - c_t)$ can be included to allow for a nonlinear relationship.[2]

[1] All control variables are not log-transformed to avoid losing about 18 percent of observations. The results are, however, not sensitive to this decision. The inclusion of athlete-fixed effects might be problematic in the sense that information is used for the estimation that was not available at time t. We thus re-ran the estimation with athlete-season-fixed effects and—in line with our main results—find a positive effect on media attention and no discontinuity in betting odds.

[2] The work by Hahn, Todd and van der Klaauw [2001] as well as Gelman and Imbens [2014] suggests to use local linear regression in an RD setting. In particular, high-order polynomials of the forcing variable should not be used. Thus, we omit squared and cubic terms of $(T_{i,t} - c_t)$ in our baseline regressions. However, the results are not sensitive to the specification.

5.4 Results

5.4.1 Effect on Media Attention

We first test whether media attention is affected by a top rank in a close race. Fitting the model specified in equation (5.4), we use the number of newspaper articles mentioning an athlete's name on the day and during the week after the race. All our regressions include athlete-fixed effects and thus use within-athlete variation. Hence, in the estimation we compare the same athlete who achieves a high ranking position (i.e., victory or podium) in one race but not in the other. The great advantage of our data set is that we have enough observations to test whether the same person receives different levels of attention by the media if she performs only marginally better than her competitors.

— Table 5.4 about here —

Table 5.4 shows that those athletes who finish on the podium in a close race are mentioned about 39.5% (= 2.95/7.48 with 7.48 being the average in the control group) more often than those who barely miss the top three ranks. For a close victory we find an additional 15.4% increase.[1] Note that the point estimate for victory is considerably smaller than the estimate for a podium finish because most athletes in the control group also finished on the podium and thus benefited from increased media attention for top-ranked athletes. The positive impact of a high ranking position in close races is still present when counting all articles published during the seven days following the race as indicated in the second column in Table 5.4. By means of Figure 5.4, we can visualize the discontinuity in media attention around the cutoff time by plotting the average absolute media attention on the day after the competition.

— Figure 5.4 about here —

[1] When restricting the sample to the period 2006–2014, we still find a large positive and significant effect of a close podium finish (and of a close victory) on media attention. The point estimates when using media attention on the day after the race are given by 4.75 and 3.66 for podium and victory, respectively.

The difference in media attention between top-ranked and not top-ranked athletes is present even when considering only close races in which the time difference was tiny. These significant differences indicate that top ranks introduce a sharp discontinuity in *absolute* media attention. However, we also have to examine whether there are differences in *relative* media attention. Not only do athletes' sponsorship contracts depend on how many times they are mentioned compared to their competitors, it is also very likely that relative media attention affects bettors' expectation about who is going to win the next race. To test the effect on relative media attention, we define $m_{i,t}$ as athlete i's share of total media attention among all athletes within our preferred bandwidth of 0.15 seconds. For example, suppose racer A wins race t, racer B trails her by less than 0.15 seconds and all other athletes have a larger distance to the winner. Then we count all newspaper articles that mention racer A and we do the same for racer B. If the winner is mentioned in 60 articles and the second in 40 articles, we have $m_{A,t} = 0.6$ and $m_{B,t} = 0.4$, respectively. Formally, $m_{i,t} = A_{i,t} / \sum_s A_{s,t}$ with s indexing all athletes who won race t (or achieved a podium finish) as well as all those missing the victory (or podium) by 15 hundredths of a second or less. Using this measure of relative media attention, we find a discontinuity of 11 percentage points at the threshold to the podium which corresponds to a 32% increase in media attention compared to the control group (see Table 5.2 in the Appendix). The equivalent relative increase in media coverage for achieving a victory is 19%. Overall, these estimates strengthen the claim that top ranks causally affect media attention.

5.4.1.1 Media Attention Before and After a Race

Despite the fact that allocation to treatment in our case is arguably random for close victories and podium finishes, the question remains whether it is the very success that affects media attention. In order to investigate this, we compare the media attention of top ranked and lower ranked athletes at various points in time before and after the race.

In the balance tests shown in Table 5.3, we observe that *before* the race, successful and non-successful athletes receive very similar levels of attention by the media. This is not surprising, since treatment and control group seem to be balanced in terms of experience and prior success. However, *after* the race the control group—those who did not win the race or did not finish on the podium—receive significantly less media attention. Although all athletes in the treatment and control group show a very similar performance, media attention is tilted heavily in favor of those finishing on the top three positions of the ranking. A second observation is that, as expected, the difference in media attention is less pronounced when considering longer time periods. However, even when we examine all articles published in the week after a competition, we observe a significant gap between successful and non-successful athletes. This gap remains significant if we subtract articles published on the day after a race. In the long run, the gap in media attention subsides. This is shown in Table 5.2 in the Appendix and driven by the fact that within a month the next tournaments took place.

5.4.2 Effect on the Betting Market

When estimating the effect of media attention on betting markets it is important to first test whether the true probability of athletes to succeed is altered by achieving a top rank. If one-time great successes have an effect on performance, we also expect a difference in betting odds because bettors update their beliefs about the future performance of athletes based on the ranking positions. In this case, it would not be possible to disentangle the effect of higher media attention from the increased performance effect. However, if rankings have no effect on athletes' performance, the effect on betting odds is likely to depend on the discontinuity in media attention introduced by the ranking scheme.

5.4.2.1 Effect on the True Probability

To test whether athletes respond to what ranking position they achieve in one tournament, we investigate their risk-taking behavior and performance in the subsequent race. First of all, it is not surprising to find a substantial amount of serial correlation in our data set. The correlation between today's distance to the podium (victory) and the average position in the next race is 0.23 (0.14).[1] The key question, however, is whether there is a discontinuity around the cutoff. We test this by plotting the average probability of achieving a victory in race $t+1$ around the time cutoff for a victory and podium finish in race t. The variable capturing whether an athlete wins the next race is our preferred measure of performance because it corresponds to the very event for which bettors can place bets.

— Figure 5.5 about here —

The results shown in Figure 5.5 indicate that there is no discontinuity at the threshold. Neither a victory nor a podium finish in a close race has a causal effect on the probability of winning the next race. One may argue that this finding is driven by the choice of the dependent variable. Very few athletes win a race in their careers and thus victory might be an imprecise measure of performance. To test whether the results are sensitive to the choice of the dependent variable, we use several alternative measures of performance including the probability of a podium finish and the average position in the next race. The results in Table 5.2 in the Appendix confirm that professional ski athletes do not change their performance following a top position in the ranking.

We also investigate whether achieving a high ranking position in one tournament changes the risk-taking behavior of athletes in subsequent races.[2] If rankings affect risk-taking, the interpretation of the performance result above would be more difficult to interpret because rankings would alter the composition of athletes who obtain a final race time and thus a performance measure. In the context of our study, it may be expected that athletes change their behavior and perception of risk after a quasi-random top rank. Athletes may misinterpret one-time victories as signals of high ability. As a consequence, they might act too ambitiously in subsequent races, leading to an increase in the probability of crashes. Hoelzl and Rustichini [2005], for example, find in an experiment that overconfidence becomes important when monetary payments are at stake. Yet our results

[1] Figure 5.6 in the Appendix provides a graphical illustration of this positive serial correlation of performance.

[2] Föllmi, Legge and Schmid [2016] investigate risk-taking in alpine skiing and document that athletes react to small changes in the perceived time difference to the leader by changing their risk strategy between two runs in the disciplines slalom and giant slalom.

indicate that athletes on a top ranking position are not more likely to finish the subsequent race. In addition, their position and the probability of finishing on the podium in the next race do not differ from those athletes who missed the top ranks (i.e., victory or podium) by a small margin. From an econometric point of view, this finding supports our estimation of the effect on performance without taking into account an attrition bias by using a principal stratification framework [Frangakis and Rubin, 2002].[1] Finally, we consider the overall time in race $t+1$ as a plausible outcome variable.[2] Again we find no difference between successful and non-successful athletes. The regression results reported in Columns (3) and (4) of Table 5.4 confirm that the point estimates of top ranks on risk-taking and performance are very close to zero and far from being statistically significant.

Before turning to the impact of top ranking positions on the betting market we briefly summarize our key findings. Whether an athlete wins or finishes on the podium in a close race has a substantial effect on the amount of media attention she receives. However, it does not affect her risk strategy and performance in the next race. Hence, if betting odds reflect true probabilities there should not be any significant difference between top ranked and lower ranked athletes.

5.4.2.2 Effect on the Betting Market

Given the large positive effect of top ranking positions on media attention, it appears likely that public expectations about the performance of top-ranked athletes in the next race increase. A natural way to test whether public expectations discontinuously rise as a consequence of an exogenous shift in media attention is the analysis of betting data. Information on betting behavior should reflect prior expectations of bettors in an incentive-compatible way because a betting agency that deviates from bettors' expectations would either incur losses (if betting odds are too high) or attract no bets (if odds are too low). We use betting odds from

[1] Consider, for example, the case in which a one-time success had a negative effect on the survival probability. This would indicate increased risk-taking among successful (i.e., treated) athletes. We discuss this concern in detail in the Appendix B.

[2] In contrast to using ranking positions in race $t+1$, the advantage of the race time is that we do not have to address the problem of potentially violating the stable unit treatment value assumption (SUTVA) which is necessary for the estimation of causal effects. In Appendix C we provide a detailed discussion.

Betfair, the world's largest Internet betting exchange, for the period of 2006–2014 to explore the impact of top ranks in close races on betting market outcomes.

The distribution of betting odds across ranking positions is shown in Panel (b) of Figure 5.1. We observe a negative gradient with better ranked athletes facing higher inverse odds in the next race. There is also a pronounced difference for close winners and non-winners as well as between podium and non-podium finishers. However, we want to examine whether an athlete who randomly achieved a top rank faces different odds in the subsequent race. Fitting the empirical model of equation (5.4) with betting odds and the number of bets as dependent variables, we obtain the results reported in columns (5) and (6) in Table 5.4. For both outcomes, the estimates are very close to zero and fall short of conventional significance levels. These results are consistent with a graphical inspection provided by Figure 5.6.

— Figure 5.6 about here —

Based on these findings there are two conclusions. First, athletes who achieve a top ranking position in a close race receive higher subsequent media attention but not a significantly higher number of bets. The results remain unchanged if we use the volume of bets instead of using the total number of bets. The second conclusion from our estimation is that the absence of any discontinuity in betting odds matches the fact that randomly assigned higher ranking positions do not increase the true probability of winning the next race. In this sense, columns (5) and (6) of Table 5.4 support the idea of using markets for predictions [Wolfers and Zitzewitz, 2006].

5.4.2.3 Betting Market Efficiency

We use our data set and investigate in more detail to what extent betting odds reflect true probabilities. In panel (a) of Figure 5.7, we show the estimated effect of finishing on the podium on the true probability of achieving a victory in the next race. The estimation is conducted for all bandwidths between 0.05 and 4.00 seconds.

— Figure 5.7 about here —

We observe that once we include athletes trailing the podium by more than a full second, those who finished on the podium are significantly more likely to win the subsequent race. Interestingly, panel (b) shows that there is also a significant difference in (inverse) betting odds between athletes on the podium and other athletes once the bandwidth is larger than one second. These results suggest that the betting market mimics true probabilities of future events.

5.4.3 Robustness Checks and External Validity

In this subsection, we discuss the robustness of our main results. In a first step, we investigate whether top ranks have an effect on betting odds at specific points in time of the betting process. Second, we discuss the bandwidth choice in our regression discontinuity design. Third, we explore whether we obtain the same results in an instrumental variables framework using betting market behavior as outcomes, media attention as treatment, and top ranking positions as instrument. Finally, we address concerns about our source of media data, market liquidity, and selective participation.

5.4.3.1 Initial, Average, Final Odds

To shed light on the betting agency's behavior and the bettors' corresponding response, we analyze initial and final betting odds. First, it might be that the estimated effect on the average odd is driven by differences in the initial odds which are exclusively determined by the betting agency. These odds change over time until the next race starts. The changes in odds are driven by new information (e.g., news about an athlete being handicapped) as well as the number of bets placed. If more and more bettors want to buy bets that athlete i will win the next race, her odds are likely to decrease. We have data not only on the initial odd but on all odds that were offered for a given individual and race. Hence we can separate the initial odd, the average odd and the final odd. This wealth of information helps us understand whether the betting agency attempts to increase its profit by offering lower odds as a result of higher media attention. We can also investigate whether such a strategy is successful or whether the market adjusts the price. Table 5.2 in the Appendix shows that the betting agency offers 'fair' odds in the first place. The point estimates for the average and final odds are very similar to those we obtain using the initial odds. This finding supports the idea that bettors' expectations are not biased as a result of selective media attention.

5.4.3.2 Bandwidth Choice

Throughout our empirical analysis we use the concept of random top ranks to identify the causal effects of media attention. As we explained in Section 5.3, it is crucial to focus on close races to overcome the identification problem. A central question in this regard is what time difference between two athletes can be attributed to random shocks. This means that it remains a priori unclear what bandwidth we should use in our estimations. In all regressions so far we have used a bandwidth of 15 hundredths of a second. The bandwidth choice was primarily based on the results of the balance tests in Table 5.3 as well as on the comparison of prior success for top-ranked and other athletes depicted in Figure 5.5 in the Appendix. We can illustrate the magnitude of 0.15 seconds by plotting the distribution of time differences to the podium. Figure 5.7 in the Appendix shows the distribution as well as a vertical line for the bandwidth we use in our estimation. When restricting the sample to those athletes trailing the podium by 0.15 seconds or less, only 10% of the sample are included.

In order to investigate the robustness of our empirical results, we re-run the RDD estimation (equation 5.4) using different bandwidths ranging from 0.10 to 0.50 seconds. In our preferred specification we include athlete-fixed effects to hold constant all individual-specific covariates. Figure 5.8 in the Appendix depicts the effect of a podium finish on media attention, performance, and betting odds using different bandwidths. There are two notable observations. First, the point estimate is very stable irrespective of the bandwidth choice. Second, when decreasing the sample size the confidence intervals become very large. Using a bandwidth of 0.10 seconds, for example, leaves us with only 304 observations in the betting odds estimation that includes athlete-fixed effects. Our preferred bandwidth of 0.15 seconds is the result of the trade-off between bias and precision: On the one hand, we can use more observations with a larger bandwidth. On the other hand, the allocation of ranking positions (and thus media attention) is only plausibly randomized for small bandwidths.

5.4.3.3 Instrumental Variables Estimation

Thus far we have analyzed the effects of media attention on the betting market using separate regressions that relate randomized top ranks to media and betting outcomes. While top ranks create sizable discontinuities in media attention, the results of the intention-to-treat (or reduced form) regression suggest that top ranks do not create a difference in betting odds. This indicates that the effect of interest, the parameter on media attention in the second-stage regression with betting market outcomes as dependent variable, is absent [Angrist and Krueger, 2001]. To explore this (absent) effect in more detail, we estimate a two-stage least-squares regression using our betting market variables as outcomes, media attention as treatment, and top ranks as instrument. We focus on close podium finishes and extend the bandwidth to half a second to avoid suffering from weak instrument problems.

— Table 5.5 about here —

Table 5.5 reports the results of these regressions using the average inverse odds as well as the total number of bets in the subsequent race as dependent variables. All regressions include the distance to the podium as well as athlete-fixed effects. The estimates highlight that the effect of the media on betting market outcomes is very close to zero and not significant. These findings add to the previous results using reduced form regressions suggesting that there is no effect of media attention on betting behavior.

5.4.3.4 Media Data from NewsLibrary

In order to examine the robustness of our empirical findings about the effect of rankings on media attention, we address the fact that our source of media data (Swissdox) is a Swiss-based, largely unknown source. An alternative source is the American newspaper archive NewsLibrary. For all top-15 athletes in all races between 1992 and 2014 we use NewsLibrary and count the number of articles mentioning the athletes' names. As before, we count the articles before and after the race over different time periods. We obtain a distribution of media attention across ranks that is very similar to the one that we observed for the Swissdox data. There is a notable gap between the winner and the runner-up as shown in Figure 5.1 in the Appendix. Even more noticeable is the difference between the amount of media attention received by the third-ranked athlete compared to those athletes who missed the podium. We can use the data from NewsLibrary to repeat the OLS and RDD estimation. The estimates are very similar to the ones we obtained when comparing athletes around the podium cutoff as we document in the Appendix Figure 5.9.

5.4.3.5 Media Attention versus Ranking Lists

Given that both Swissdox and NewsLibrary also contain ranking lists, one could argue that we do not find an effect of media attention on betting markets simply because our data is a weak proxy for true media attention. We address this potential issue by an additional scraping procedure which eliminates all ranking lists from our sample.[1] Note that this is a conservative approach as it removes all ranking lists but also some additional articles which makes it more challenging to obtain significant effects in the first stage regressions.

Estimates shown in Table 5.3 of the Appendix indicate that a close podium finish or victory has a positive and significant effect on media attention after a race. Compared to our main results, we find that the relative impact of success on media attention is even larger if we exclude ranking lists.

5.4.3.6 Liquidity of the Betting Market

One potential concern about the interpretation of our results is that the alpine skiing betting market is very thin or illiquid. This might be worrisome for two reasons. First, in a very illiquid market betting agencies might not adjust their odds

[1] Technically, we again search for an athlete's name during various periods around the race date. However, we leave out all articles that contain ranking positions such as "7." or "8." which every ranking list includes.

in response to participants' beliefs. If only few individuals participate, betting agencies may anticipate that there is very little dynamic movement in the market and consequently do not take into account bettors' beliefs but rather charge a fixed mark-up. As a result, the finding that media attention has no effect on betting odds stems from the supply (betting agencies) and not—as we suggest in this paper—from the demand side (bettors). Second, in very illiquid markets it might not be profitable to adjust odds as a response to bettors' beliefs because the betting volume is simply to small. Moreover, the predictive power of betting markets might be limited in the absence of liquidity because information is not efficiently and timely aggregated if only few individuals participate.

To address this concern, we explored several aspects of the betting data. First, even though the average individual bet is only about \$35, the total volume per race is \$23,313 and the average number of bets is 660. Furthermore, betting odds change substantially over time. For more than 99.7% of all bets there is at least one change of odds. The average number of changes is eight. In addition, these changes are substantial: the standard deviation of inverse odds is 0.07 which is about a quarter of the overall mean of 0.27. Second, we probe whether our main result that media attention has no effect on betting odds also holds in a very liquid market. To do so, we split our sample by two measures of market liquidity, namely the total volume and the number of bets per race. As indicated in Table 5.4 in the Appendix, we find no effect of media coverage on betting odds in less liquid but also in highly liquid markets.

These results are in line with the literature that investigates the efficiency properties of betting markets. Several studies have explored whether prediction markets can be manipulated.[1] Camerer [1998] finds that placing temporary bets in order to manipulate horse race markets is unlikely to affect odds in the long term. Similarly, Rhode and Strumpf [2004, 2008] examine various prediction markets and conclude that these markets cannot be systematically manipulated beyond short time periods. Finally, Hanson, Oprea and Porter [2006] as well as Oprea et al. [2008] provide experimental evidence that a group of traders cannot

[1] Meng [2016] provides a discussion of how the illiquidity of a prediction market might be a serious concern.

manipulate the accuracy of forecasts. The only exception to this literature is a study by Rothschild and Sethi [2015] which finds some evidence of possible manipulation in the 2012 Intrade U.S. presidential prediction market.

5.4.3.7 Selective Participation

Another important potential concern is that athletes who finished on the podium might participate in different races than athletes who barely missed the podium. In particular, it might be that those athletes on the podium compete in more races with fiercer competition. To explore this concern, we first present institutional evidence on participation in FIS races and then explicitly test whether athletes selectively participate in World Cup races. In general, participation in the FIS World Cup series is reserved to the best athletes in a certain discipline. Less experienced and less successful athletes are competing in the FIS European Cup and the Continental Cup. In all three series, athletes can win FIS points that allow participation in the World Cup which is the most prestigious series in terms of prize money and media attention. Yet, most athletes in our sample have a high score of FIS points and thus a fixed starting position in the World Cup. Hence, the additional World Cup points from a close podium finish or a close victory should not affect eligibility for the next race. Furthermore, it is important to note that organizers have no power to select the competing athletes except that they can give wild cards to certain athletes who would not be classified (mostly to young local athletes). There might, however, be selective participation due to injuries or strategic reasons.

To explore the effect of a random victory and podium finish on participation due to injuries or strategic reasons, we reran the main estimations using participation as an outcome. As indicated by the first two columns of Table 5.5 in the Appendix, the point estimates are close to zero and not significant. We also examined whether athletes select themselves into races with stiffer competition by regressing the total sum of victories among top 5 athletes in the next race on our treatment indicators, victory and podium. Columns (3) and (4) of Table 5.5 in the Appendix report that there is no significant difference.[1] To explore whether

[1] Note that these results are robust to using alternative measures of competition in the next

athletes select themselves into races with higher prize money, we split the sample for which we have data on prize money into races with above- and below-median total prizes. Again, we find no evidence that previously successful athletes select into races with more prize money. Overall, these results suggest that future participation decisions are not driven by today's ranking positions and thus the sample of racers just above and below the podium cutoff is not selective.

5.5 Conclusion

This paper investigates how media attention affects betting markets. In a first step, we use a novel data set on all World Cup tournaments in alpine skiing between 1992 and 2014 to show that media attention is highly skewed in favor of successful athletes. Even if performance differences are tiny, there is a significant gap in media attention between athletes on the podium and those athletes who miss it. We exploit this discontinuity in media attention to estimate the causal effect on betting market outcomes.

Our results reveal that ranking positions significantly affect the amount of media attention individuals receive after a tournament. Although prior theoretical and empirical work suggests that top ranks increase athletes' self-confidence and goal-setting behavior, we find no effect of high ranking positions on subsequent performance or risk-taking behavior. Since the true probability of winning the next tournament is not affected by a top rank in a close race, we expect to find no difference in betting odds for athletes with different amounts of media attention if the betting market is efficient. Using data from Betfair, we find that increased media attention has neither an effect on average odds nor on the number of bets. The betting agency offers initial odds that reflect the unchanged true probability of athletes succeeding in the subsequent tournament. While Thaler and Ziemba [1988] suggest that bettors can achieve a positive rate of return by placing bets on extreme favorites, we find no such opportunity. Our results suggest, in contrast to Levitt [2004], that bookmakers do not gain from being more proficient at predicting future events than bettors.

race (e.g., number of podiums among top 5 and top 10 athletes).

Our findings with respect to the distribution of media attention are consistent with theoretical studies on markets for attention [Falkinger, 2007, 2008]. In an information-rich world, media have to concentrate information and sport athletes compete for the scarce resource of media attention. Since athletes in World Cup tournaments draw a large share of their earnings from sponsorship contracts, being among the top-3 is of particular importance to get higher media attention which translates into better sponsorship deals.

Overall, our findings add to the growing literature on the economic impact of the media [DellaVigna and La Ferrara, 2015]. We show that media attention is highly biased in favor of successful individuals but this bias neither affects prices nor quantities in the betting market. While our empirical approach adds credible causal evidence on the effects of the media using field data, exploring the relationship between attention and expectations about future events in different settings or in a lab environment appears to be a fruitful area for future research.

Tables and Figures

Table 5.1: Descriptive Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
I. World Cup Alpine Skiing:					
Position	7.96	4.32	1	15	23,761
Victory	0.07	0.25	0	1	23,761
Podium	0.20	0.40	0	1	23,761
Time Distance to Victory	148.73	140.34	0	3,705	23,761
Time Distance to Podium	71.40	107.45	-968	3,220	23,761
Male	0.51	0.50	0	1	23,761
Age	26.42	3.67	16.38	41.68	23,761
Experience (# Races)	101.29	80.02	1	443	23,761
# Victories at Time of Race	3.99	7.62	0	59	23,761
# Podiums at Time of Race	11.18	17.40	0	109	23,761
Finished Next Race	0.82	0.39	0	1	19,430
Position in Next Race	11.66	9.25	1	65	15,845
Time in Next Race	11,199.10	2,675.59	5,307	25,529	15,845
II. Media Attention:					
Articles the Day after a Race	6.62	10.38	0	125	23,761
Articles the Week after a Race	15.07	25.61	0	381	23,761
Articles the Month after a Race	45.65	82.17	0	974	23,761
III. Betting Market:					
Initial Odds in Next Race	17.15	17.52	1.07	200	2,840
Final Odds in Next Race	17.72	25.42	1.30	1,000	2,840
Average Odds in Next Race	17.33	17.45	1.31	200	2,840
Inverse initial Odds in Next Race	0.13	0.12	0	0.93	2,840
Inverse final Odds in Next Race	0.12	0.12	0	0.77	2,840
Inverse average Odds in Next Race	0.12	0.11	0	0.76	2,840
Volume of a Bet in Next Race (in Dollars)	378.25	861.53	0.06	10,290.18	2,840
Number of Bets in Next Race	12.18	15.18	2	213	2,840

Note: The table presents descriptive statistics for all variables used in the empirical analysis, covering all athletes with a final rank between 1 and 15. Panel (I) presents information on the Alpine skiing data set. Time is always measured in hundredths of a second. Data on media attention in Panel (II) is drawn from our Swissdox database. Panel (III) is based on Betfair. Note that one bet observation is defined as a bet on a specific event ("athlete A wins race X"). The number of bets is the total of all individual bets for a specific event. The volume of bets is also defined at the bet observation level. The time period is 1992–2014 for the World Cup and media data and 2006–2014 for the betting data.

Table 5.2: Ordinary Least Squares Regression

	Media Attention		Performance		Betting Market	
	Day After	Week After	Survival	Victory	Inv. Odds	# Bets
Mean of dep. var.:	11.02	23.56	0.84	0.12	0.16	2.22
I. Podium:						
Podium	5.302***	7.765***	0.016	0.004	0.001	0.095
	(0.377)	(0.701)	(0.011)	(0.013)	(0.006)	(0.058)
Distance to Podium	-0.015***	-0.027***	-0.000	-0.001***	-0.001***	-0.003***
	(0.004)	(0.008)	(0.000)	(0.000)	(0.000)	(0.000)
Experience	0.033***	0.056***	0.000	0.000	-0.000	0.002*
	(0.008)	(0.021)	(0.000)	(0.000)	(0.000)	(0.001)
Prior Podiums	0.100***	0.297***	-0.001	-0.000	0.003***	0.003
	(0.038)	(0.100)	(0.001)	(0.001)	(0.001)	(0.004)
Competition	-0.000	0.013	0.000***	0.000	-0.000	0.000
	(0.002)	(0.008)	(0.000)	(0.000)	(0.000)	(0.000)
Fixed Effects	Athlete	Athlete	Athlete	Athlete	Athlete	Athlete
Observations	7,408	7,408	6,478	6,478	1,625	1,625
R-squared	0.257	0.188	0.003	0.022	0.222	0.067

	Media Attention		Performance		Betting Market	
	Day After	Week After	Survival	Victory	Inv. Odds	# Bets
Mean of dep. var.:	11.02	23.56	0.84	0.12	0.16	2.22
II. Victory:						
Victory	5.151***	10.186***	-0.006	0.026*	0.046***	0.146*
	(0.387)	(0.903)	(0.014)	(0.015)	(0.007)	(0.076)
Distance to Victory	-0.013***	-0.020***	-0.000***	-0.000***	-0.000***	-0.002***
	(0.002)	(0.003)	(0.000)	(0.000)	(0.000)	(0.000)
Experience	0.035***	0.059***	0.000	0.000	-0.000	0.002*
	(0.008)	(0.021)	(0.000)	(0.000)	(0.000)	(0.001)
Prior Podiums	0.086**	0.273***	-0.001	-0.000	0.003***	0.003
	(0.037)	(0.098)	(0.001)	(0.001)	(0.001)	(0.003)
Competition	-0.002	0.011	0.000***	0.000	0.000	0.000
	(0.003)	(0.008)	(0.000)	(0.000)	(0.000)	(0.000)
Fixed Effects	Athlete	Athlete	Athlete	Athlete	Athlete	Athlete
Observations	7,408	7,408	6,478	6,478	1,625	1,625
R-squared	0.257	0.188	0.003	0.022	0.226	0.067

Note: The table shows the results of twelve separate linear regressions using six different dependent variables as indicated in the top rows. The data on media attention is taken from Swissdox. Media attention is measured by the number of articles published on the day and during the week after the race t. In columns 3 and 4, the dependent variable is the probability of finishing (col 3) or achieving a victory (col 4) in race $t+1$. In the last two columns, log inverse average odds as well as the log of the total number of bets for the race $t+1$ are used as dependent variable. The sample includes all athletes who finished in the top-5 in race t. Columns (1) to (4) include tournaments from 1992-2014 while columns (5) and (6) are based on 2006-2014. Numbers in brackets indicate standard errors clustered at the athlete level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 5.3: Balance Tests for Close Podium Finish

	Mean Value			
	Treatment	Control	Difference	p-value
A: Athlete Characteristics				
Male	0.55	0.56	-0.00	0.98
Experience	114.80	117.50	-2.70	0.48
Average Survival	0.93	0.93	-0.00	0.89
Survival in Last Race	0.86	0.87	-0.01	0.64
Number of Victories	5.63	5.53	0.10	0.76
Number of Podiums	15.58	15.33	0.25	0.76
First Prize Possible	0.60	0.59	0.01	0.50
B: Competition				
Total Podiums among Top 5	76.49	76.72	-0.22	0.92
Total Victories among Top 5	31.33	31.40	-0.07	0.93
Total Podiums among Top 10	135.18	135.38	-0.19	0.95
C: Media Attention				
Day Before Race t	5.79	5.34	0.45	0.23
Week Before Race t	15.26	14.74	0.51	0.60
Month Before Race t	49.69	48.76	0.93	0.78
D: Betting Market				
Average Odd Race t	13.25	14.03	-0.78	0.50
Volume in Race t	378.34	388.55	-10.21	0.90

Note: The table shows mean comparisons (t-tests) for all relevant pre-treatment variables. The sample includes all athletes within a bandwidth of 0.15 seconds around the podium. Experience is measured by the total number of races prior to the race, survival in last race is the indicator for successfully finishing in the preceding race, victory and podium measure the total number of an athlete's victories and podiums prior to the race. Media attention is measured by the total number of articles mentioning an athlete's name in the Swissdox archive. The sample in A–C includes tournaments from 1992-2014 while part D is based on 2006-2014.

Table 5.4: RDD for Random Top Ranking Positions

	Media Attention		Performance		Betting Market	
	Day After	Week After	Survival	Victory	Inv. Odds	# Bets
Mean of dep. var.:	9.36	20.83	0.85	0.09	0.12	1.96
I. Close Podium:						
Podium	2.951***	4.735***	0.030	0.023	-0.022	-0.290
	(0.753)	(1.815)	(0.037)	(0.034)	(0.014)	(0.202)
Distance to Podium	-0.146***	-0.226*	-0.002	-0.001	-0.002	-0.018
	(0.045)	(0.115)	(0.003)	(0.002)	(0.001)	(0.019)
Experience	0.035***	0.069**	0.000	0.000	-0.000	0.001
	(0.010)	(0.030)	(0.000)	(0.000)	(0.000)	(0.002)
Prior Success	0.063	0.198	-0.000	0.000	0.002***	0.005
	(0.041)	(0.131)	(0.001)	(0.001)	(0.001)	(0.005)
Competition	0.002	0.019*	0.000	-0.000	0.000	0.001*
	(0.004)	(0.010)	(0.000)	(0.000)	(0.000)	(0.001)
Fixed Effects	Athlete	Athlete	Athlete	Athlete	Athlete	Athlete
Observations	2,227	2,227	1,966	1,966	455	455
R-squared	0.185	0.143	0.004	0.002	0.133	0.029

	Media Attention		Performance		Betting Market	
	Day After	Week After	Survival	Victory	Inv. Odds	# Bets
Mean of dep. var.:	14.91	31.71	0.88	0.15	0.18	2.42
II. Close Victory:						
Victory	2.294**	5.586*	0.000	-0.025	-0.018	-0.174
	(1.149)	(3.034)	(0.045)	(0.064)	(0.020)	(0.192)
Distance to Victory	-0.172*	-0.357	-0.005	0.001	-0.001	-0.019
	(0.091)	(0.253)	(0.004)	(0.004)	(0.001)	(0.025)
Experience	0.059***	0.110***	0.001**	-0.000	-0.000	0.003
	(0.015)	(0.033)	(0.000)	(0.000)	(0.000)	(0.002)
Prior Success	0.021	0.098	-0.004**	0.000	0.003***	0.004
	(0.066)	(0.117)	(0.002)	(0.002)	(0.001)	(0.006)
Competition	0.015***	0.054***	0.001**	0.001**	0.000**	0.003***
	(0.005)	(0.016)	(0.000)	(0.000)	(0.000)	(0.001)
Fixed Effects	Athlete	Athlete	Athlete	Athlete	Athlete	Athlete
Observations	916	916	811	811	221	221
R-squared	0.255	0.197	0.029	0.005	0.227	0.074

Note: The table shows the results of twelve separate linear regressions using six different dependent variables as indicated in the top row. In Part (I) the treatment variable is a podium finish while in part (II) treatment is defined by victory. The sample includes all athletes within a bandwidth of 15 hundredths of a second and to athletes finishing first or second in Part II. The data on media attention is taken from Swissdox. In the last two columns, inverse average odds as well as the log of the total number of bets for the race $t+1$ are used as dependent variable. Columns (1) to (4) include tournaments from 1992-2014 while columns (5) and (6) are based on 2006-2014. Numbers in brackets indicate standard errors clustered at the athlete level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

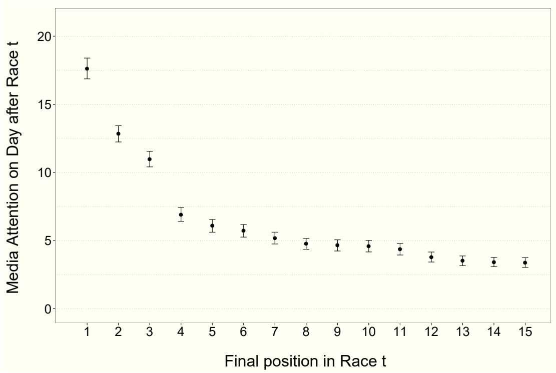
Table 5.5: Instrumental Variables Estimation

	Inv. Odds		# Bets	
Mean of dep. var.:	0.12	0.12	1.95	1.95
Log Media Attention	0.0015	0.0015	-0.0202	-0.0175
	(0.0013)	(0.0012)	(0.0165)	(0.0146)
Fixed Effects	Athlete	Athlete	Athlete	Athlete
Covariates	No	Yes	No	Yes
Instrument	Podium	Podium	Podium	Podium
Observations	1,356	1,356	1,356	1,356
F-value	8.64	10.56	8.64	10.56

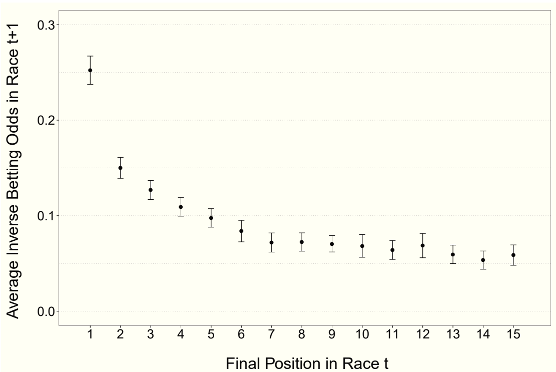
Note: The table shows the results of four separate linear regressions using two different dependent variables as indicated in the top row. The sample includes all athletes within a bandwidth of 50 hundredths of a second. The data on media attention is taken from Swissdox. The dependent variables are log inverse average odds as well as the log of the total number of bets for the race $t+1$. All columns include tournaments from 2006-2014. The F-value is based on the Kleibergen and Paap [2006] rk Wald F statistic. The values of the Cragg and Donald [1993] F statistic are given by 21.00 and 26.16 for the specification without and with control variables. Numbers in brackets indicate standard errors clustered at the athlete level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Figure 5.1: Media Attention and Betting Odds for each Position

(a) Articles on the Day after Race t

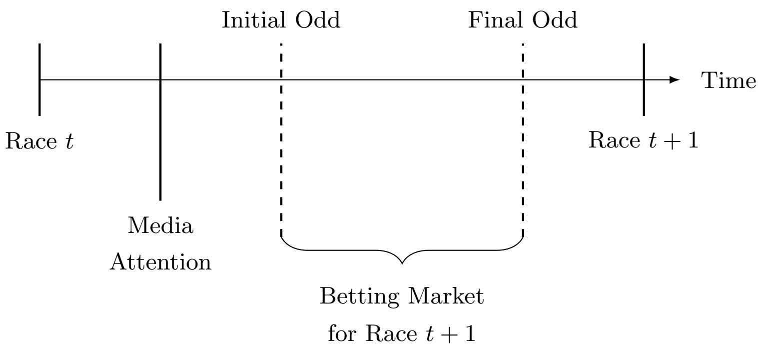


(b) Average Betting Odds in Race $t+1$



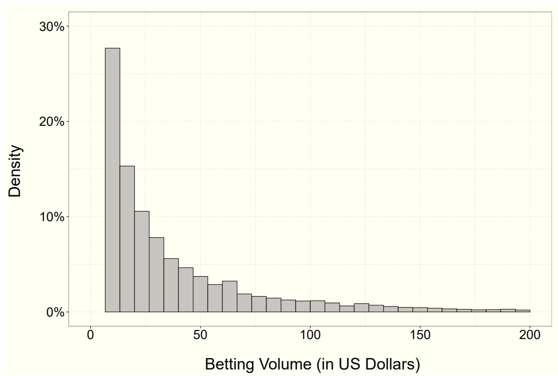
Note: The figure in Panel (a) shows the average number of articles that mention an athlete who finished on a specific position in race t. Media attention is measured on the day after the race took place. In Panel (b), we show the average of all inverse betting odds for bets offered before the race $t+1$ took place for each ranking position in race t. The bars indicate 95% confidence intervals.

Figure 5.2: Illustration of the Timing of Events



Note: The figure illustrates the timing of events. After race t, the betting market for the next race opens with an initial odd for each athlete. We measure media attention on the day after race t took place. The betting market closes with a final odd before race $t+1$ takes place.

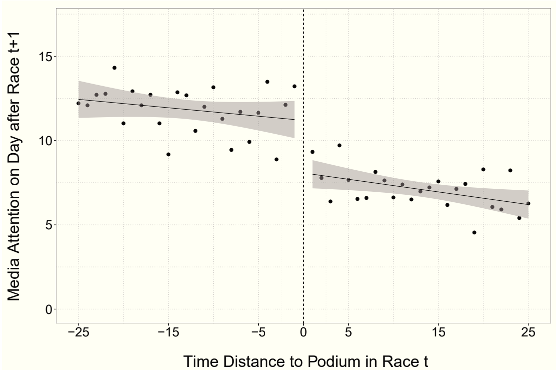
Figure 5.3: Distribution of Betting Volumes



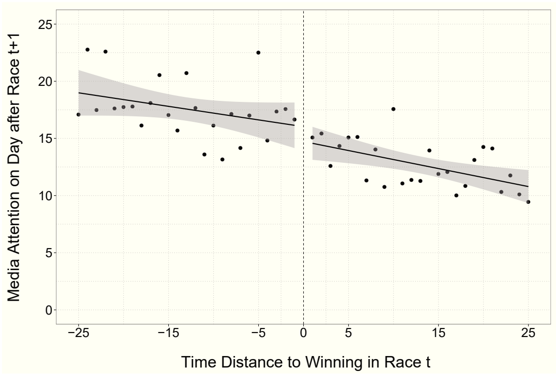
Note: The figure shows the distribution of the volume of each individual bet (in US Dollars). Note that we do not show observations greater than \$200 that account for 2.6% of total observations.

Figure 5.4: Effect of Top Rank on Media Attention

(a) Close Podium Finish



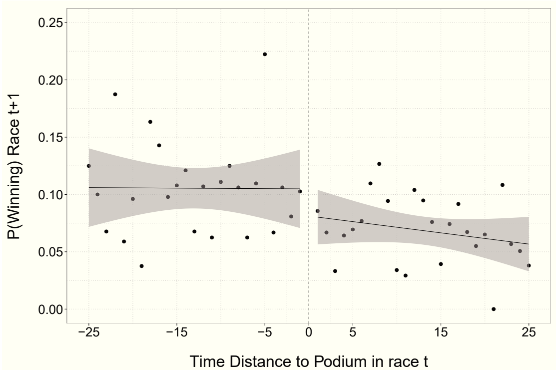
(b) Close Victory



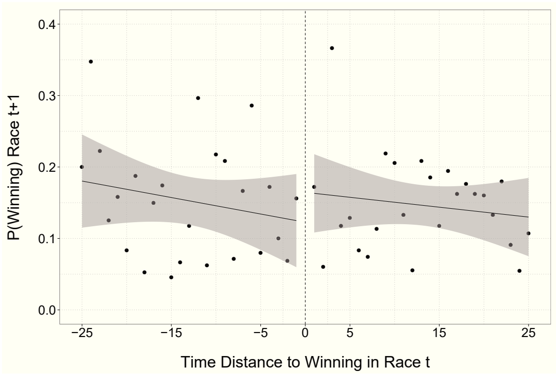
Note: The figures show a linear fit for the number of media articles that mention an athlete (on the day after a race took place) for both treatment (left) and control group (right). In panel (a) we compare victory and non-victory while in panel (b) athletes on the podium are compared with those missing it by 25 hundredths of a second or less. Both variables on the x-axis are expressed in units of hundredths of a second. The grey region indicates the 95% confidence interval.

Figure 5.5: Effect of Top Rank on Performance

(a) Close Podium Finish



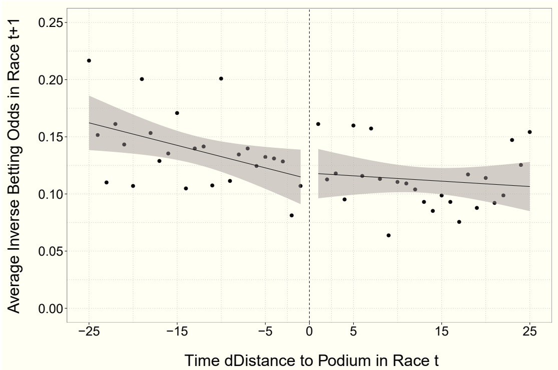
(b) Close Victory



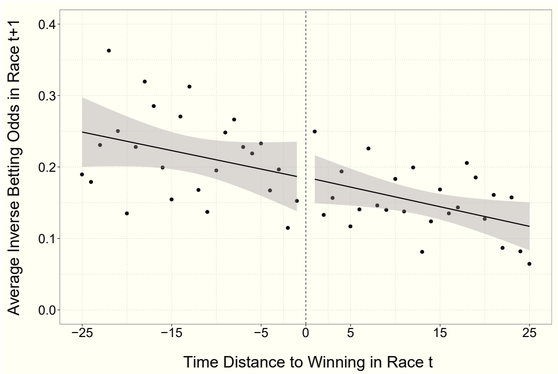
Note: The figures show a linear fit for the probability of achieving a victory in the subsequent race for both treatment (left) and control group (right). In panel (a) we compare victory and non-victory while in panel (b) athletes on the podium are compared with those missing it by 25 hundredths of a second or less. Both variables on the x-axis are expressed in units of hundredths of a second. The grey region indicates the 95% confidence interval.

Figure 5.6: Effect of Top Rank on the Betting Market

(a) Close Podium Finish



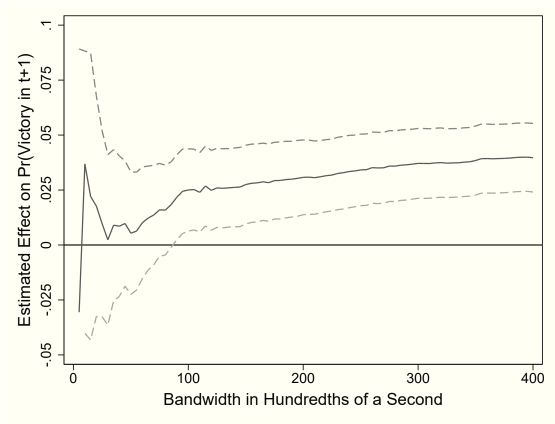
(b) Close Victory



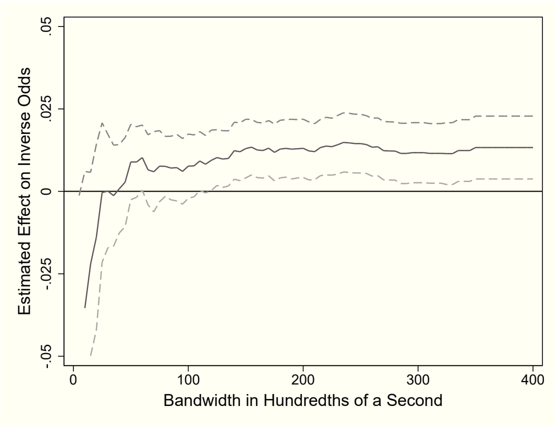
Note: The figures show a linear fit for the average inverse betting odds in the next race for both treatment (left) and control group (right). In panel (a) we compare victory and non-victory while in panel (b) athletes on the podium are compared with those missing it by 25 hundredths of a second or less. Both variables on the x-axis are expressed in units of hundredths of a second. The grey region indicates the 95% confidence interval.

Figure 5.7: Betting Market Efficiency

(a) Victory in Race $t + 1$



(b) Inverse Odds in Race $t + 1$



Note: The figures show the estimated effect of a close podium finish on the probability of winning the subsequent race (Panel a) as well the inverse average odds in the subsequent race (Panel b). Both variables on the x-axis are expressed in units of hundredths of a second. The dashed lines indicate 95% confidence intervals.

Appendix A: Additional Tables and Figures

Table 5.1: Balance Tests for Close Podium Finish Controlling for Distance to Podium

	Mean Value			
	Treatment	Control	Difference	p-value
A: Athlete Characteristics				
Male	0.55	0.56	0.00	0.28
Experience	114.80	117.50	-2.70	0.48
Average Survival	0.93	0.93	0.00	0.87
Survival in Last Race	0.86	0.87	-0.01	0.77
Number of Victories	5.63	5.53	0.10	0.10
Number of Podiums	15.58	15.33	0.25	0.15
First Prize Possible	0.60	0.59	0.01	0.60
B: Competition				
Total Podiums among Top 5	76.49	76.72	-0.22	0.94
Total Victories among Top 5	31.33	31.40	-0.07	0.92
Total Podiums among Top 10	135.18	135.38	-0.19	0.87
C: Media Attention				
Day Before Race t	5.79	5.34	0.45	0.89
Week Before Race t	15.26	14.74	0.51	0.88
Month Before Race t	49.69	48.76	0.93	0.75
D: Betting Market				
Average Odd Race t	13.25	14.03	-0.78	0.45
Volume in Race t	378.34	388.55	-10.21	0.77

Note: The table shows the mean for all relevant pre-treatment variables for treatment and control group as well as the difference. The p-value comes from a regression of the pre-treatment variable on podium controlling for the assignment variable (distance to podium). The sample includes all athletes within a bandwidth of 0.15 seconds around the podium. Experience is measured by the total number of races prior to the race, survival in last race is the indicator for successfully finishing in the preceding race, victory and podium measure the total number of an athlete's victories and podiums prior to the race. Media attention is measured by the total number of articles mentioning an athlete's name in the Swissdox archive. The sample in A–C includes tournaments from 1992-2014 while part D is based on 2006-2014.

Table 5.2: Random Top Ranks and Additional Outcomes

	Media Attention		Performance		Betting Odds	
	Relative	Month	Position	Time	Initial	Final
Mean of dep. var.:	0.34	55.88	9.58	10,790	0.12	0.12
I. Close Podium:						
Podium	0.108***	1.628	-0.775	-233.664	-0.019	-0.018
	(0.022)	(5.285)	(0.773)	(266.732)	(0.016)	(0.015)
Distance to Podium	-0.001	-0.367	0.029	-29.082	-0.000	-0.002
	(0.001)	(0.472)	(0.066)	(19.345)	(0.002)	(0.001)
Experience	-0.000	0.241**	-0.007	-6.028***	-0.000	-0.000
	(0.000)	(0.102)	(0.012)	(1.921)	(0.000)	(0.000)
Prior Success	0.002	0.532	0.006	21.128**	0.002***	0.002***
	(0.001)	(0.522)	(0.047)	(9.648)	(0.001)	(0.001)
Competition	-0.000***	0.004	-0.002	-3.428*	0.000	0.000*
	(0.000)	(0.031)	(0.006)	(1.859)	(0.000)	(0.000)
Fixed Effects	Athlete	Athlete	Athlete	Athlete	Athlete	Athlete
Observations	2,054	2,227	1,698	1,698	455	455
R-squared	0.171	0.120	0.008	0.016	0.082	0.117

	Media Attention		Performance		Betting Odds	
	Relative	Month	Position	Time	Initial	Final
Mean of dep. var.:	0.42	82.35	7.87	10,764	0.19	0.18
II. Close Victory:						
Victory	0.049*	-9.790	0.528	-9.247	-0.031	-0.018
	(0.027)	(8.506)	(1.186)	(429.756)	(0.021)	(0.021)
Distance to Victory	-0.003	-1.469**	0.110	-3.533	-0.001	-0.002
	(0.002)	(0.568)	(0.084)	(28.459)	(0.001)	(0.002)
Experience	-0.000**	0.374***	-0.003	-0.480	-0.000	-0.000
	(0.000)	(0.127)	(0.010)	(3.758)	(0.000)	(0.000)
Prior Success	0.002*	0.075	0.017	-3.260	0.003***	0.003***
	(0.001)	(0.502)	(0.034)	(14.905)	(0.001)	(0.001)
Competition	-0.000	0.173***	-0.014*	-3.888	0.000	0.000**
	(0.000)	(0.062)	(0.008)	(2.836)	(0.000)	(0.000)
Fixed Effects	Athlete	Athlete	Athlete	Athlete	Athlete	Athlete
Observations	848	916	724	724	221	221
R-squared	0.196	0.161	0.009	0.009	0.195	0.215

Note: The table shows the results of twelve separate estimations using six different dependent variables as indicated in the top row. We use the inverse of both initial and final odds in columns (5) and (6). In Part (I) the treatment variable is a podium finish while in part (II) treatment is defined by victory. The sample includes all athletes within a bandwidth of 15 hundredths of a second. Numbers in brackets indicate standard errors clustered at the athlete level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 5.3: Media Attention without Lists

	Day After	Week After	Day After	Week After
	(1)	(2)	(3)	(4)
Mean of dep. var.:	4.84	11.02	8.02	17.58
Podium	1.585***	2.636**		
	(0.550)	(1.324)		
Distance to Podium	-0.088***	-0.128*		
	(0.032)	(0.076)		
Victory			1.848**	4.022**
			(0.830)	(2.039)
Distance to Victory			-0.065	-0.209
			(0.065)	(0.158)
Experience	0.018**	0.039*	0.030**	0.056*
	(0.007)	(0.023)	(0.012)	(0.033)
Prior Success	0.058**	0.176*	0.051	0.154*
	(0.027)	(0.090)	(0.045)	(0.091)
Competition	-0.000	0.006	0.009*	0.026**
	(0.003)	(0.006)	(0.005)	(0.012)
Observations	2,209	2,209	907	907
R-squared	0.138	0.133	0.202	0.178

Note: The table shows the results of four separate linear regressions using media attention on the day or week after a race as dependent variable. We exclude those articles from media attention that (likely) reflect mere ranking lists. In columns (1) and (2), the treatment variable is a close podium finish while in the last two columns, treatment is defined as winning a close race. The sample includes all athletes within a bandwidth of 15 hundredths of a second. Control variables include experience, gender, prior success, and competition. Numbers in brackets indicate standard errors clustered at the athlete level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 5.4: Effect on Betting Market by Liquidity

	Dependent variable: avg. odds			
Liquidity Measure	Volume		Number	
	Low	High	Low	High
Mean of dep. var.:	0.12	0.12	0.12	0.12
I. Close Podium:				
Podium	0.006	-0.044	-0.010	-0.026
	(0.017)	(0.035)	(0.016)	(0.031)
Distance to Podium	-0.001	-0.003	-0.002	-0.002
	(0.001)	(0.003)	(0.001)	(0.002)
Prior experience	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Prior success	0.002***	0.002***	0.001***	0.002***
	(0.000)	(0.001)	(0.000)	(0.001)
Competition	0.000**	0.000	0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	221	224	216	229
R-squared	0.305	0.313	0.300	0.318

	Dependent variable: avg. odds			
Liquidity Measure	Volume		Number	
	Low	High	Low	High
Mean of dep. var.:	0.18	0.18	0.20	0.16
II. Close Victory:				
Victory	-0.016	-0.037	-0.037	-0.018
	(0.033)	(0.028)	(0.028)	(0.029)
Distance to Victory	-0.003	-0.002	-0.003	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Prior experience	-0.000*	-0.000**	-0.000*	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Prior success	0.003***	0.004***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Competition	0.000	0.000*	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	115	106	107	114
R-squared	0.445	0.506	0.557	0.376

Note: The table shows the results of eight separate linear regressions using average odds as dependent variable split by two measures of market liquidity, namely the total volume and the number of bets. The first column uses only data on races with a total betting volume lower than the median; the second column uses races with above-median volumes. The third column uses data on races with a below-median number of bets; the fourth column uses races with an above-median number bets. In Part (I) the treatment variable is a podium finish while in part (II) treatment is defined by victory. The sample is restricted to all athletes within a bandwidth of 15 hundredths of a second. Numbers in brackets indicate standard errors clustered at the athlete level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

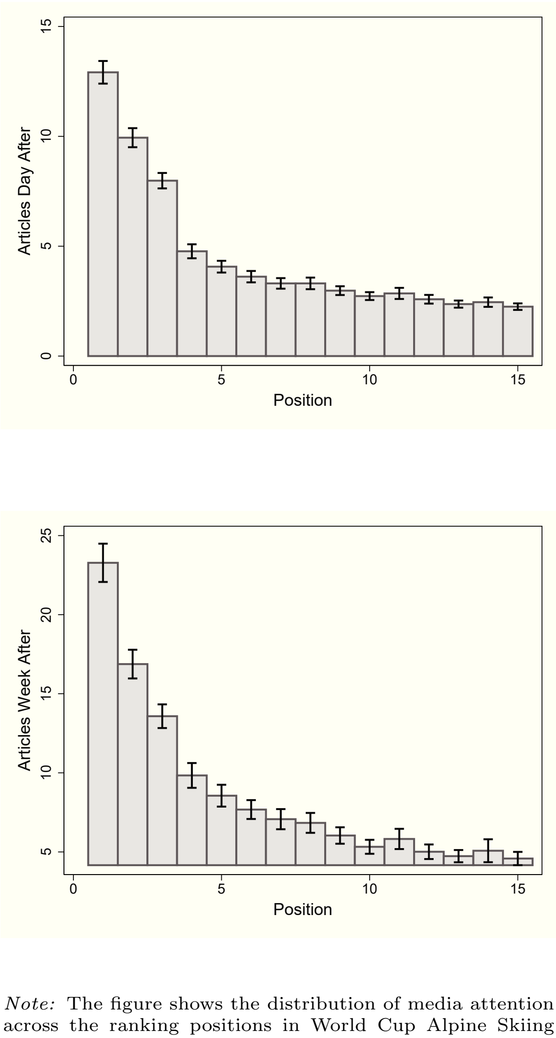
Table 5.5: Endogenous Participation Decision

Dep. var.	Participation Next Race		Competition Next Race		High Prize Next Race	
	(1)	(2)	(3)	(4)	(5)	(6)
Mean of dep. var.:	0.66	0.66	33.05	33.05	0.42	0.42
I. Close Podium:						
Podium	0.042	0.033	-1.062	-0.097	-0.039	-0.062
	(0.044)	(0.047)	(1.668)	(1.766)	(0.148)	(0.192)
Distance to Podium	-0.000	-0.003	-0.132	-0.067	0.007	0.008
	(0.003)	(0.004)	(0.108)	(0.122)	(0.009)	(0.014)
Prior experience	0.000	0.000	0.017	0.024	-0.001	0.001
	(0.000)	(0.000)	(0.011)	(0.015)	(0.001)	(0.001)
Prior success	0.003***	-0.000	0.091	0.073	0.000	-0.008*
	(0.001)	(0.001)	(0.058)	(0.072)	(0.003)	(0.004)
Competition	-0.000**	0.000	0.299***	0.263***	0.001	0.002*
	(0.000)	(0.000)	(0.010)	(0.012)	(0.001)	(0.001)
Fixed Effects		Athlete		Athlete		Athlete
Observations	1,966	1,966	1,508	1,508	247	247
R-squared	0.024	0.006	0.671	0.538	0.070	0.061

Dep. var.	Participation Next Race		Competition Next Race		High Prize Next Race	
	(1)	(2)	(3)	(4)	(5)	(6)
Mean of dep. var.:	0.74	0.74	31.98	31.98	0.42	0.42
II. Close Victory:						
Victory	-0.041	-0.006	0.587	1.008	-0.027	0.091
	(0.059)	(0.063)	(2.166)	(2.048)	(0.211)	(0.329)
Distance to Podium	-0.005	-0.009*	0.058	0.279*	-0.017	-0.003
	(0.004)	(0.005)	(0.152)	(0.155)	(0.016)	(0.017)
Prior experience	0.000	0.001*	0.013	-0.001	-0.001	0.005
	(0.000)	(0.000)	(0.014)	(0.018)	(0.001)	(0.004)
Prior success	0.001	-0.005**	0.053	0.119	-0.002	-0.023
	(0.002)	(0.002)	(0.049)	(0.077)	(0.002)	(0.014)
Competition	0.000	0.001**	0.311***	0.288***	0.001	0.002
	(0.000)	(0.000)	(0.015)	(0.022)	(0.001)	(0.002)
Fixed Effects		Athlete		Athlete		Athlete
Observations	811	811	688	688	93	93
R-squared	0.024	0.028	0.707	0.610	0.181	0.098

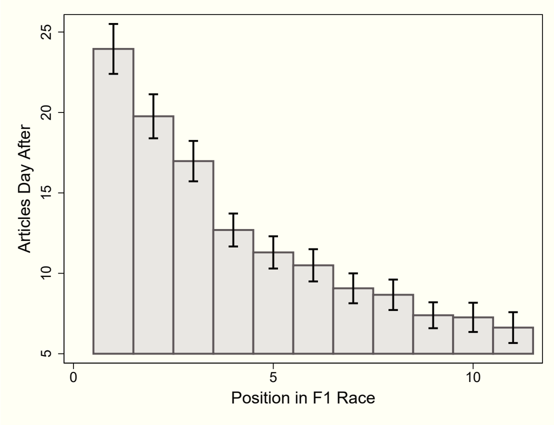
Note: The table shows the results of twelve separate linear regressions. Columns (1) and (2) use a binary indicator that captures whether an athlete participated in the subsequent race as dependent variable. Columns (3) and (4) use a measure of competition in the next race as dependent variable, namely the total number of victories among top 5 athletes. Columns (5) and (6) use a dummy variable that captures whether an athlete's next race prize money was above the median prize money. In Part (I) the treatment variable is a podium finish while in part (II) treatment is defined by victory. The sample is restricted to all athletes within a bandwidth of 15 hundredths of a second. Note that the number of observations for columns (3) and (4) are lower because competition in the next race is only defined for athletes who participated in this race. In addition, the number of observation for the prize money estimations in columns (5) and (6) is lower as we have only data on prize money for about 17% of all completed races. Numbers in brackets indicate standard errors clustered at the athlete level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

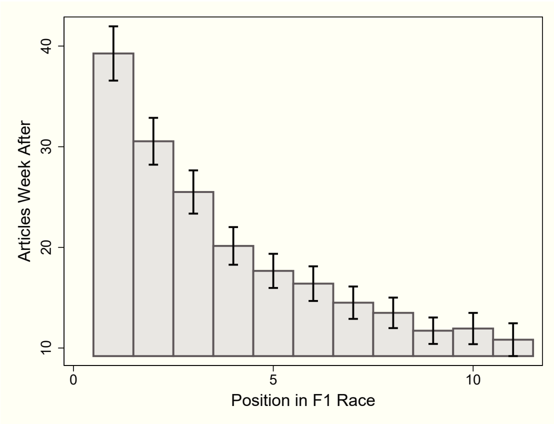
Figure 5.1: Distribution of NewsLibrary Media Attention



Note: The figure shows the distribution of media attention across the ranking positions in World Cup Alpine Skiing between 1992–2014. The data is based on NewsLibrary. On the left-hand side, we show media attention on the day after a race while on the right-hand side, we show the number of articles published in the week after the race. Whiskers indicate 95% confidence intervals.

Figure 5.2: Distribution of Media Attention in Formula One

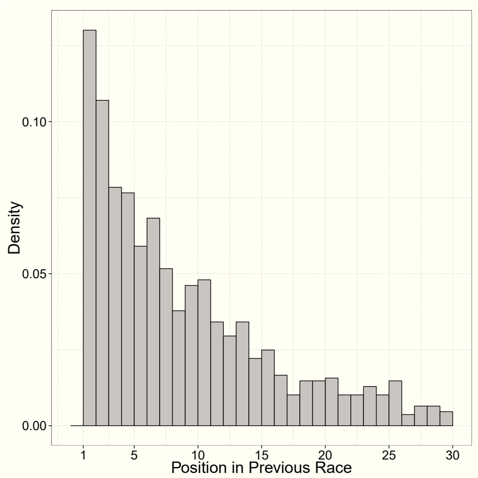




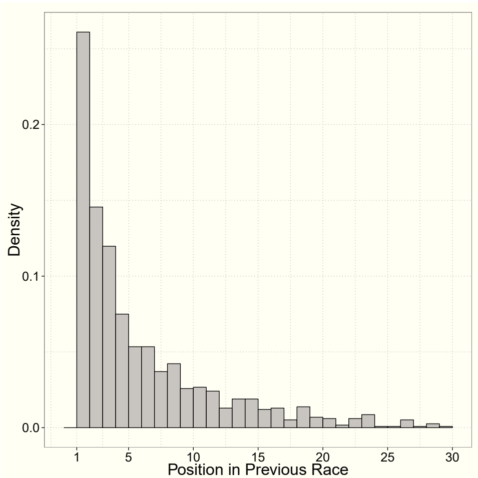
Note: The figure shows the distribution of media attention across the ranking positions in Formula 1 races between 1992–2014. The data on media attention is based on Swissdox. On the left-hand side, we show media attention on the day after a race while on the right-hand side, we show the number of articles published in the week after the race. Whiskers indicate 95% confidence intervals.

Figure 5.3: Previous Positions of Winners and 3rd-Ranked Athletes

(a) Podium



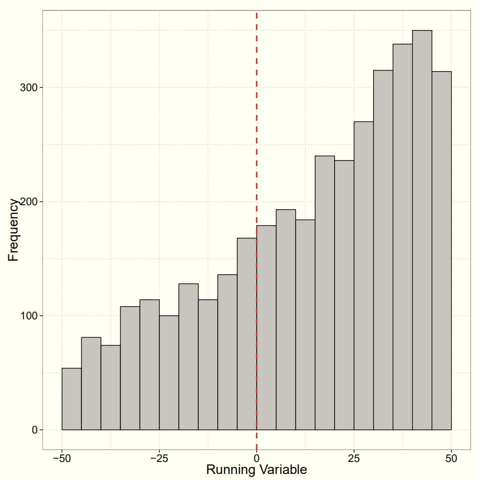
(b) Victory



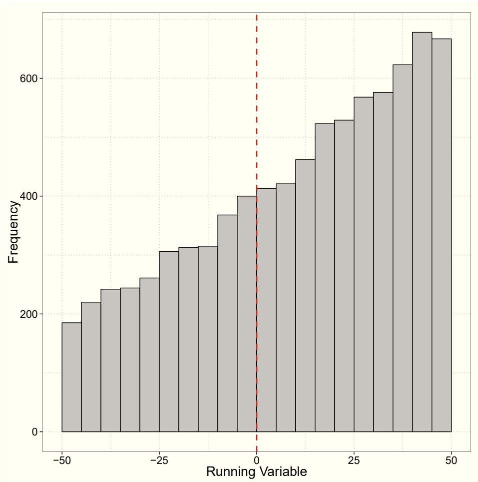
Note: The figures show histograms of the ranking positions in the previous race ($t - 1$) of third-ranked athletes and winners in the current race (t).

Figure 5.4: Observations around the Cutoff

(a) Podium



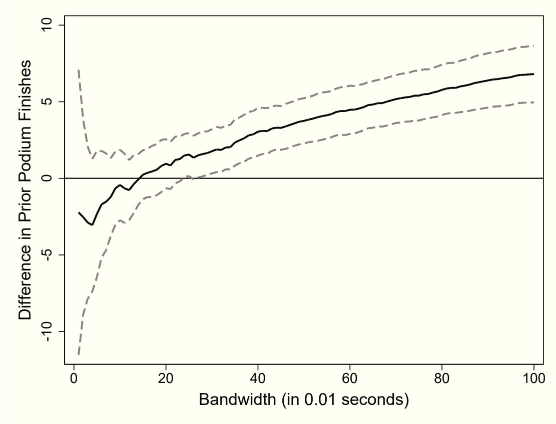
(b) Victory



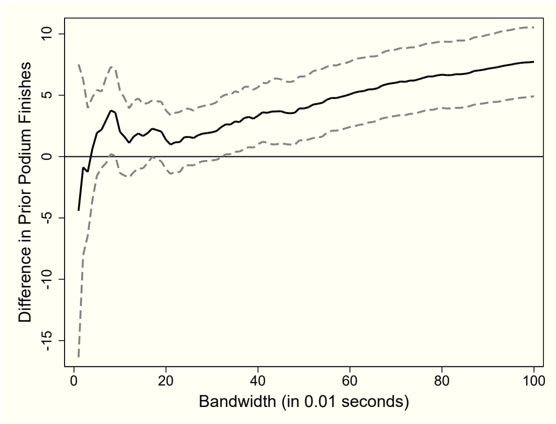
Note: The histograms show the densities of race times around the cutoff for a podium and victory finish. Both running variables are expressed in units of hundredths of a second. In line with McCrary [2008], this indicates that there is no manipulation around the respective cutoffs. In panel (a) we plot the distance to rank 2 (from left-hand side) and 1 (from right-hand side), respectively. For the distance to the podium in panel (b), we plot the distance to rank 4 and 3, respectively.

Figure 5.5: Balance Tests by Bandwidth

(a) Prior Success by Podium



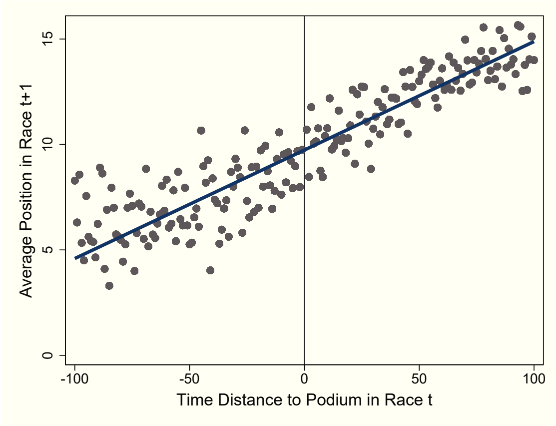
(b) Prior Success by Victory



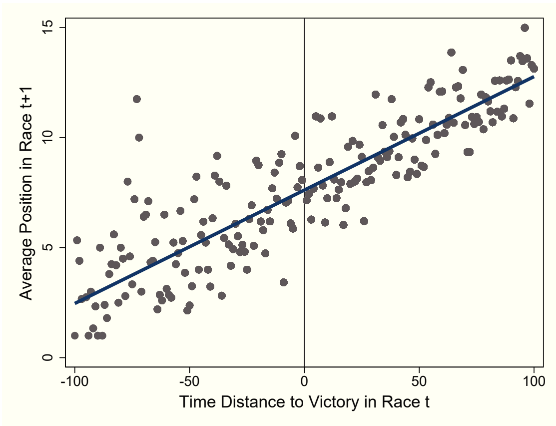
Note: The figures show the results of t-tests on athletes' number of podium finishes prior to the race determining treatment. In plot (a), treatment is defined by podium while in plot (b) it is based on whether an athlete finished as the winner. The sample includes all tournaments between 1992–2014. Confidence intervals at 95% are shown.

Figure 5.6: Serial Correlation in Individual Performance

(a) Distance to Victory

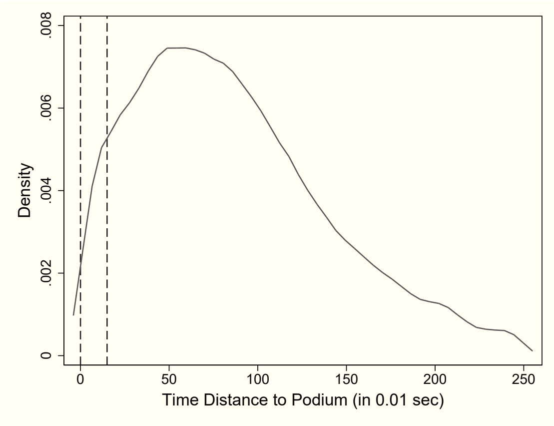


(b) Distance to Podium



Note: The figure shows the average position in race $t + 1$ for each distance to the winner (left-hand side) or podium (right-hand side) in race t. In addition, we add a linear regression line. The sample includes all athletes between 1992–2014 who finished within a bandwidth of one second to the victory (a) or podium (b).

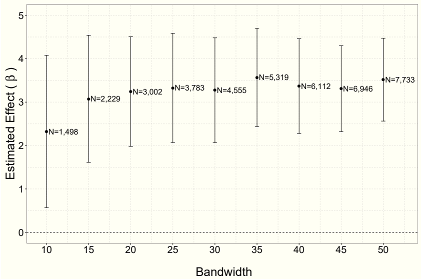
Figure 5.7: Distribution of Running Variable and Bandwidth Choice



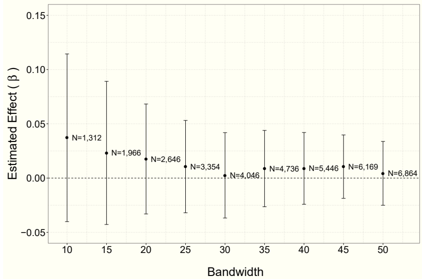
Note: The figure shows the distribution of the time distance to the podium (i.e., running variable). We plot an Epanechnikov kernel function with a bandwidth of 0.05 seconds. The dashed vertical lines illustrate our preferred bandwidth choice of 0.15 seconds.

Figure 5.8: Regression Results using Different Bandwidths

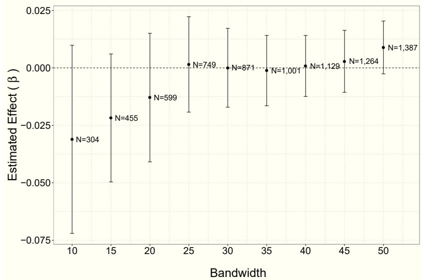
(a) Effect on Media Attention



(b) Effect on Performance



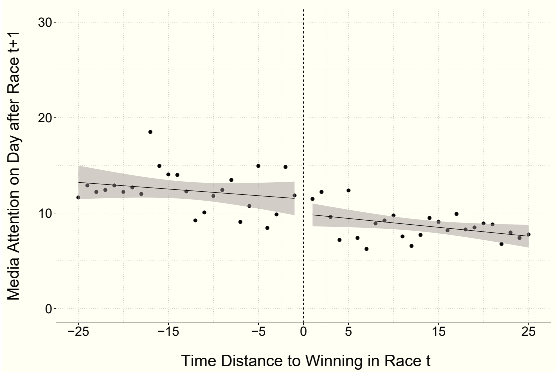
(c) Effect on Inverse Betting Odds



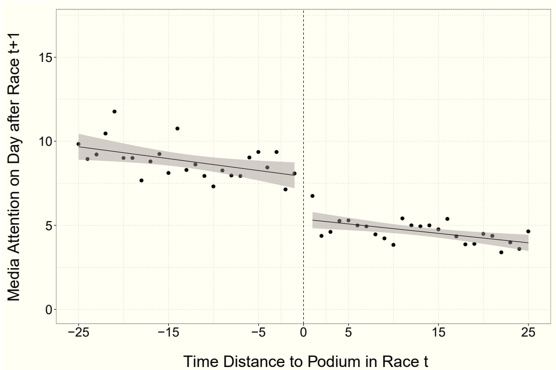
Note: The figure shows the results from estimating the effect of a podium finish on media attention on the day after the competition (Panel (a)), the probability to win the next race (Panel (b)), and average inverse betting odds in the next race (Panel (c)) using bandwidths from 10 to 50 hundredths of a second. The dots indicate the point estimate, while the bars depict the 95% confidence interval.

Figure 5.9: Effect of Top Ranks on NewsLibrary Media Attention

(a) Close Victory



(b) Close Podium Finish



Note: The figures show a linear fit for the number of media articles (using NewsLibrary data) that mention an athlete (on the day after a race took place) for both treatment (left) and control group (right). In panel (a) we compare victory and non-victory while in panel (b) athletes on the podium are compared with those missing it by 0.25 seconds or less. The grey region indicates the 95% confidence interval.

Appendix B: Attrition Bias

One problem that may arise when estimating the effect of a ranking position on future performance is that athletes differ with respect to their probability of survival, i.e. not crashing. If this probability is related to success our estimates for performance would be biased. It could be, for example, that athletes who are very successful once adopt a riskier behavior in subsequent races in order to be successful again.

Following Frangakis and Rubin [2002], we denote athletes with a constant low (high) probability of survival by DD (LL). While the survival probability of this set of athletes is unaffected by the treatment, other athletes adjust their behavior when being treated, in other words after a quasi-random top rank. The athletes in subset LD adopt a more risky strategy after treatment while those in subset DL follow a low-risk strategy in case they achieve a quasi-random top rank. In the regression of the probability of survival ($s_{i,j+1}$) on treatment $D_{i,j}$ and controls for a athlete's own characteristics $\mathbf{X}_{i,j}$ and competitors' characteristics $\mathbf{Z}_{i,j}$

$$s_{i,j+1} = D_{i,j}\tau + \mathbf{X}_{i,j}\,\gamma + \mathbf{Z}_{i,j}\,\delta + \varepsilon_{i,j} \tag{5.5}$$

we should expect $\tau = 0$ for the two types with constant behavior (DD and LL). For types DL we expect $\tau < 0$ and for types LD we should see an increase in the probability of survival. Thus we have two problems if the coefficient τ is significantly negative: First, our estimates with respect to subsequent performance would be biased upwards because we would only observe treated athletes in case they are successful in subsequent races. Second, the overall gain from imposing a ranking will be reduced and perhaps negative if top ranks lead to a strong increase in risky behavior.[1]

[1] This finding would be in line with research on addiction to success.

Appendix C: Relative Performance Measures and SUTVA

The Stable Unit Treatment Value Assumption (SUTVA) is fundamental to most estimators used in the program evaluation literature. It allows to write the treatment status of individual i only dependent on her assignment, and the outcome of individual i only dependent on her assignment and treatment status. More formally, SUTVA is defined as follows (according to Angrist, Imbens and Rubin [1996, p.446]):

(a) If $Z_i = Z_i'$, then $D(\mathbf{Z}) = D(\mathbf{Z}')$

(b) If $Z_i = Z_i'$, then $Y_i(\mathbf{D}, \mathbf{Z}) = Y_i(\mathbf{D}', \mathbf{Z}')$

This allows us to write $D_i(\mathbf{Z}) = D_i(Z_i)$ and $Y_i(\mathbf{D}, \mathbf{Z}) = Y_i(D_i, Z_i)$.

Applying this assumption to our paper, let us define the vector of final times $\mathbf{T}$ in race j as well as vectors of assignments (to the podium) ($\mathbf{Z}$) and treatments ($\mathbf{D}$):

$$\mathbf{T} = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ \vdots \\ t_N \end{pmatrix} \qquad \mathbf{Z} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_N \end{pmatrix} \qquad \mathbf{D} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_N \end{pmatrix} \tag{5.6}$$

We assume that there is a positive probability of a crash. We model survival as

$$S_i = \begin{cases} 1 \text{ if } S_i^* > 0 \\ 0 \text{ if } S_i^* \leq 0 \end{cases} \tag{5.7}$$

with

$$S_i^* = \theta_i a_1 + \mu_j + \varepsilon_i \tag{5.8}$$

where S_i^* is a latent variable, θ_i is an athlete's skill, $a_1 > 0$ is a coefficient, μ_j is a race fixed effect, ε_i is an unobserved component. With $a_2 < 0$ being some coefficient, the final time can be written as

$$T_i = \begin{cases} \text{NA} & \text{if } S_i = 0 \\ \theta_i a_2 + \delta_j + v_i & \text{if } S_i = 1 \end{cases} \tag{5.9}$$

We consider a specific race with three top athletes under two circumstances. First, conditions are equal for all athletes. Second, the three top athletes ($i \in \{1, 2, 3\}$) suffer from bad weather conditions, which makes it impossible for them to attain a place on the podium.

$$\mathbf{Z} = \mathbf{D} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \qquad \mathbf{Z}' = \mathbf{D}' = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{5.10}$$

In this case, it is obvious that the assignment status of individuals 4–6 depends upon the assignment of the three top athletes. However, since $D_i = Z_i$, the individual treatment status D_i can still be written as a function of the assignment Z_i.

Turning to the implication (b) of SUTVA, we first note that our outcome can be written purely as a function of the assignment, i.e. $Y_i(\mathbf{Z}, \mathbf{D}) = Y_i(\mathbf{Z})$. This comes from the fuzzy design where $\mathbf{Z} = \mathbf{D}$.

Any measure of relative performance —such as the position— depends on a athlete's own time as well as the competitors' times: $P_{i,j} = g(T_{i,j}, T_{s,j}) =$

$g(\theta_i, \lambda_{i,j}, n_{i,j}, \theta_s, \lambda_{s,j}, n_{s,j}) \quad \forall s \neq i$. The winner's time is often the benchmark and can be written as

$$T_{\text{win},j} = \min_{i \in S}(T_{i,j}) \tag{5.11}$$

where S indicates the set of survivors. For the sake of illustration, we specify equation (5.9) for survivors as

$$T_{i,j} = \theta_i a_2 + \delta_j + \text{Exp}_i a_3 + \text{Exp}_i^2 a_4 + \tau_i a_5 + u_i \tag{5.12}$$

where Exp_i is experience and treatment τ_i equals one if athlete i won the last race and zero otherwise. Imagine that the winner in a given race j is determined by a tiny time difference between athlete 1 and 4. Assume both athletes to have the same skill level θ_i, but while athlete 1 is a rookie, athlete 4 is an experienced and successful athlete. Athlete i's relevant outcome $Y_i = P_{i,j+1}$ in the next race depends on the performance of the best athlete in that race. So if athlete 4 wins today and $a_5 \neq 0$, $T_{\text{win},j+1}$ is likely to be lower than if athlete 2 wins (because athlete 4 is more experienced and experience positively affects performance).

Therefore, the outcome of athlete i in race $j + 1$ is likely to depend on the assignment of the winner (note that treatment and assignment are henceforth defined for the victory treatment and not the podium treatment as above)

$$\mathbf{Z} = \mathbf{D} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad \mathbf{Z}' = \mathbf{D}' = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \qquad \text{and} \qquad Y_i(\mathbf{Z}) \neq Y_i(\mathbf{Z}')$$

which violates definition (b) of SUTVA. This problem arises to different extents with all kinds of relative performance measures.

Appendix D: Sample Race

Table 5.1: Sample Race for Illustrative Purpose

Rank	Name	Nationality	Time	Difference	Articles
1	Benjamin Raich	AUT	1:36.66		13
2	Akira Sasaki	JAP	1:36.83	0.17	9
3	Thomas Grandi	FRA	1:37.17	0.51	15
4	Michael Janyk	USA	1:37.19	0.53	4
5	Ted Ligety	USA	1:37.54	0.88	5

Table 5.1 shows the result of the 2006/07 Men World Cup Slalom race in Shigakogen, Japan. We observe that Thomas Grandi achieved a podium finish because he was 0.02 seconds ahead of Michael Janyk. This race result can be used to illustrate how we define treatment and control group in our estimation. First, we take the time of the third-ranked athlete (1:37.17 in this case). Then we compute a 0.15 seconds window around this race time. All athletes within this time window (1:37.02 to 1:37.32) are part of our estimation sample. Every athlete in this group who finished on the podium is in the treatment group. Every other athlete in the estimation sample serves as part of the control group.

Appendix E: Two Examples of Media Attention

The following two clips from newspapers illustrate the kind of articles we find using Swissdox or NewsLibrary to measure media attention.

Figure 5.1: Examples of Media Attention

Pittsburgh Post-Gazette (PA)

January 27, 2013
Edition: TWO STAR
Section: SPORTS
Page: D-18

Column: MORNING BRIEFING

VONN TOPS RIVAL IN GS
Author: *From wire dispatches*

Article Text:

American **Lindsey Vonn** beat rival Tina Maze of Slovenia on Maze's home snow and in her best discipline Saturday, earning a surprising victory in a giant slalom in Maribor.

Maze had a chance to secure her third giant slalom discipline title with a victory and led after a near-perfect first run, but a poor start to the second cost her valuable time and she finished 0.08 seconds behind **Vonn**.

Vonn proved again that she is back to her best after an illness by putting down two good runs in what is traditionally her weakest event to win in 2:22.2.

Vonn was third after the first run but overcame several errors to have the fastest time in the second and create a margin that Maze couldn't bridge.

"In the second run I decided, OK, it's all or nothing, I had to go for it," **Vonn** said. "It's been a rough year for me in GS, so it's just perfect."

Vonn earned her first GS victory since March 2012, and her second win in a week after taking the downhill at Cortina D'Ampezzo.

With her 59th World Cup win overall, she's just three away from equaling Annemarie Moser Proell's record on the women's side.
More skiing
Dominik Paris became the second Italian skier to win the classic men's World Cup downhill in Kitzbuehel, Austria, Saturday. Paris came down the approximately 2-mile Streif course in 1:57.56 to follow in the footsteps of Kristian Ghedina, who won in 1998. World downhill champion Erik Guay of Canada trailed Paris by 0.13 in second place, and Hannes Reichelt of Austria came 0.36 back in third.
Horse racing
Hall of Fame jockey Russell Baze rode the 50,000th thoroughbred race of his career. Baze, 54, finished third in his historic race Friday at Golden Gate Fields near San Francisco, although he was allowed in the winner's circle afterward to pose for a photo. He swept the first three races on the eight-race card. Baze has notched 11,839 victories from his 50,000 mounts, and both totals are North American racing records, according to Equibase.

© 24 heures; 28. Januar 2013

Ausgaben-Nr. Seite 15VC1
sports

swissdox.ch

Tina Maze continue d'affoler les statistiques

Ski alpin2edu géant samedi derrière Lindsey Vonn , la Slovène domine le slalom sur sa piste de Maribor

Tina Maze n'est pas restée longtemps dans l'ombre à Maribor. Eclipsée samedi en géant par Lindsey Vonn , la Slovène a signé un succès éclatant lors du slalom dominical. Comme la veille, Tina Maze pointait en tête après la manche initiale. Mais la comparaison s'est arrêtée là. Battue d'un souffle (0"08) samedi par une Lindsey Vonn ressuscitée en géant, la Slovène n'a pas essuyé un deuxième camouflet sur ses terres. Meilleur temps sur les deux tracés, elle s'est offert un triomphe devant un public en délire, reléguant sa dauphine suédoise Frida Hansdotter (à 0"86) au rang de faire-valoir.

«En franchissant la ligne, je n'ai pas vu mon classement. Mais quand j'ai entendu la clameur de la foule, j'ai su que c'était bon», a raconté la superstar nationale.

Avec cette nouvelle victoire, Tina Maze n'en finit plus d'affoler les statistiques. En 25 courses cet hiver, elle est montée à 17 reprises sur le podium, dont sept fois sur la plus haute marche. Sauf accident, elle va exploser le mythique record de Hermann Maier, qui avait accumulé 2000 points en 1999-2000. La Slovène en est actuellement à 1654 unités, alors qu'il reste douze épreuves.

La fille aux skis suisses (Stöckli) s'apprête aussi à faire une razzia parmi les globes. Celui du géant est mathématiquement assuré et celui du classement général devrait l'être bientôt. Hier, elle s'est aussi replacée en slalom à seulement 13 points de Mikaela Shiffrin. Le prodige américain, qui restait sur deux succès, a payé cher une grosse faute en première manche et a dû se contenter de la 8e place.

Côté suisse, Wendy Holdener a continué de jouer les métronomes. Toujours classée entre le 9e et le 17e rang cet hiver en slalom, la Schwytzoise a terminé 10e à Maribor. Une régularité remarquable pour une skieuse de 19 ans. Il ne lui manque plus qu'une place aux avant-postes pour franchir un nouveau palier dans sa prometteuse carrière. Il a ensuite fallu remonter au 21e rang pour retrouver Michelle Gisin. Un classement qui reste bon à prendre pour la «rookie» obwaldienne, qui dispute à 19 ans sa première saison en Coupe du monde. La déception suisse du week-end est venue de Lara Gut, éliminée samedi en géant. Brillante en début de saison dans la discipline avec des 4es places à Aspen et à Saint-Moritz, la Tessinoise n'y arrive plus avec un 17e rang comme meilleur classement lors de ses quatre dernières sorties. SI

Note: The figures show two examples of media attention found in NewsLibrary (left-hand side) as well as Swissdox (right-hand side) for Linsey Vonn and Tina Maze in the aftermath of World Cup races in late January of 2013.

Chapter 6

Limited Attention and Risk-Taking Behavior

This chapter is based on joint work with Reto Föllmi and Lukas Schmid from the University of St.Gallen and published as "Do Professionals Get It Right? Limited Attention and Risk-Taking Behaviour" in *The Economic Journal* (2016) Vol. 126 (592), p.724–755.

6.1 Introduction

Individuals often have to make decisions under uncertainty that involve risk-return trade-offs. Although traditional economic models assume perfect information processing and foresight, a large body of research in behavioral economics has documented the limits of individuals' cognitive abilities [DellaVigna, 2009]. The literature has focused in particular on the question of how limited attention affects consumption choices and has provided evidence for a left-digit bias, the empirical regularity of people's tendency to focus on the leftmost digit of a number and pay only partial attention to other digits [Korvorst and Damian, 2008; Lacetera, Pope and Sydnor, 2012].

Despite this empirical evidence, three challenges remain. First, estimating the causal effects of the left-digit bias is difficult due to bunching of data. Second, it remains unclear whether limited attention also influences the process of individual decision making with respect to risk taking. Finally, there is an on-going discussion about whether individual experience or high stakes situations mitigate behavioral biases.[1]

In this paper, we propose a new approach to addressing these three challenges, namely, by estimating the impact of behavioral biases on risk-taking in a setting involving professional and experienced athletes engaged in fierce competition. We investigate the presence of a left-digit bias by using detailed data on 1,865 athletes in World Cup alpine skiing over the period of 1992–2014. Our empirical analysis exploits the fact that slalom and giant slalom races consist of two separate runs. After the opening run, each athlete obtains information about her own time as well as her distance in relation to the current leader. We explore whether athletes exhibit a left-digit bias when processing this time difference to the leader. In particular, we test whether the use of heuristic thinking affects the way athletes choose their risk strategy in the second run. In the presence of a left-digit bias, our theoretical model shows that athletes misinterpret distances such as nine hundredths of a second to be significantly smaller than, for example, ten hundredths of a second. This behavioral bias in turn leads to the adoption of a more risky strategy because achieving the great success (i.e., winning the race) appears to be more likely if the gap to the current leader is small rather than large. In our empirical analysis, we apply a regression discontinuity design for the estimation of causal effects by exploiting the fact that the allocation of right digits in athletes' time distance to the leader can be regarded as quasi-random.

Our empirical findings suggest that professional athletes exhibit a substantial left-digit bias. Individuals with an opening-run time difference to the leader just below a tenths-of-a-second threshold are significantly more likely to adopt a risky behavior, which increases the probability of not successfully finishing the race by up to 28.0%. Moreover, the standard deviation of race times in the second run increases by approximately 26.1%. The estimated effect is robust when us-

[1] The presence of behavioral biases in the context of experience, competition and high stake situations has been subject to widespread scepticism [Feng and Seasholes, 2005; Levitt and List, 2008; List, 2003; Pope and Schweitzer, 2011].

ing different bandwidths and including race-fixed effects. To account for genetic determinants of risk-taking behavior, we add athlete-fixed effects to our baseline specification and obtain very similar estimates. As expected, we find the effect to be present only among athletes close enough to the leader after the first run to have a plausible chance of winning the race. In a placebo test, we construct left-digit breaks based on time differences expressed in minutes, a figure that is not shown to athletes, and find no relationship with second-run behavior. These results are consistent with our theoretical prediction that athletes receive a signal about their time distance to the leader and pay only limited attention to right digits. In contrast to previous evidence by List [2003], as well as Gardner and Steinberg [2005], we find that the behavioral bias does not disappear when restricting the sample to older, more experienced athletes. Furthermore, the left-digit bias is also present in races with particularly high stakes.

To examine the sensitivity of our empirical findings, we conduct a series of robustness checks. First, we document that there is no difference in predetermined covariates between the treatment and control group. Second, we test alternative digit breaks, finding that all other cutoffs such as 0-1, 2-3, or 6-7, exhibit no discontinuity in survival rates. In our third robustness test, we calculate time distances to the second- and third-ranked athlete. Because these differences are not shown to athletes, they can be used as placebo treatments. All estimates on placebo treatments are very close to zero, thus increasing our confidence that the main findings are in fact driven by limited attention. We also explore whether the effect of the left-digit bias is driven by nervousness and provide evidence that the bias is also present among athletes with arguably low levels of nervousness.

Our results contribute to a number of studies in psychology and economics. In particular, the observation of a persistent behavioral bias in the context of large stakes and highly experienced professionals appears puzzling [Levitt and List, 2007, 2008]. We argue that our results can be explained by different ways of thinking [Kahneman, 2011; Stanovich and West, 2000], including the concept of ego-depletion. Baumeister et al. [1998] argue that 'all variants of voluntary effort—cognitive, emotional, or physical—draw at least partly on a shared pool of mental energy'. If individuals exert a large amount of physical or mental effort on one particular task, they are less likely to pay full attention to or exert full effort on a subsequent task. The very high stakes in World Cup competitions cause athletes to exert extreme effort during the race, thus making them vulnerable to behavioral biases afterwards. The aforementioned placebo tests reveal that it is the very information provided to athletes that shapes their behavior. Neither time differences expressed in minutes or the time gap to the second or third contestant is correlated with second-run behavior. Athletes that are physically exhausted after the opening run appear to use heuristics when processing information about performance differences. Hence, the left-digit bias is present only with respect to information that is readily available.

Our study makes several contributions to the literature on limited attention and risk-taking. In the field of behavioral economics, several studies have investigated how individuals deal with signals and information. We link this research to the literature on the determinants of risk preferences. In particular, we provide evidence that heuristic information processing affects not only consumption choices but also risk-taking behavior. Our findings are closely related to the work by Lacetera, Pope and Sydnor [2012], who find that the left-digit bias is present in product markets. A notable advantage of our research design is the smoothness of the assignment variable—the distance to the leader—at the respective cutoff. It allows us to avoid the problem of clustered observations. We can also rule out any kind of manipulation, which is more likely to occur in product markets. Second, our findings suggest that even professional and experienced actors appear to suffer from limited attention. This complements previous research by Busse et al. [2013] who document that professional dealers in the wholesale market for cars anticipate that final customers in the retail market will focus on the left digit of the odometer. In contrast, we provide a psychological explanation for why behavioral biases can exist despite high stakes and individual experience. Third, our results show that limited attention affects risk-taking behavior even though all relevant information is easily observable. Much of the literature on limited attention has focused on settings in which, to some extent, information is shrouded [Brown, Hossain and Morgan, 2010]. The importance of heuristic thinking appears to be much greater if it can even be documented in settings where information is readily available. Again, we provide a psychological rationale for the persistence of behavioral biases in our setting.

Finally, our findings also contribute to a growing literature on the heterogeneity of risk-taking behavior across individuals. Understanding the determinants of risk preferences is particularly important because the assumptions about individual risk behavior are key to economic models.[1] A large body of literature has investigated the various determinants of risk preferences and found both genetics [Barnea, Cronqvist and Siegel, 2010; Cesarini et al., 2009] and personal experiences [Booth and Nolen, 2012; Dohmen et al., 2012; Malmendier and Nagel, 2011] to be relevant. We find that irrespective of an individual's genetics and experience, the way of processing information also shapes behavior under uncertainty.

The paper proceeds as follows. In Section 6.2, we illustrate how limited attention can generate discontinuities in risk behavior among professional athletes. Section 6.3 presents some general information on World Cup alpine skiing and descriptive statistics on our data set. Section 6.4 provides a description of our econometric approach. In Section 6.5, we show the main empirical findings, discuss effect heterogeneity, provide a psychological explanation and numerous robustness checks. Finally, Section 6.6 concludes.

6.2 Theoretical Considerations

6.2.1 Left-Digit Bias

Following the seminal paper by Simon [1955], a large body of literature has examined imperfect individual information processing. Tversky and Kahneman [1974] point out that people tend to rely on specific heuristic principles that reduce the complexity of difficult tasks. These heuristics may be useful in many occasions but they can also lead to severe biases that have been documented in various fields of economics. Conlin, O'Donoghue and Vogelsang [2007] find that individ-

[1] Recent press coverage emphasises the increased attention paid to the heterogeneity of risk preferences across individuals. See, for instance, the article 'Risk off', in the *The Economist*, January 25, 2014.

uals' decisions are excessively influenced by current weather conditions. Gabaix and Laibson [2006] show how shrouding may occur in an economy if at least some customers are myopic. An extension of their framework provides an explanation for thinking in categories, shedding light on the causes of uninformative advertising [Mullainathan, Schwartzstein and Shleifer, 2008]. Chetty, Looney and Kroft [2009] show that people have a lower demand for products if those are tagged including commodity taxes than if taxes are not included in posted prices. These findings suggest that the salience of taxes plays an important role for individual tax responses.[1] Further research by Ashton [2013], however, indicates that a left-digit bias is the main channel through which tax salience affects consumer decisions.

There is also empirical evidence on the question for which types of goods and transactions people tend to use heuristics. Köszegi and Szeidl [2013] document that individuals tend to focus more on attributes in which disparities are large. In professional skiing, as in many other fields of sport, differences in performance can be very small. Yet focusing on left digits can subjectively generate a sharp distinction between time differences that are in fact very similar. Thus, left-digits may serve as a reference point as described in Köszegi and Rabin [2006] as well as Gill and Prowse [2012].

The recent literature has paid particular attention to a heuristic technique known as the left-digit bias. Basu [1997] examines the prevalence of 99-cent pricing and argues that it can be explained in a model of full rationality. Schindler and Kirby [1997] examine pricing strategies using advertisements data from newspapers and present evidence that the over-representation of 9 endings is due to reference points that are linked to the decimal system. Anderson and Simester [2003] conduct three experiments that randomly vary the ending digits of prices and find that a price of $9 yields significantly higher demand than slightly lower or higher prices, particularly when products are unfamiliar and customers pay limited information. Lacetera, Pope and Sydnor [2012] as well as Busse et al. [2013] advance this work by showing that information-processing heuristics mat-

[1]This evidence is consistent with the results of Finkelstein [2009], suggesting that the introduction of an electronic collection system for highways increases total tolls because driving becomes less elastic with respect to the toll.

ter even in markets with large stakes and easily observed information. In addition to this, Backus, Blake and Tadelis [2015] provide evidence that round numbers can be used as signals in economic negotiations.

6.2.2 Model

We discuss the implications of a left-digit bias in World Cup alpine skiing by means of a simple theoretical model. Following the seminal work by Atkinson [1957], risk-taking behavior is affected by both the motive to achieve and the motive to avoid failure. This is particularly relevant in the context of World Cup alpine skiing in which athletes have to choose very carefully the level of risk they are willing to take. Even small mistakes can lead to errors that cause a substantial loss of time, reducing the probability of being successful.

For simplicity, we assume that the utility of an athlete is comprised of three elements only, winning, not-winning but finishing, and not finishing the race at all. The athlete's risk choice impacts both the variance of time if finishing the race, and the probability to finish the race. We normalise the direct utility of a victory to U_W, the direct utility of finishing behind to $U_L < U_W$, the utility of not finishing the race at all to zero.

We consider the decision problem of an athlete i who trails the leader after the first run. Athlete i is endowed with talent ξ_i, incurred a distance d_i in the first run and chooses which level of risk r_i to take in the second run. Distance is negatively related to talent but randomly—due to weather and wind conditions—the distance may be higher or lower than predicted by talent alone, thus $d_i = \delta(\xi_i, \varepsilon_i)$ with $\delta_\xi < 0$ and $\delta_\varepsilon > 0$, while $E(\varepsilon_i) = 0$.

Athlete i takes the risk decision of the leader as given. More talent increases the chance of finishing the race $\phi(\cdot)$. Taking more risk decreases the probability of finishing ($\phi' < 0$) but increases the variance of the race time in the second run. We omit the index i in what follows as long as this causes no confusion. Given the race is finished, the chance to win the race is given by $\pi(d + rz)$ where z is a normally distributed random variable. We are free to choose the unit of measurement concerning risk, hence we assume z has unit variance. We assume that a larger distance after the first run reduces the probability to win. If the outcome of risk-taking is success (i.e., a low time in the second run) the realization of z is negative and the winning probability higher. Furthermore, we assume that the marginal winning probability shall be concave. Thus, the derivatives alternate in signs, hence $\pi' < 0$, $\pi'' > 0$, and $\pi''' < 0$ for positive arguments. This is also true in our data, as shown in Figure 6.1 in the Appendix.

For our theoretical exposition, we make a slightly stronger assumption and we impose that $\pi(\cdot)$ is convex enough such that $\pi'/\pi > \pi'''/\pi''$. This condition holds, for example, for the Pareto distribution.

— Figure 6.1 about here —

In Figure 6.1, the probability to win given the distance d in the first run is calculated and the path of $\pi(\cdot)$ is estimated non-parametrically.[1] We see that the winning probability follows a decreasing but convex path as a function of distance d.[2] Omitting the talent variable ξ in what follows, these arguments let us write the utility function in the following way:

$$U(r) = E\left[\phi(r) \times \left(\pi(d + rz)U_W + (1 - \pi(d + rz))\, U_L\right)\right]. \tag{6.1}$$

We use this cost-benefit framework to investigate the implications of a particular behavioral bias: the left-digit bias. Drawing on work by Lacetera, Pope and Sydnor [2012] we can incorporate limited attention to performance differences. Athletes in the model are assumed to pay full attention to the left digit (i.e., the more visible component) of the time distance to the leader but only partial attention to the right digit. Let T be the time of the athlete in the first run, expressed in hundredths of a second. Then we can define her distance to the leader as $d \equiv (T - T_1)$. This distance can be broken down into two parts. The first part, d_l, indicates the number of tenths of a second in the distance (i.e., the left digit). The second part, $mod(T - T_1, 10)$, is the modulo of the distance to the leader with respect to ten (i.e., the right digit). The modulo finds the remainder

[1] Note that the estimated winning probability in Figure 6.1 measures the winning probability as a function of distance d. However, risk changes along the distance, hence the estimated winning probability is not exactly identical to the theoretical $\pi(d - r)$ curve.

[2] Similar patterns of success by time distance to the leader can be shown for a finish in the top 3 and top 5, as shown by Figure 6.2 in the Appendix.

of the division of the time distance by ten, that is the number of single hundredths of a second. For example, a distance of 39 hundredths of a second yields 3 tenths of a second and a modulo of 9. We can express the *perceived* distance $\hat{d}$ as

$$\hat{d} = d_l 10 + (1 - \theta) \times mod(T - T_1, 10), \tag{6.2}$$

where $\theta \in [0, 1]$ is the inattention parameter. Note that $mod(T - T_1, 10) \in \{0, 1, 2, ..., 9\}$ while $d_l \in \mathbb{N}$. For example, a time distance of 119 hundredths of a second would be perceived as $\hat{d} = 11 \times 10 + (1 - \theta) \times 9 = 110 + (1 - \theta) \times 9$.

This approach generates discontinuities at distinct cutoffs. In particular, at each tenths-of-a-second threshold, the perceived distance jumps. Consider, for example, a distance of 20 hundredths of a second. As long as d is below that threshold, the athlete will perceive a change of $(1 - \theta)$ for every one-unit increase in d. However, when crossing the cutoff from 19 to 20, the perceived increase will be $1 + \theta \times 9$. With a maximum bias ($\theta = 1$) the perceived increase is 10, while in the absence of any bias ($\theta = 0$) it is 1.

— Figure 6.2 about here —

The biased perception of distances is illustrated by the step function in the lower half of Figure 6.2. The absolute discontinuity is the same at each threshold and given by $\Delta = 10 \times \theta$. In the upper half we see the implications of this discontinuity for the perceived chance of winning as a function of distance. While the *actual* change of the winning probability $-\pi' \ (d + rz)$ is a smooth downward sloping function, the *perceived* change of the winning probability $-\pi'(\hat{d} + rz)$ is again a step function. The important observation is that to the left of each tenths-of-a-second threshold it holds that $d > \hat{d}$. And because the winning probability is a negative function of the distance to the leader, it also holds that $-\pi'(d + rz) < -\pi'(\hat{d} + rz)$ to the left of each threshold.

Let us study the optimal risk decision of the athlete. While taking the actions of the leader as given, the perceived first order condition of the athlete reads

$$\begin{aligned} \hat{U}'(r) =& \phi'(r)E\left[\pi(\hat{d}+rz)U_W + \left(1-\pi(\hat{d}+rz)\right)U_L\right] \\ &+ \phi(r)E\left[\pi'(\hat{d}+rz)z\left(U_W-U_L\right)\right] =: 0 \end{aligned} \tag{6.3}$$

which is equal to zero in the optimum.
The negative first term $\phi'(r)E\left[\pi(\hat{d}+rz)U_W + \left(1-\pi(\hat{d}+rz)\right)U_L\right]$ denotes the marginal cost of taking additional risk, it equals the product of increased probability of non-finishing times the expected utility of finishing. The positive second term $\phi(r)E\left[\pi'(\hat{d}+rz)z\left(U_W-U_L\right)\right]$ captures the benefit of risk taking which is the product of the increased probability to win times the utility gain of $U_W - U$. Note that the presence of a left-digit bias ($\theta > 0$) causes $\hat{d}$ to be smaller than d. The following proposition states that a reduced value of $\hat{d}$, through the left-digit bias, increases risk taking and lowers the probability to finish the race.

Proposition 1. *The presence of a left-digit bias leads athletes to overestimate the winning probability and choose a higher risk level. This decision (i) increases the probability of not finishing the second run, (ii) raises the variance of time in the second run, (iii) leaves the average race time, given finishing the race, unaffected, and (iv) increases the probability to win, given finishing the race.*

Proof. We do a second order Taylor approximation of equation (6.3). Note that $E[z] = E[z^3] = 0$ and $E[z^2] = 1$. We get

$$\hat{U}'(r) \tilde{=} \phi'(r)\left[\left(\pi(\hat{d}) + \pi''(\hat{d})r^2/2\right)\left(U_W-U_L\right) + U_L\right] + \phi(r)\pi''(\hat{d})r\left(U_W-U_L\right) = 0. \tag{6.4}$$

We are interested how the derivative changes when the perceived distance changes. We take the derivative to get $\partial\hat{U}'(r)/\partial\hat{d} = \phi'(r)\left(\pi'(\hat{d}) + \pi'''(\hat{d})r^2/2\right)\left(U_W-U_L\right) + \phi(r)\pi'''(\hat{d})r\left(U_W-U_L\right)$. Using equation (6.4) we may replace $\phi(r)$ and get the following expression

$$\begin{aligned} \frac{\partial}{\partial\hat{d}}\hat{U}'(r) =& \phi'(r)\left(U_W-U_L\right)\left(\pi'(\hat{d}) - \pi'''(\hat{d})\pi(\hat{d})/\pi''(\hat{d})\right) \\ &- \phi'(r)\left(\pi'''(\hat{d})/\pi''(\hat{d})\right)U_L < 0. \end{aligned}$$

Since $U_L > 0$ and we assume $\pi'/\pi > \pi'''/\pi''$ and therefore $\pi' > \pi'''\pi/\pi''$, we see that both terms are negative. Thus, whenever $\hat{d} < d$, we have $\hat{U}'(r, \hat{d}) > \hat{U}'(r, d)$. A lower perceived distance reduces the marginal utility of taking risk. Denote the optimal risk choice by r^*. Therefore, $r^*(\hat{d}) > r^*(d)$, because $\hat{U}''(r^*) < 0$ in the optimum. Increased risk-taking $r^*(\hat{d})$ directly increases the variance of race time $\left[r^*(\hat{d})\right]^2$ but leaves expected race time unaffected, proving claims (i) to (iii). The increased variance of race time raises the probability to win. The latter equals $E\left[\pi(d + rz)\right] \tilde{=} \pi(d) + \pi''(d)\left[r^*(\hat{d})\right]^2/2$, applying the same Taylor approximation as above, see equation (6.4). This proves claim (iv). $\square$

The intuition behind Proposition 1 is the following. Taking more risk raises the variance of race time and thereby increases the expected chance of winning, since the probability to win π is a convex function of distance. If the distance to the leader is perceived as too small, the perceived gain from risk taking seems higher. This leads to higher risk taking and therefore a lower probability of finishing the race $\phi(r)$ and an increased variance of race time, given the athlete finishes the race.

The model makes an interesting prediction that an athlete with left-digit bias overestimates both the cost and benefit of taking more risk. The lower perceived distance increases the perceived utility of finishing the race $\pi(\hat{d} + rz)U_W + \left(1 - \pi(\hat{d} + rz)\right)U_L$. This raises the first term in (6.3), hence the athlete becomes more risk averse when the distance is perceived too low. However, when the probability to win π is sufficiently convex (as guaranteed by our assumption $\pi'/\pi > \pi'''/\pi''$ and supported by the data), the benefits of risk taking—the second term in the first order condition (6.3)—increase more than the costs.

To illustrate the trade-off that athletes face when choosing the risk level for the second run, consider two individuals trailing the leader by nine and ten hundredths of a second, respectively. In the presence of a left-digit bias, the *actual* difference in their time distance to the leader is smaller than the *perceived* difference. The athlete trailing the leader by 0.09 seconds perceives her distance to be significantly smaller. However, the perception of being closer to the victory has two effects: On the one hand, taking *more* risk in the second run appears to be more profitable since winning the race is perceived to be more likely. On the other hand, the athlete wants to take *less* risk in order not to crash and squander the good chance to achieve a high rank behind the winner. We find that the former effect dominates the latter both in our theoretical model and in the empirical analysis. The overestimation of the winning probability thus leads athletes to the left of a tenths-of-a-second cutoff to increased risk-taking in the second run. In Section 6.5.2 we document empirically that athletes with low left digits indeed have a higher variance of performance, while their average performance is unaffected. While we find supporting evidence for claims (i), (ii), and (iii), claim (iv) on the winning probability is not borne out in the data. Our analysis shows no significant impact of a left-digit bias on the chance of winning, given an athlete finishes the race. We come back to this issue in more detail in Section 6.5.2 below.

To conclude this section, we discuss two further aspects of the model. First, the model predicts that the probability of finishing the race decreases with the distance after the opening run if talent (ξ) strongly affects the survival probability (ϕ). The analysis of our data confirms this prediction: athletes with a larger distance to the leader are more likely not to finish the second run.[1] Second, while our model puts forward that athletes rationally behave based on biased information processing, an alternative explanation for the decreased survival probability ϕ below the threshold could be nervousness. Feeling closer to the leader could make the athlete more nervous, prompting more mistakes in the second run. In turn, the increased number of mistakes decreases the survival probability and raises the variance of final race times. However, nervousness would also worsen the average race time given the athlete finishes the race. This prediction cannot be observed in the data as we discuss in Section 6.5.4 in more detail.

[1] While athletes trailing by less than 25 hundredths of a second have a 3.45% probability of not finishing the race, the respective figure is 4.73% for athletes between 26 and 50 hundredths of a second, 5.30% for those trailing between 51 and 75 hundredths of a second, and 5.52% for athletes with a time distance between 76 and 100 hundredths of a second.

6.3 Data

6.3.1 World Cup Alpine Skiing

The first alpine skiing races were organised in the 1930s, but it was not until 1967 that the *Fédération Internationale de Ski* (FIS) launched the FIS World Cup. During the first couple of years, the disciplines included only slalom, giant slalom, and downhill races. In 1974, combined races were included, while super G was added to the FIS World Cup in 1983. The main interest of our paper is the question how athletes take into account information about their relative distance to the leader. We thus focus only on slalom and giant slalom races because their final standing is calculated by adding up the individual times of two separate runs. The time distance to the leader after the first run indicates how close athletes are to achieving a victory.

Alpine skiing provides a unique real-world setting to examine the effects of left-digit biases on subsequent risk behavior. All athletes are highly intrinsically motivated, yet the extrinsic motivation—in the form of monetary rewards and international fame—is likely to play a central role for individual performance as well.[1] Besides the prize money, success in World Cup races can also lead to better sponsorship contracts. Alpine ski races are, particularly in Europe, very popular, which makes competition fierce. Only a few junior athletes make it to the national World Cup team and among them only a small group is very successful.The goal in each race is to slide down a course in the fastest overall time. Each track consists of a series of gates. All of them have to be passed correctly, so that all athletes run the same course.

6.3.2 Data Set and Descriptive Statistics

We use a panel data set on 995 male and 869 female athletes in all 787 slalom and giant slalom ski races for the period of 1992–2014. The data include information on whether an athlete finished a race, the exact time (in hundredths of a second),

[1] Following Kahn [2000], we use sports data as an empirical laboratory for the evaluation of individual behavior. The benefit is that we have exact information on individual performance in a real-world setting with high stakes and competition.

the time difference to the winner, as well as gender, age, and discipline of competition. A detailed description of the full data set is provided by Legge and Schmid [2016]. The descriptive statistics for all relevant variables are shown in Table 6.1. In addition to the full sample statistics, we also provide all information based on the sample that only contains observations used for the estimation.

— Table 6.1 about here —

There are only minor differences between the two samples which is not surprising given the equal distribution of right digits (cf. Figure 6.3). On average athletes are 25.8 years old, do not finish the race with a probability of 5%, and have an average time distance to the leader of 2.02 seconds.

The numbers indicate that a victory is a likely outcome only for athletes who are in the top fifteen after the first run. While the average winning probability for a top fifteen athlete is 6.65%, it is only 0.03% for athletes outside of the top fifteen after the opening leg. Therefore, we should see only these athletes to respond to a high or low left digit in their time distance to the leader. Figure 6.3 in the Appendix shows several measures of success by athletes' rank after the first run.

6.4 Econometric Approach

Since World Cup alpine skiing is an outdoor event, external weather and snow conditions vary significantly over the course of a single race and can alter individual race times. However, the mere presence of unstable external conditions does not lead to cancellation and is broadly accepted as a natural source of variation among competitors. Thus, the impact of random wind, weather, and snow conditions is crucial and can even be amplified by the fact that individual race times critically depend on the performance in key sections of the course. An error in these sections, caused by external conditions, not only leads to an immediate time loss but also affects speed, and thus time, in the following sections. We refer to these external conditions as quasi-random noise and argue that it has sufficiently large effects on individual race times in order to randomly affect race times. In our estimation, this random noise is particularly relevant because we test whether athletes adopt a more risky behavior based on their time distance to the leader in the first run. This distance is affected by random weather shocks, which implies that athletes cannot locate themselves strategically to the left or right of the cutoff (i.e., a tenth-of-a-second threshold). The histogram shown in Figure 6.3 supports this assumption. There is a smooth distribution of left digits and all possible right digits are almost equally likely.[1]

— Figure 6.3 about here —

Our data set contains detailed information about whether athletes competed in a race and whether they successfully finished the race or not. We define survival (i.e., finishing the race) of athlete i in any race j as

$$S_{i,j} = \begin{cases} 1 & \text{if athlete } i \text{ successfully finished race } j \\ 0 & \text{if athlete } i \text{ did not successfully finish race } j. \end{cases} \tag{6.5}$$

Based on our considerations in Section 6.2, we expect survival rates to be a discontinuous function of the time distance to the leader after the opening run. If athletes are subject to a left-digit bias there should be a negative effect on survival if an athlete randomly achieves a distance with a low left digit. Comparing, for example, two athletes with almost identical distances 9 and 10 hundredths of a second, we suggest that the former should have a lower probability of survival. This is because the athlete with a low left digit underestimates the magnitude of the time distance to the leader.[2]

In order to estimate the effect of interest, we assume that except for the distinct effect of left digits, there is no reason why survival $S_{i,j}$ should be a discontinuous function of the distance to the leader. This is supported by balance tests shown in Table 6.2. The statistics indicate that athletes close to these cutoffs are not systematically different in their baseline characteristics.

— Table 6.2 about here —

[1]The fact that all right-digits in our sample are equally likely rules out the possibility of a non-uniform distribution of digits as in collections of many natural numbers [Benford, 1938].

[2]We provide a detailed description of the treatment variables in the Appendix.

In our estimation, we apply a regression discontinuity design with two different bandwidths. The narrow bandwidth includes only athletes with a right digit of 0 and those with right digit of 9. Both have almost the same time difference to the leader but the first one has a lower left digit and thus a larger *perceived* chance of being successful. Using this bandwidth, any athlete with a time distance to the leader of 9, 19, 29, etc. hundredths of a second is in the treatment group. For the control group, we take all athletes with a time distance of 10, 20, 30, etc. hundredths of a second. The broader bandwidth compares athletes with a right digit of 8 and 9 with those having a 0 and 1. Under the assumption that in general $S_{i,j}$ should be a continuous function of the distance to the leader, any observed discontinuity in $S_{i,j}$ at tenths-of-a-second cutoff levels is identified as the causal effect of the treatment. Using the narrow bandwidth, we estimate this effect, denoted by τ, by fitting the linear regression

$$\begin{aligned} S_{i,j} =& \alpha + \tau D[\mathrm{mod}(T_{i,j} - T_{1,j}, 10) = 9] \\ &+ \sum_{k=1}^{5} \gamma_k (T_{i,j} - T_{1,j})^k + \beta T_{i,j} + \mathbf{X}_{i,j}\,\delta + \varepsilon_{i,j} \end{aligned} \tag{6.6}$$

where $D[\mathrm{mod}(T_{i,j} - T_{1,j}, 10) = 9]$ is an indicator function, taking the value one if the modulo of the distance to the leader with respect to ten is equal to 9, $T_{i,j}$ denotes athlete i's time in the first run, $T_{1,j}$ is the time of the leader of the first run, and $\mathbf{X}_{i,j}$ captures individual characteristics: age, experience, as well as a prior successes. The error term, $\varepsilon_{i,j}$, is clustered at the athlete level.[1] The modulo finds the remainder of the division of the time distance by ten. For example, a distance of 0.39 seconds yields a modulo of 9. Note that equation (6.6) includes a fifth-order polynomial of athlete i's distance to the leader. In our analysis, we use this higher-order polynomial to control for nonlinear fits. However, the particular

[1] We can also cluster standard errors at the race-level and obtain virtually identical results. Note that we do not explicitly control for the presence of superstars in our main regressions. However, we find that —in line with previous research by Brown [2011]— this does not affect our results.

choice of the polynomial does not affect our estimates.[1]

An important potential bias in our results would arise if athletes with low and high left digits are systematically different with respect to pre-determined covariates. To explore this possibility, we compare the characteristics of treated and non-treated athletes. Balance tests reported in Table 6.2 show that all tested variables have very similar means and the differences between treatment and control group fall short of conventional levels of statistical significance. In addition, Figure 6.3 indicates that there is no clustering of observations around the cutoff. While left digits are roughly normally distributed, right digits are evenly distributed.[2] Following McCrary [2008] we argue that athletes cannot influence their location to the left or right of the threshold. For one thing, this is because they have no influence on the leader's race time. In addition, as we argue in Section 6.3, random weather shocks are sufficiently large to affect each individual race times. Overall the balance tests shown in Table 6.2 support the assumption that treatment is randomly assigned. Together with the evidence on the smooth distribution of the running variable depicted in Figure 6.3, we are confident that the main assumptions for identifying causal effects are satisfied.

We add control variables and fixed effects in some specifications to address previous research by Dohmen et al. [2011] suggesting that risk-taking behavior correlates significantly with individual characteristics and decreases, for example, with age. Moreover, the inclusion of covariates may rule out observable and time-fixed non-observable confounders and can improve the precision of the estimation [Frölich, 2007]. For a comparison, we report results both with and without using control variables.

[1] Adding any combination of polynomials up to the order of 15 yields virtually identical results. The fifth-order polynomial is chosen based on significance levels when regressing survival on distance [Lee and Lemieux, 2010].

[2] This feature of our data differs from previous studies using other settings in which the allocation of left digits is not entirely random (cf. Lacetera, Pope and Sydnor, 2012 or Englmaier, Schmoeller and Stowasser, 2013).

6.5 Results

6.5.1 Main Effects

In our main analysis, we test Proposition 1 that predicts that athletes change their behavior based on left digits in their distances to the leader of the first run. Before turning to the econometric estimates, we provide descriptive evidence suggesting that in fact survival rates differ between athletes with low or high left-digits. Figure 6.4 plots the average probability of finishing the race (henceforth, the probability of survival) by an athlete's right digit in the distance to the leader of the opening run. Each mean survival rate is based on approximately 2,000 observations.

— Figure 6.4 about here —

We observe that on average about 95 % of athletes successfully finish a race. This probability, however, differs substantially depending on the right-digit in the time distance to the leader. In particular, athletes with a relatively low *left*-digit, i.e. those with a *right*-digit of '8' or '9', exhibit a visibly lower probability of survival.

This simple empirical evidence suggests that athletes may respond to having low or high left digits in their time distance to the leader of the first run. We explore this in more detail by applying the econometric approach outlined in the previous section. Our main estimation results are shown in Table 6.3.

— Table 6.3 about here —

Using the full sample and different sets of controls, we find that a low left digit has a significant negative effect on the probability of survival. Applying the wide bandwidth (N=8,482), the results indicate that those athletes with a distance to the leader of 8, 9, 18, 19, 28, 29, etc. hundredths of a second are about 28.5 % (or 1.4 percentage points) more likely not to finish the second run than those with a distance of 10, 11, 20, 21, 30, 31, etc. hundredths of a second. When we choose the narrow bandwidth (comparing only digits 9 and 0; N=4,141), the effect is similar in magnitude while the significance is reduced due to the smaller sample size. Applying a regression discontinuity design, we exploit the fact that athletes quasi-randomly receive a low or high left-digit in their time distance to the leader. Therefore adding control variables to the regression should not affect our estimates. The observation that point estimates are not sensitive to the specification lends confidence to our results. The estimated effects in columns 2-5 of Table 6.3 are all very similar to the baseline results, even when adding race- or athlete-fixed effects.[1]

We can show that this discontinuity is present at multiple thresholds. In Figure 6.5, we plot the average survival rate for various time distances to the leader of the opening run. Each tenth-of-a-second is a threshold.

— Figure 6.5 about here —

We observe that in total, 16 (or 76%) differences in time window of the first two seconds are negative, 5 (24%) are positive.[2] The average difference in survival between athletes with a high left digit and athletes with a low left digit is 0.012 and thus very similar in magnitude to the main estimates obtained in Table 6.3. This indicates that the left-digit bias is present at multiple thresholds. In several cases, those athletes with a slightly smaller difference to the leader are more than 50% (or 2.5 percentage points) more likely to crash.[3]

6.5.2 Individual Success and Variance of Performance

The analysis thus far has shown that the left digit in an athlete's time distance to the leader affects her subsequent risk behavior, measured as the probability of not finishing the race. However, it remains unclear whether the left digit also affects individual race times or the probability of winning the race. Our previous

[1] The addition of athlete-fixed effects can be motivated by recent research on the determinants of risk preferences. For example, Cesarini et al. [2009] as well as Barnea, Cronqvist and Siegel [2010] find that up to one third of the variance in stock market participation and asset allocation can be attributed to genetic factors.

[2] We do not show confidence intervals. These are fairly large due to the fact that we only have about 80 observations for each threshold (e.g., combined observations for 8, 9, 10, 11).

[3] Note that we obtain very similar results when using a framework that controls for the time distance to the leader. Figure 6.4 in the Appendix shows predicted survival rates and differences in survival rates at multiple cutoffs based on a regression of survival on a fifth-order polynomial of the time distance variable and dummies for the left-digit cutoffs. Both levels and differences are virtually identical to the raw descriptive statistics shown in Figure 6.5.

findings suggest that athletes with a low *perceived* distance to the leader take high risks. A fraction of them does not finish the race. But what happens to those who successfully finish the race? It may be that their increased risk-taking pays off and leads to a better overall time and thus a higher probability of winning the race. However, it is also possible that the higher risk leads to errors that translate into worse final times, a large variance in race times, and a lower probability of winning.

— Table 6.4 about here —

To examine the effects of left digits on the set of finishers, we perform five separate tests reported in Table 6.4. The first column compares the winning probability of treatment and control group and shows the results of a test for equality of means. Individuals with a low left digit seem to have a slightly lower probability of winning the race but the difference is not statistically significant. Both the second and third column perform a test for equality of means for two standardised measures of time in the second run. Time measure 1 divides individual race time in the second run by the final second run time of the subsequent winner. Time measure 2 does the standardization by the final second run race time of the best athlete in the second run. Both measures are comparable in magnitude in treatment and control group and not statistically significant. These findings suggest that athletes taking more risk due to a left-digit bias are not more successful in the second run in case they do finish the race.

In order to explain this result we test whether increased risk-taking affects the variance of race times in the second run. Results in columns 4 and 5 indicate that the standard deviation of athletes with a low left digit is between 21.4% and 26.1% higher when compared to athletes with a high left digit. The variance-comparison tests indicate that both differences are significant at the 1% level. This suggests that the increased risk-taking behavior of athletes with low left digits does not change their *average* time. It does, however, lead to a situation in which some athletes who finish with few errors perform very well and finish with a low final time, while others who make errors finish with a large final time. This increases the tails of the distribution of race times in the second run and thus the respective standard deviation.

Overall, we find empirical evidence for claims (i), (ii), and (iii) of Proposition 1, namely that a low left digit (i) raises the probability of not finishing the second run, (ii) increases the variance of race time, but (iii) leaves the average race time unaffected. However, there is no evidence for claim (iv) that low left digits lead to an increase in the probability to win.[1]

6.5.3 Left-Digit Bias in Tournaments

The setting of World Cup alpine skiing is characterised by high stakes, fierce competition and experienced athletes. Our main results suggest that in this setting a left-digit bias is present, causing athletes to be more likely not to finish the race. In what follows, we provide a discussion of why large stakes, experience and competition do not eliminate the behavioral bias.

Previous research in psychology distinguishes two fundamentally different ways of thinking, commonly referred to as 'System 1' and 'System 2' [Kahneman, 2011; Stanovich and West, 2000]. The former deals automatically and quickly with signals and information. This occurs without effort and with no sense of voluntary control. In contrast, System 2 allocates attention to those mental activities which demand it because of their complexity.[2] In our setting of World Cup alpine skiing, the athlete's System 1 is used to recognise and judge the distance to the leader. This information is easily observable and athletes are familiar with this type of information. If an athlete, however, wants to know the distance to other ranks she has to actively compute that distance. This task is performed by System 2.

[1] One explanation for the absence of this result is the fact that the left digit bias seems to be most pronounced for athletes trailing the leader by 70 to 120 hundredths of a second as indicated by Figure 6.5. Only few of them are able to compensate their substantial time difference to the leader after the first run.

[2] Kahneman [2011] provides a simple illustration for the difference between System 1 and System 2. When an individual sees the image of an angry woman, for example, System 1 immediately recognises that the person in the picture is angry. It takes no effort to recognise the anger and individuals do not control whether or not to see the anger. The same happens with familiar and simple problems to which one has an immediate solution (e.g., 2+2=4). If individuals face, however, a complex problem like '17 x 24', System 2 takes over because solving this problem takes effort, people are usually not familiar with it and do not know the answer without spending time on calculating the solution.

A large body of research documents that behavioral biases such as heuristics arise in System 1. Thus, athletes relying on System 1 when dealing with time distances to the leader are prone to error. The automatic way of thinking tends to simplify information. One way of doing this is to concentrate on left-digits in any number. In theory, System 2 could intervene and prevent a left-digit bias. However, using System 2 requires conscious effort. As Kahneman [2011] explains, the common expression of 'paying attention' is apt. Individuals have a limited budget of mental resources. As a result, athletes cannot pay full attention to every information all the time.[1] However, if stakes are sufficiently high athletes should have a strong incentive to pay attention to the information they receive after the opening run. This reasoning has led prior research to question the existence of behavioral biases in the context of large stakes.

High Stakes and Behavioral Biases. — We consider a tournament setting with large stakes reflected not only in substantial prize money, but also in lucrative sponsorship contracts. The incentive structure in such a setting has been subject to previous research by Ehrenberg and Bognanno [1990]. An important question is whether high stakes—and the concentration thereof among the most successful athletes—yields higher effort levels and reduces behavioral biases.

As the results in Table 6.3 indicate, we observe a left-digit bias despite the large stakes present in World Cup alpine skiing. One way of explaining this finding is to refer to the concept of *ego-depletion*. Baumeister et al. [1998] argue that 'all variants of voluntary effort—cognitive, emotional, or physical—draw at least partly on a shared pool of mental energy'. If individuals exert a lot of physical or mental effort on one particular task, they are less likely to pay full attention to or exert full effort on a subsequent task. This phenomenon is called ego-depletion. If System 2 is exhausted because of some current or prior activity, System 1 takes over. The existence of this process has been demonstrated in numerous experiments. If individuals, for example, had to keep in mind a seven-digit number for one or two minutes they respond differently to various questions

[1] According to Kahneman [2011], 'Constantly questioning our own thinking would be impossibly tedious, and System 2 is much too slow and inefficient to serve as a substitute for System 1 in making routine decisions.'

or tasks. In particular, while being cognitively busy they are more likely to make superficial judgments [Kahneman, 2011].

In World Cup slalom races, athletes have to exert a lot of effort and pay full attention during the first run of a race. Compared to other disciplines, slalom races are considered the technical events of alpine ski racing. A course consists of more than fifty gates, all of which must be passed correctly at a speed of about 40 km/h. The vertical offset between gates is around 9 meters while the horizontal offset is about 2 meters. After the opening run, athletes' mental resources are depleted and they rely on System 1 to deal with information and signals. Thus when looking at the distance to the leader they suffer from a left-digit bias.

Large stakes are likely to magnify this bias. Prize money in World Cup alpine skiing is huge. The winner of a single race earns, on average, about $40,000 while the athlete ranked second only receives $23,000. Moreover, ski races attract a large audience of up to one million per race, thus leading to a considerable sponsorship market. These figures indicate that athletes are subject to significant pressure. At the same time, tiny mistakes can have huge effects on performance, thus causing financial implications for individual athletes. Choking under such immense pressure is a well-known phenomenon. In the seminal work by Baumeister [1984], choking is defined as increased pressure which raises the attention to individuals' own process of performance, thus disrupting the automatic nature of the execution. Experimental evidence suggests that a stressful environment is detrimental to the working memory [Beilock, 2008; Beilock and Carr, 2005]. In many fields of sports stakes are also large due to the presence of spectators. Dohmen [2008] investigates whether choking can be observed among professionals performing their usual tasks. In the setting of the German Premier football league Dohmen finds professional players to choke more frequently when playing a home match. Higher stakes or the importance of success, however, do not seem to be associated with more choking.

Turning to World Cup skiing, it is worth noting that pressure caused by other (trailing) contestants does not significantly differ between treated and non-treated athletes in our sample. The balance tests in Table 6.2 show that the number of athletes within a tenths-of-a-second time window after the first run does not differ between treated and non-treated individuals. Both athletes in the control and treatment group face on average one contestant ahead and behind them whose distance is one tenth of a second or less.

In order to investigate the role of pressure on performance and behavior, we first split our data by the prize money for the winner. In Table 6.5 we estimate the effect of low left-digits on survival in a sample of races with above-median prize money. The results of column 2 indicate that the left-digit bias is similar in magnitude to our baseline estimate. In high-prize races, athletes with a low left-digit in their distance to the leader are about 30.5% more likely not to finish the second run.

— Table 6.5 about here —

As a second test for the role of stakes we split our sample into two periods of the season. Public attention is exceptionally high at the beginning of a new season. Our results in column 3 indicate that the point estimate is also substantially larger for this sample. Moreover, we find similar evidence for races toward the end of the season. When only two races are left, stakes are usually higher since the overall classification is determined, sponsorship contracts are renewed, and athletes for the national team are selected.

Instead of mitigating behavioral biases, high stakes in the setting of World Cup skiing actually do not affect the presence of a left-digit bias. One explanation for this finding could be that high stakes cause mental stress and lead to higher effort level during the first run. This may cause an increased reliance on heuristics when dealing with information like time distances after the opening run.

Experience and Behavioral Biases. — Many behavioral biases were first documented in experimental settings. Participants of lab experiments, however, are usually unfamiliar with the tasks they are asked to solve. In contrast, contestants in World Cup tournaments usually have many years of experience. Following two studies by List [2003, 2004], individual experience could render athletes in our setting less likely to suffer from behavioral biases. List [2003] reports experiments in which subjects gained experience over the course of several weeks. Based on his findings he concludes that 'useful cognitive capital builds up slowly, over days or years' (List 2003, p.67). This experience results in a reduction, although not complete elimination, of the endowment effect.[1] In our setting one could also argue that more experienced athletes are less affected by a low left-digit.

We investigate this relationship between the left-digit bias and athletes' experience. The median athlete has about 60 World Cup races of experience.[2] This translates into six or more years of World Cup level experience. Moreover, virtually all athletes in our sample have participated in at least ten races (or one season) when we use them for our estimation. However, despite the large experience the behavioral bias does not vanish. In columns 4 and 5 of Table 6.5 we restrict the sample to athletes with above-median age or experience. In both cases the point estimate for the effect of low left-digits is very similar to our baseline estimate. Even when restricting the sample to athletes with more than 100 World Cup races, we obtain evidence of a large and significant left-digit bias.

This finding raises the question why experience does not eliminate the behavioral bias in our setting. While we cannot directly examine the mechanisms through which experience affects the left-digit bias, we present recent evidence on the causes of behavioral biases that is consistent with our findings. Following Kahneman [2011], 'As you become skilled in a task, its demand for energy diminishes. Studies of the brain have shown that the pattern of activity associated with an action changes as skill increases, with fewer brain regions involved. Talent has similar effects.' In other words, experience and talent make System 2 work more efficiently and quickly. A trained mathematician, for instance, can solve 17 x 24 much quicker than ordinary people. This advantage, however, is mostly limited to System 2. And because the left-digit bias arises in System 1, experience does not mitigate the problem. In contrast, experience could in theory actually magnify the bias. Having participated in numerous races may render looking at the classification and time distances into a routine task. Thus experienced athletes are likely to rely more on System 1 when dealing with time distances. System 1,

[1] In a subsequent study, List [2011] randomises market experience and obtains empirical findings that support the premise that market experience successfully eliminates the gap between willingness-to-accept and willingness-to-pay.

[2] Figure 6.5 in the appendix plots the distribution of experience.

however, is prone to making mistakes. In the setting of World Cup tournaments, however, all athletes irrespective of their experience are used to deal with information about performance differences. Hence we only find evidence that the left-digit bias does not vanish with experience and no support for the idea that it is either mitigated or magnified among experienced athletes.

Overall, we find that large stakes, competition and experience do not eliminate behavioral biases. First, experience is likely to render looking at the classification scheme into a routine task and thus may increase the use of heuristics. Second, large stakes and competition cause mental stress (i.e., pressure). Together with heavy physical activity, athletes are likely to rely on System 1 when processing information after the opening run. As a result, professionals do not get it right.

Learning Effects. — Given the result that individual experience does not reduce the existence of the left-digit bias, one might ask whether athletes learn from 'mistakes'. Are individuals less prone to error if, in prior races, they did not successfully finish the race after having (potentially) misinterpreted a time distance due to a low left-digit bias? We test this hypothesis by reducing the sample to all those athletes who were in the treatment group and did not finish the race before a given race j. Fitting the same regression as in our main estimation of Table 6.3, we find that the coefficient on the low left-digit is virtually identical in both groups (-0.013). Hence, we conclude that there is no evidence of learning effects. This is not surprising because, unlike in other settings (e.g., Hart, 2005, Malmendier and Nagel, 2011 or Stango and Zinman, 2014), athletes are not aware of the left-digit bias.

6.5.4 Alternative Explanations

Differences in Risk Preferences and Race Tracks. — One immediate concern with our findings could be that athletes in the treatment and control group have systematically different risk preferences. A large body of literature has investigated the determinants of risk preferences. Since we do not observe individuals' preferences, we can only infer them from prior behavior. In particular, we can compute probabilities of not finishing the race at the individual level. We take these probabilities into account to test whether individual risk preferences affect our results. In a first step, the balance tests in Table 6.2 indicate that there is no significant difference in observed probabilities of not finishing the race prior to a particular race. Athletes in the treatment group do not show systematically more risky behavior than athletes in the control group. In a second step, we add athlete-fixed effects to our baseline regression.[1] The estimates in column 4 of Table 6.3 show that this does not change our results. In contrast, the point estimates are virtually identical to the ones in our main specification. The fact that adding athlete-fixed effects does not alter our findings also rules out the possibility that our estimates suffer from an omitted variable bias (i.e., from not observing skill or talent).

A second concern addresses the differences across race tracks. In World Cup alpine skiing, race tracks differ substantially in terms of length and difficulty. Thus, the optimal level of risk chosen by the athletes varies significantly across race tracks. Moreover it is possible, for instance, that in some locations it is more likely for athletes ranked second or third to surpass the leader by means of an exceptional performance in the second run. We take these considerations into account by adding race-fixed effects to our baseline regression. The results are reported in Table 6.3. In column 3, race-fixed effects are added. This, however, does not alter the finding that low left-digits are correlated with a significantly lower probability of survival in the second run.

Reference Points. — Throughout our theoretical considerations we assume that athletes use victory (or better: the distance to the leader) as a natural reference point when making decisions about risk levels in the second slalom run. In a study by Köszegi and Rabin [2006], the authors argue that rational expectations can serve as reference points. Individuals facing reference-dependent choices use their expectations when making decisions. Empirical support for this idea is provided by Abeler et al. [2011].[2] There are good reasons to assume that

[1] Note that a consequence of including athlete-fixed effects is that we can only use observations of those athletes who are observed both as part of the control and the treatment group.

[2] In a related study, Gill and Prowse [2012] argue that a rational agent anticipates possible

some athletes in a World Cup contest do not expect to achieve a victory but rather aim for a podium or top-5 finish. While we cannot test for individual-specific reference points, we can examine whether athletes using reference points other than victory also suffer from a left-digit bias. This exercise, however, causes a major problem. In order to use, for example, rank 2 or 3 for a podium finish as reference point, athletes have to compute themselves the time distance. This involves the use of System 2 which suggests that we should not observe any significant difference in survival rates. In fact, results shown in Table 6.6 show no evidence of a left-digit bias concerning the time distance to other ranks.

— Table 6.6 about here —

It seems to be the very information about the distance to the leader that drives risk behavior in the second run. We can refer to our explanations in Section 6.5.3 to understand this zero result. Each athlete uses System 1 when dealing with the easily observable time distance to the leader. However, in order to know the distance to the second, she has to compute the time gap herself. This is carried out by System 2 which is not prone to behavioral biases such as heuristic thinking.

Information Availability. — We can test whether the availability of information is crucial for the behavioral bias to arise. To do this, we use the fact that time is not measured using the decade system. Since one minute is sixty seconds we can re-calculate all time distances from seconds to minutes. This turns, for example, 19 hundredths of a second into 0.0032 minutes. As before, we then take the modulus of this number rounded to the nearest integer (this is 2 in the example above) and code the treatment status according to athletes' right digit. If limited attention is the source of our findings, we should not observe any discontinuities in the survival function at thresholds using time distances expressed in minutes. In Table 6.6 we show that the effect is in fact insignificant. This gives

disappointment. In our case, this applies to athletes that are close to the leader after the opening run and want to avoid forfeiting the opportunity to win the race. Similarly, the leader of the opening run is likely to take into account the expected disappointment that would arise if she does not defend the first rank in the second run.

us confidence that the time distances actually shown to the athletes (measured in seconds) are the important signal.

Another way to test for the existence of the left-digit bias is to compare athletes around the one-second threshold. Throughout the paper, we assume that athletes pay full attention to the left-most digit in their time distance to the leader. This implies that athletes trailing the leader by 0.90 to 0.99 seconds should have a lower survival rate than those with a distance between 1.00 and 1.09 seconds. Using our data, we find that the former group has a 1.7 percentage points (or 34%) higher probability of not finishing the second run.

In addition to the time difference to the leader, each athlete is provided with the classification after the opening run. As shown in Figure 6.6 in the appendix, this classification includes each racer's time to finish the first run. We can use this and test whether there is evidence of a left-digit bias in the processing of this information. For example, consider the case in which the leader of the opening run finished with a time of 54.91 seconds, while the second and third have a time of 54.99 and 55.00 seconds, respectively. In this setting, our main specification would consider both second and third as part of the treatment group (their time differences are 0.08 and 0.09 seconds). However, we can also define the treatment and control group based on whether athletes share the same integer on the full second count. In the example above, the first two athletes have a full-second of 54 while the third one's time begins with 55.[1] We examine whether athletes misinterpret time differences because of a left-digit bias with respect to the time (not time difference) of the opening run.

Table 6.2 in the Appendix presents the results. We employ basically the same specification as in our main analysis. However, we replace the treatment variable as described above. The point estimates show that athletes close to the leader (i.e., trailing by less than 0.3 seconds) are significantly more likely not to finish the second run if they share the leader's count of full seconds.[2] This finding is

[1]We provide a detailed description of the different treatment variables in Table 6.1 in the Appendix.

[2]The cutoff of 0.3 seconds refers to the maximum time distance between the leader and trailing athletes. Our choice of 0.3 seconds as a cutoff is not crucial and we obtain very similar results using 0.2 or 0.4 seconds.

strongly in line with our previous results and suggests that there is a left-digit bias with respect to the time of the opening run as well.[1] The fact that this time is easily observable by athletes underscores the importance of information availability for the existence of a left-digit bias.

Alternative Digit Breaks. — A crucial question concerning our estimation strategy is whether the discontinuity in survival rates is a particular phenomenon of the 9-10 cutoff. Typically we compare, for example, athletes with a time distance to the winner of 9 versus 10 hundredths of second. If limited attention explains our findings above, we should not obtain significant results when using other digits for the cutoff. That is, there should be no difference in survival rates when comparing, for example, a distance of 10 versus 11 hundredths of a second. The descriptive data from Figure 6.4 suggests that there is only a discontinuity when comparing right-digits 8 and 9 with 0 and 1. Moreover, a placebo test in Table 6.6 confirms that none of the other possible thresholds exhibits the pattern we find at the cutoff when the left-digit changes. In particular, we test whether athletes perform rounding when processing information about time differences. This would imply that athletes perceive a distance with a right-digit of 4 (compared to a right-digit of 5) as significantly smaller. Following the same idea outlined in Section 6.2, we should find a discontinuity at the 4-5 cutoff. This, however, is not confirmed by the estimation result shown in Table 6.6.

Distance to Leader. — Following our theoretical model, we expect the left-digit bias to have a significant effect only among athletes close to the leader of the opening run. These athletes pay attention to their distance to the leader and may over-estimate their winning probability due to inattention to right digits. To test this, we split our sample of athletes based on to their time difference to the leader after the opening run. Each subsample includes athletes in a range of 150 hundredths of a second. We then run separate regressions for every subsample

[1] Not surprisingly, the variation is very limited if we add race-fixed effects to the estimation. Hence, the coefficient is insignificant in column 3. More important is the observation that even when estimating with athlete-fixed effects, we obtain a significant negative effect.

using our main specification.[1]

— Figure 6.6 about here —

The results in Figure 6.6 demonstrate that the treatment effect is significant only among athletes in close distance to the victory. Once we reduce the sample to athletes with low winning probabilities, the treatment effect is no longer significantly different from zero.

Nervousness. — Our theoretical and empirical analysis has stressed the idea that the left-digit bias we find is a form of rational strategic decision-making mixed with biased information processing. An alternative explanation is that the misperceived time distance to the leader of the first run leads to higher nervousness, which then translates into choking under pressure in a competitive environment with high stakes [Dohmen, 2008]. To disentangle rational decision-making mixed with biased information processing from nervousness, we explore the left-digit bias for a sample of athletes with low levels of nervousness. As a measure of nervousness, we use an athlete's average improvement in the second run over the career. Athletes who have proven to considerably increase their performance in the second run are very unlikely to suffer from nervousness. Table 6.3 in the Appendix reports the results of estimating the main specification for athletes in the lowest quartile of our nervousness measure. The estimated changes in the probability of not surviving for athletes with a low left-digit range from -54.0% to -56.0% and are thus even larger than the estimates found for the full sample. The finding that even racers with low levels of nervousness exhibit a substantial change in the probability of not finishing the competition helps us understand the main channel of our results. It suggests that athletes do not fully account for the actual time distance to the leader and act rationally based on this mis-perceived distance.

There are two other pieces of evidence that are not consistent with the hypothesis that athletes with a smaller perceived distance are simply nervous and

[1]This rolling regression procedure has the advantage that the point estimates remain relatively stable across subsamples with similar distances to the leader.

thus finish with a lower probability. First, nervousness should be greatest in the sample of athletes trailing the leader by a very small distance. However, we do not find any particularly large effect among those athletes as documented in Figure 6.5. Second, nervousness should lead to a worse performance in general. In our context, we would expect athletes with a low left-digit to perform worse in the second run if nervousness is the main channel. Yet the prediction that average performance decreases as a consequence of nervousness induced by the left-digit bias cannot be supported by the data. Athletes to the left and right of a tenth-of-a-second threshold show a very similar average performance in the second run.

6.6 Conclusion

This paper investigates the effects of heuristic information processing and left-digit biases on risk-taking behavior in a setting of high stakes, competition, and experienced professionals. By applying a regression discontinuity design in a sample of professional World Cup alpine ski athletes, we present new empirical evidence that individuals misinterpret performance differences even in a professional setting. We find that athletes with a low perceived distance after the first run adopt a more risky strategy in the second run. Our estimates suggest that the probability of not finishing the race increases by up to 28% when comparing individuals with lower and higher left-digits (e.g., nine compared to ten hundredths of a second).

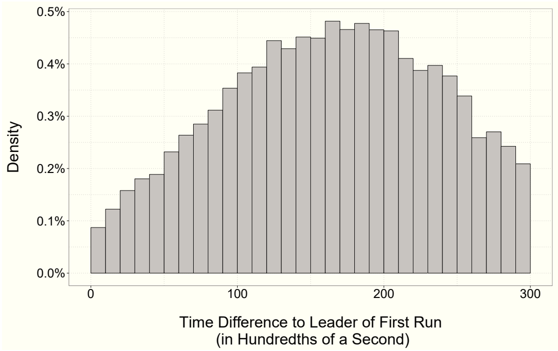
Our findings show that limited attention can be present even in a setting of professional athletes who compete for large prizes. Alpine ski athletes are aware of the fact that tiny differences in their distance to the leader after the first run hardly matter for their prospects in the second run. However, when comparing individuals with arguably similar distances, we find large discontinuities in survival rates. Our results are robust to the inclusion of several control variables as well as race- and athlete-fixed effects. In addition, a large set of robustness checks gives us confidence that limited attention is the source of the significant effect of low left-digits on survival.

Our findings have implications beyond alpine skiing as they add to the understanding of the causes of individual risk behavior that is crucial for many economic questions. Limited attention is likely to affect our everyday risk behavior. Prior research by Barber and Odean [2008] highlights the role of salience of stocks for attention-driven buyers. Due to a left-digit bias traders may interpret the magnitude of a 0.9% stock market change in a different way than a 1.0% change. This can then have implications for the amount of risk they are willing to take, even in the presence of large stakes. The behavioral bias we document may also have an impact in other areas. Our results suggest that individuals may have a different perception of a 90 km/h speed limit when compared to a 100 km/h speed limit. As a consequence, risk behavior under the two regimes can differ substantially. These and other aspects of limited attention are left to future research.

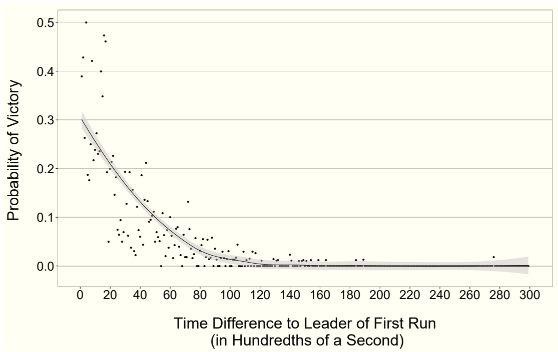
Figures and Tables

Figure 6.1: Distance to the Leader and Winning Probability

(A) Distribution of Time Differences

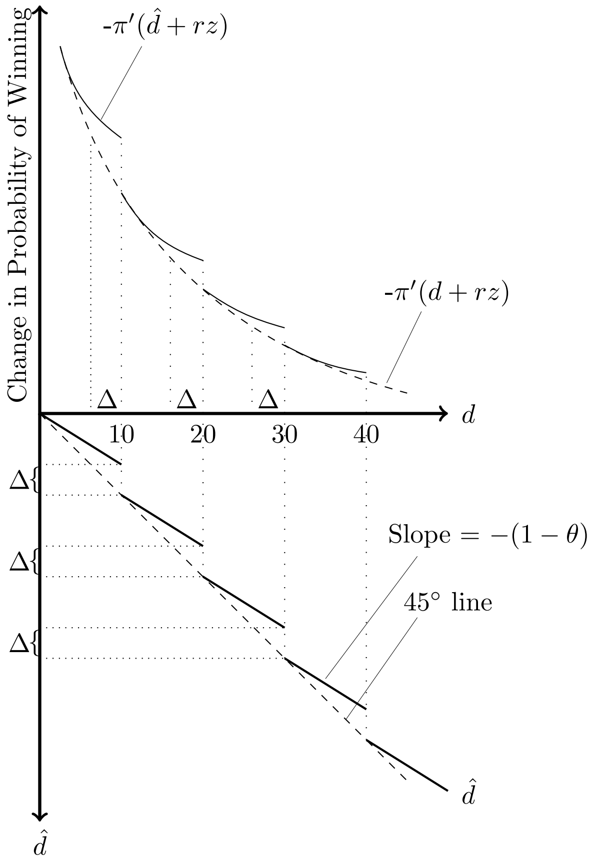


(B) Prob. of Victory by Difference to the Leader



Note: The figure in panel (A) shows the distribution of time differences to the leader after the opening run using the full sample. The figure in panel (B) plots the average probability of winning the race for each hundredth of a second. The solid line is a local linear regression, while the shaded area depicts a 95% confidence interval.

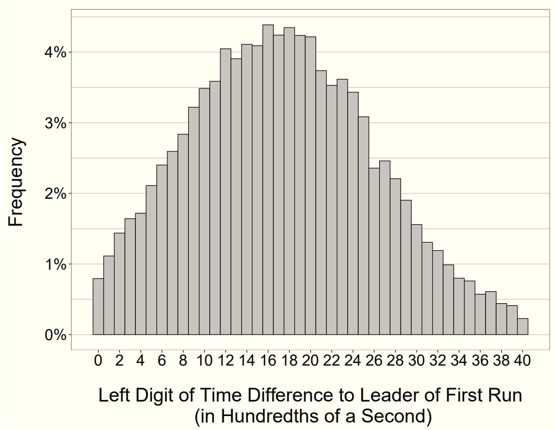
Figure 6.2: Perceived Distance and the Probability of Winning



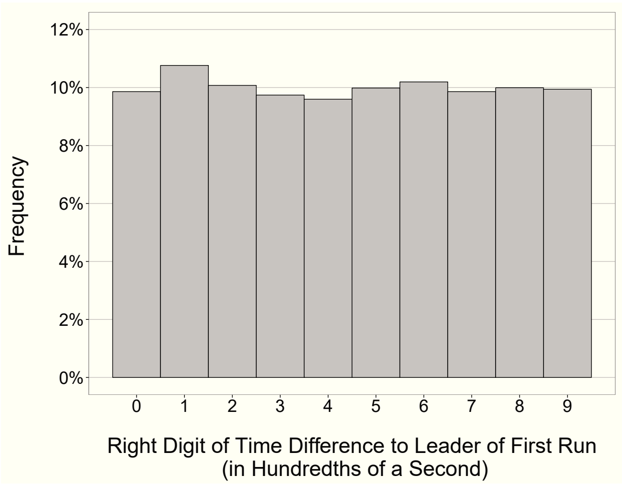
Note: The upper half of the figure illustrates how the discontinuity in the perceived distance $\hat{d}$ causes a discontinuity in the change of the winning probability, $-\pi'(\hat{d}+rz)$, drawn for a realization of $z=0$. In lower half we show the discontinuities between actual distance d (x-axis) and perceived distance $\hat{d}$ (y-axis).

Figure 6.3: Distribution of Left- and Right-Digits

(A) Histogram of Left Digits

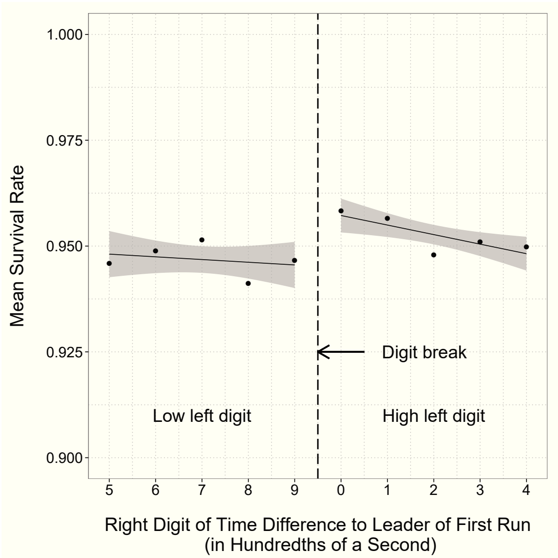


(B) Histogram of Right Digits



Note: The figure shows two histograms of the time difference to the leader after the first run. Panel (A) depicts the left-digit of an athlete's time difference to the leader of the first run, while panel (B) depicts the right-digit of the time difference to the leader of the first run. For example, the time distance of 0.27 seconds can be decomposed into a left digit of 2 and a right digit of 7.

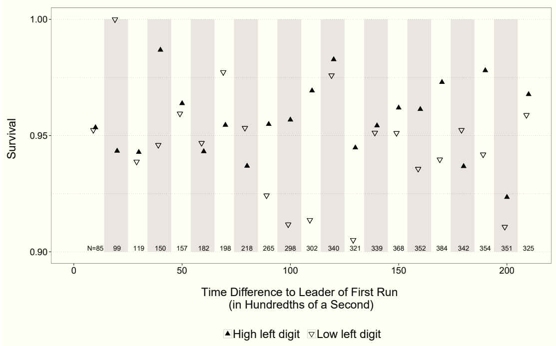
Figure 6.4: Survival Rate by Right-Digit



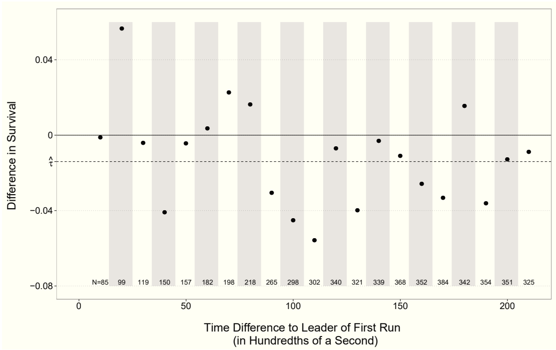
Note: The dots in the figure depict the average survival rate for every right digit in the time difference to the leader of the first run. In addition, we show a linear fit for both low and high left digits. Note that we remove all athletes with a distance of 0 to 5 hundredths of a second to avoid that slightly more skilled athletes are to the right of the digit break. The shaded area indicates the 90% confidence interval.

Figure 6.5: Discontinuities in Average Survival Rates

(A) Avg. Survival Rate by Distance to the Leader

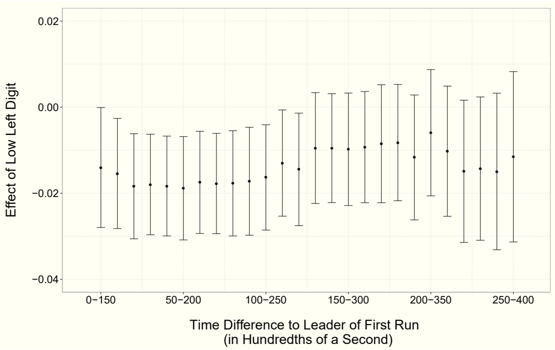


(B) Difference in Average Survival Rate by Time Difference to the Leader



Note: The figure in Panel (A) shows the average survival rate for various time distances to the leader of the opening run. The shaded and white bars indicate tenths-of-a-second brackets. For each bracket, we show the average survival rate slightly below (right digits 8 and 9) as well as slightly above (right digits 0 and 1) for each digit break. For example, the first two triangles on the left hand side show the average survival rate of athletes trailing the leader by 8 and 9 hundredths of a second (white triangle) as well as the survival rate for athletes with a difference of 10 and 11 hundredths of a second (black triangle). The figure in Panel (B) shows the difference between the two survival rates for each bracket. The horizontal dashed line depicts the estimated treatment effect, $\hat{\tau} = 0.014$, from our main regression in Table 6.3. At the bottom of both figures, we show the numbers of observation for each bracket.

Figure 6.6: Treatment Effect by Distance to the Leader



Note: The figure shows the estimated treatment effect (y-axis) for samples restricted by the distance to the leader (x-axis). The estimation sample includes all athletes with a right digit of 8, 9, 0, and 1 in their time distance to the leader of the first run, expressed in hundredths of a second. Confidence intervals at the 95% level are shown.

Table 6.1: Descriptive Statistics

	Athletes in Estimation Sample					Athletes in Full Sample				
Variable	Mean	SD	Min	Max	N	Mean	SD	Min	Max	N
Survival	0.95	0.22	0	1	8,482	0.95	0.22	0	1	13,225
Dist. to Leader	183.30	99.83	1	2959	8,482	171.28	102.47	0	1892	13,225
Pos. in 1st Run	14.92	7.89	2	30	8,482	14.19	8.43	1	30	13,225
Final Position	13.47	8.04	0	31	8,482	12.88	8.32	0	31	13,225
Age	25.79	3.72	16.89	37.64	8,482	25.88	3.7	15.82	37.65	13,225
Male	0.50	0.5	0	1	8,482	0.51	0.5	0	1	13,225
Experience	36.70	26.49	1	151	8,482	37.31	26.32	1	154	13,225
No. Podiums	7.46	14.46	0	101	8,482	8.14	14.73	0	110	13,225
No. Victories	2.57	6.12	0	58	8,482	2.87	6.24	0	53	13,225

Note: The table shows descriptive statistics for all relevant variables. Columns (2)–(6) show statistics for the sample used for the main regression analysis with a wide bandwidth. This includes all athletes with a right digit of 8, 9, 0, and 1 in their time distance to the leader of the first run, expressed in hundredths of a second. Columns (7)–(11) show statistics based on the full sample. Experience is measured by the number of races in the discipline of competition (slalom and giant slalom).

Table 6.2: Balance Tests

	Mean Treatment	Mean Control	Difference	p-value
A: Athlete Characteristics				
Age	26.19	26.19	-0.00	0.97
Male	1.50	1.50	-0.00	0.95
Experience	36.58	36.83	-0.25	0.67
B: Risk Preferences				
Average Survival	0.96	0.96	-0.00	0.29
Survival in Last Race	0.95	0.95	0.00	0.61
C: Overconfidence				
Athlete's No. of Victories	2.58	2.56	0.02	0.89
Athlete's No. of Podiums	7.47	7.44	0.03	0.92
D: Athlete Skill				
Victories in Career	5.68	5.30	0.38	0.20
Podiums in Career	16.06	15.36	0.70	0.23
E: Competition				
Total Podiums in Top 5	73.21	72.83	0.38	0.70
Total Victories in Top 5	27.42	27.19	0.23	0.58
Total Podiums in Top 10	118.71	118.51	0.20	0.89
F: Pressure				
Mass of Athletes Ahead	1.02	1.01	0.01	0.63
Mass of Athletes Behind	1.01	1.02	-0.00	0.92
FIS World Cup Points	159.51	152.73	6.78	0.18
First Prize Possible	0.32	0.32	-0.00	0.95

Note: The table shows mean comparisons (t-tests) for all relevant pre-treatment variables. The sample includes all athletes with a right digit of 8, 9, 0, and 1 in their time distance to the leader of the first run, expressed in hundredths of a second. Age is the exact age in years at the time of the race, experience is measured by the total number of races prior to the race, survival in last race indicates having successfully finished the preceding race, victory and podium measure the total number of an athlete's victories and podiums prior to the race, and mass of athletes ahead (behind) is the total number of athletes who are leading (lagging) by 10 hundredths of a second after the first run.

Table 6.3: Effect of Low Left Digits on Survival

	Dependent variable = 1 if athlete finished the race			
	Logit Estimation		OLS Estimation	
	(1)	(2)	(3)	(4)
A: Wide Bandwidth				
Low Left-Digit	-0.014***	-0.014***	-0.013***	-0.014***
	(0.004)	(0.004)	(0.005)	(0.005)
Change in Probability of not Surviving	-28.0%	-28.0%	-26.0%	-28.0%
Observations	8,482	8,482	8,482	8,182
R-squared	0.001	0.008	0.149	0.005
B: Narrow Bandwidth				
Low Left-Digit	-0.011*	-0.012*	-0.010	-0.011*
	(0.006)	(0.006)	(0.007)	(0.007)
Change in Probability of not Surviving	-22.0%	-24.0%	-20.0%	-22.0%
Observations	4,141	4,141	4,141	3,821
R-squared	0.006	0.010	0.231	0.005
Controls	-	Yes	Yes	Yes
Fixed Effects	-	-	Race	Athlete

Note: The table shows the results of eight separate regressions. In Panel A, we estimate the effect of a low left digit on the probability of survival using the wide bandwidth. The sample includes all athletes with a right digit of 8, 9, 0, and 1 in their time distance to the leader of the first run, expressed in hundredths of a second. In Panel B, we estimate the same regression using the narrow bandwidth that includes all athletes with a right digit of 9 and 0 in their time distance to the leader of the first run. Thus, in the latter specification, we estimate the effect of digit 9 on the probability of survival. Controls include age, experience, and prior successes. Columns 1 and 2 report marginal effects, while Columns 3-4 show OLS estimates. When applying athlete fixed effects, the sample is reduced to athletes with at least one observation in the treatment and control group. Standard errors (in parentheses) are clustered at the athlete level. Significance at the 10% level is indicated by *, at the 5% level by **, and at the 1% level by ***.

Table 6.4: Effect on Performance and Variance of Performance

	Individual Performance			Variance of Performance	
	(1)	(2)	(3)	(4)	(5)
	Mean Probability of Winning	Mean Time 2nd Run Measure 1	Mean Time 2nd Run Measure 2	Std. Dev. Time 2nd Run Measure 1	Std. Dev. Time 2nd Run Measure 2
High Left-Digit	0.019	1.017	1.024	0.023	0.022
Low Left-Digit	0.015	1.018	1.025	0.029	0.028
Difference	-0.004	0.001	0.001	0.005	0.006
P-value	0.169	0.246	0.264	0.000	0.000

Note: The table shows the results of five different tests between treatment and control group. All samples include athletes with a right digit of 8, 9 (low left digit) as well as 0 and 1 (high left digit) in their time distance to the leader of the first run, expressed in hundredths of a second. Columns (1) to (3) report the results of a t-test for equal means. The variables are the probability of winning the race (Column 1), the time in the second run scaled by the time of the subsequent winner (Column 2), and the time in the second run scaled by the time of the best second run athlete (Column 3). Columns (4) and (5) report the results from a variance-comparison tests for the two time measures.

Table 6.5: Effect with High Stakes and Individual Experience

	Dependent variable equals 1 if athlete finished the race				
	Baseline	Estimation sample split by			
	Regression	Prize Money	Early Season	Age	Experience
	(1)	(2)	(3)	(4)	(5)
Low Left-Digit	-0.013***	-0.014***	-0.052**	-0.013**	-0.015***
	(0.004)	(0.005)	(0.022)	(0.006)	(0.005)
Change in Prob. of not Surviving	-26.0%	-28.0%	-104.0%	-26.0%	-30.0%
Observations	8,482	7,455	475	4,766	6,246
R-squared	0.015	0.018	0.054	0.003	0.013
Controls	Yes	Yes	Yes	Yes	Yes

Note: The table shows the results of five separate logistic regressions. We always estimate the effect of a low left digit on the probability of survival and report marginal effects. All estimation samples include athletes with a right digit of 8, 9, 0, and 1 in their time distance to the leader of the first run, expressed in hundredths of a second. In Column (1) we report the baseline regression from Table 6.3. In Column (2) we restrict the sample to races with above-median prize money for the winner (>35,000 CHF). In Column (3) we restrict the sample to races at the beginning of the season when 10% (or less) of the races have been completed. In Column (4) we only consider athletes with above-median age (25 years or more), while in Column (5) we restrict the sample to athletes with above-median experience (18 races within discipline). Standard errors (in parentheses) are clustered at the athlete level. Significance at the 10% level is indicated by *, at the 5% level by **, and at the 1% level by ***.

Table 6.6: Information Availability and Placebo Tests

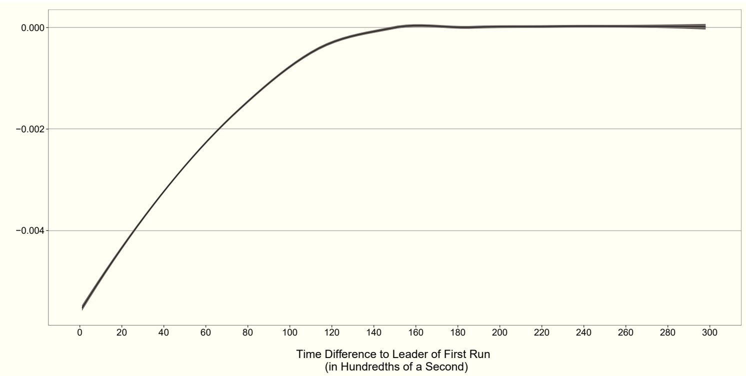
	Dep. Variable equals 1 if athlete finished the race			
	Distance 2nd	Time in Min.	Digits 0-1	Digits 4-5
Low Left-Digit	-0.002	0.000	-0.001	-0.004
	(0.007)	(0.005)	(0.006)	(0.007)
Change in Prob. of not Surviving	-4.0%	0.0%	-2.0%	-8.0%
Observations	3,970	6,907	4,313	4,095
R-squared	0.002	0.004	0.001	0.001
Controls	Yes	Yes	Yes	Yes

Note: The table shows the results of four separate OLS regressions. We always estimate the effect of a low left digit on the probability of survival. In Column (1), we use the left-digit of the time difference to the athlete on the second position after the first run. This information is not given to the athletes but has to be computed individually. In Column (2), time distances are converted to minutes. Note that after the conversion a right digit of 9 is impossible. Thus, we compare 8 (treated) versus 0 and 1 (control). Because the converted distances are never shown to athletes, left-digits should not have a significant effect on risk-taking. In Columns (3) and (4), we define treatment as having a right digit of 1 (or 5) versus 0 (or 4). Controls include age, experience, and prior successes. Standard errors (in parentheses) are clustered at the athlete level. Significance at the 10% level is indicated by *, at the 5% level by **, and at the 1% level by ***.

Appendix

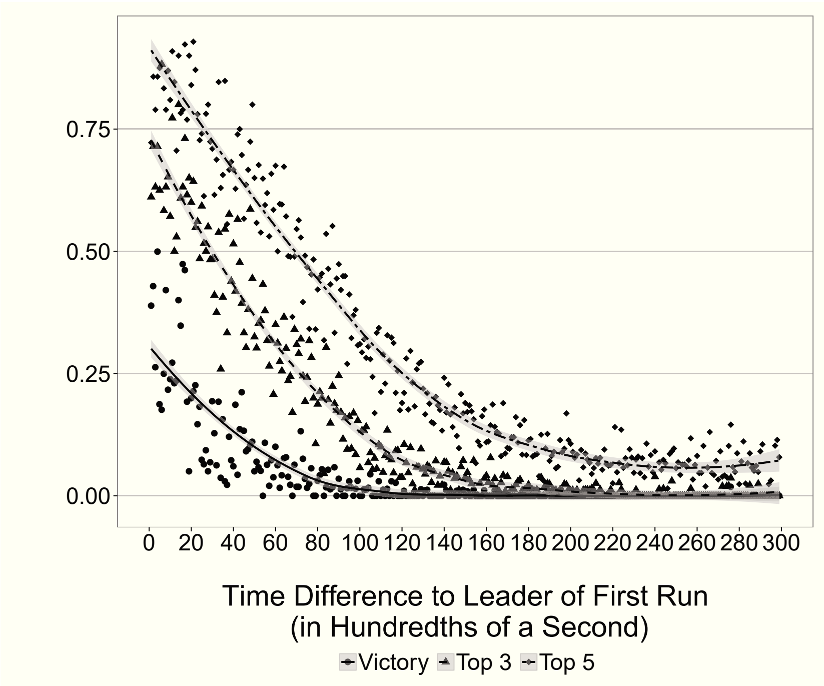
Additional Figures

Figure 6.1: First Derivative of the Winning Probability Function $\pi'(\cdot)$



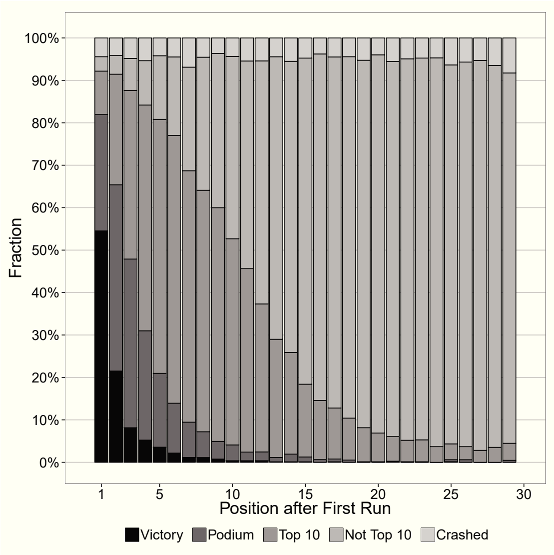
Note: The figure shows an estimate of the first derivative of the winning probability function, $\pi'(\cdot)$. The relationship is estimated non-parametrically. It is based on the actual time differences of the winning probability between athletes with different time differences to the leader of the first run as depicted in Panel (B) of Figure 6.1.

Figure 6.2: Final Position by Distance to Leader after First Run



Note: The figure shows the average probability of finishing first (solid line), in the top three (dashed line), and in the top five (twodashed line) for every time distance to the leader after the first run, expressed in hundredths of a second.

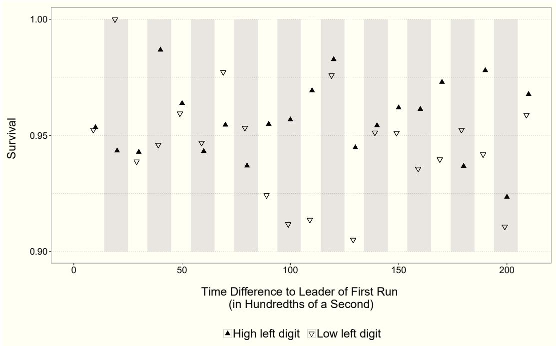
Figure 6.3: Relationship Position after First Run and Final Position



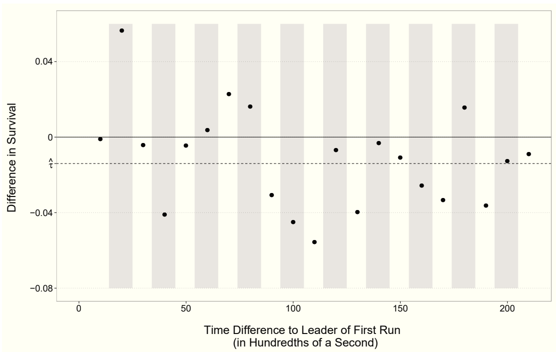
Note: The figure shows the final race position for every position after the first run. The first bar (black) is the fraction of athletes who won the race. The second bar is the fraction of athletes who finished on the podium (but did not win the race). The third bar is the fraction of athletes who made it into the top ten (but not on the podium). The fourth bar is the fraction of athletes who are outside of the top ten. The fifth bar (light gray) is the fraction of athletes who did not finish the race.

Figure 6.4: Residual Plot

(A) Residual of Average Survival Rate by Time Difference to Leader

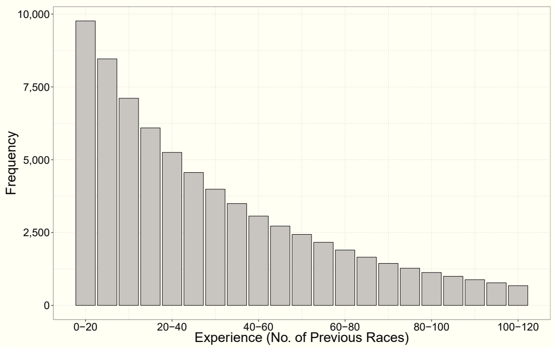


(B) Difference in Residuals of Average Survival Rate by Time Difference to Leader



Note: The figure in Panel (A) shows a residual plot from regressing survival probabilities on the time distance to the leader (including polynomials up to an order of five) plus dummy variables for being left or right of a tenths-of-a-second. The shaded and white bars indicate tenths-of-a-second brackets. For each bracket, we show the residuals slightly below (right digits 8 and 9) as well as slightly above (right digits 0 and 1) the digit break. For example, the first two triangles on the left hand side show the residuals for athletes trailing the leader by 8 and 9 hundredths of a second (white triangle) as well as the residuals for athletes with a difference of 10 and 11 hundredths of a second (black triangle). The figure in Panel (B) shows the difference between the two residuals for each bracket. The horizontal dashed line depicts the estimated treatment effect, $\hat{\tau} = 0.014$, from our main regression in Table 6.3. At the bottom of both figures, we show the numbers of observation for each bracket.

Figure 6.5: Distribution of Athlete Experience



Note: The figure plots the distribution of athlete experience. Each completed race within the slalom and giant slalom discipline counts as one unit of experience. The median of the distribution is 24. Across disciplines, the distribution has a similar shape with a median of about 60 races.

Figure 6.6: Example of FIS World Cup Results Table

Rank	Bib	FIS Code	Name	Year of Birth	NSA	Time	Diff.	Ski
1	2	491879	SALARICH Joaquim	1994	SPA	55.44		
2	6	481103	ANDRIENKO Aleksander	1990	RUS	56.67	1.23	Rossignol
3	10	30149	SIMARI BIRKNER C.	1980	ARG	56.80	1.36	
4	17	430429	BYDLINSKI Maciej	1988	POL	56.99	1.55	Atomic
5	13	40594	PERAUDO Ross	1992	AUS	57.11	1.67	
6	16	320293	KYUNG Sung-hyun	1990	KOR	57.12	1.68	Salomon
7	21	30266	GASTALDI Sebastiano	1991	ARG	57.24	1.80	Rossignol
7	12	92562	PRISADOV Stefan	1990	BUL	57.24	1.80	Fischer
9	34	221213	RAPOSO Charlie	1996	GBR	57.34	1.90	
10	50	400237	MEINERS Maarten	1992	NED	57.36	1.92	Rossignol

Note: The figure shows an example of how the FIS plots the results of an opening slalom run. The important performance measures for our analysis are shown in column 7 ('Time') and 8 ('Diff.').
Source: http://data.fis-ski.com/pdf/2015/AL/0219/2015AL0219RLR1.pdf

Additional Tables

Table 6.1: Generic Results and Treatment Variables

Rank	Name	Time	Distance	Low Left-Digit Narrow BW	Low Left-Digit Wide BW	Same Full Second
1	A	54.91	-	-	-	-
2	B	54.98	0.07	-	-	T
3	C	54.99	0.08	-	T	T
4	D	55.00	0.09	T	T	C
5	E	55.01	0.10	C	C	C
6	F	55.02	0.11	-	C	C

Note: The table shows a generic result after an opening run as well as the three different treatment variables used in the analysis. Being part of the treatment group is denoted by a 'T', while the control group is denoted by 'C'.

Table 6.1 shows a generic result for an opening run of a slalom race. All six contestants shown have a very similar performance with athletes B to E trailing the leader by up to 0.11 seconds. The first treatment variable ('Low Left-Digit, Narrow BW') is based on the time distance to the leader. In particular the *right* digit in the time distance determines the treatment status. If the right digit is 9, the athlete is part of the treatment group. In contrast, any athlete with a right digit of 0 is part of the control group. In a similar vein, the second treatment variable ('Low Left-Digit, Wide BW') is defined. The only difference is that the wide bandwidth includes right digits 8 for the treatment and 1 for the control group.

The definition of the third treatment variable is based on each athlete's time (shown in column 3). Any athlete whose time distance to the leader is less than 0.3 seconds is either part of the treatment or control group. To be considered as treated, an athlete must have the same full second count as the leader. In the example above, athletes B and C have a time of 54 seconds plus some hundredths of a second. In contrast, athletes D, E and F have a time that begins with 55. Hence, the former define the treatment and the latter the control group.

Table 6.2: Left-Digit Bias in Opening Run Time

	Dependent variable = 1 if athlete finished the race			
	Logit Estimation		OLS Estimation	
	(1)	(2)	(3)	(4)
Same Full Seconds in Opening Run	-0.033*** (0.009)	-0.041*** (0.011)	-0.022 (0.052)	-0.033*** (0.013)
Change in Probability of not Surviving	-66.0%	-82.0%	-44.0%	-66.0%
Observations	691	691	691	689
R-squared	0.006	0.021	0.717	0.023
Controls	-	Yes	Yes	Yes
Fixed Effects	-	-	Race	Athlete

Note: The table shows the results of four separate regressions. We estimate the effect of having the same number of full seconds as the leader of the opening run on survival in the second run. Treated athletes have the same full seconds and trail the leader by less than 0.3 seconds while the control group has a different full second (but still lagging 0.3 seconds at most). Controls include age, experience, and prior successes. Columns 1 and 2 report marginal effects while columns 3-5 show OLS estimates. When applying athlete fixed effects, the sample is reduced to athletes with at least two observations in the estimation sample. Standard errors (in parentheses) are clustered at the athlete level. Significance at the 10% level is indicated by *, at the 5% level by **, and at the 1% level by ***.

Table 6.3: Rational Decision-Making versus Nervousness

	Dependent variable = 1 if athlete finished the race			
	Logit Estimation		OLS Estimation	
	(1)	(2)	(3)	(4)
Low Left-Digit	-0.028**	-0.027**	-0.022	-0.035**
	(0.012)	(0.012)	(0.015)	(0.014)
Change in Probability of not Surviving	-56.0%	-54.0%	-44.0%	-70.0%
Observations	2,153	2,153	2,153	1,988
R-squared	0.003	0.019	0.309	0.011
Controls	-	Yes	Yes	Yes
Fixed Effects	-	-	Race	Athlete

Note: The table shows the results of four separate regressions using a sample of athletes who are unlikely to suffer from nervousness. As a measure of nervousness we use an athlete's average improvement in the second run over the career. The sample includes only all athletes in the highest quartile of this measure how finished the first run with a right digit of 8, 9, 0, and 1 in their time distance to the leader of the first run, expressed in hundredths of a second. Standard errors (in parentheses) are clustered at the athlete level. Significance at the 10% level is indicated by *, at the 5% level by **, and at the 1% level by ***.

References

Aaberge, Rolf, Anthony B. Atkinson, and Jorgen Modalsli. 2013. "The Ins and Outs of Top Income Mobility." *IZA Discussion Paper No. 7729.*

Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman. 2011. "Reference Points and Effort Provision." *American Economic Review*, 101(2): 470–492.

Abraham, Arpad. 2008. "Earnings Inequality and Skill-Biased Technological Change with Endogenous Choice of Education." *Journal of the European Economic Association*, 6(2/3): 695–704.

Acemoglu, Daron. 1998. "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality." *The Quarterly Journal of Economics*, 113(4): 1055–1089.

Acemoglu, Daron. 1999. "Changes in Unemployment and Wage Inequality: An Alternative Theory and Some Evidence." *American Economic Review*, 89(5): 1259–1278.

Acemoglu, Daron. 2002*a*. "Directed Technical Change." *The Review of Economic Studies*, 69(4): 781–809.

Acemoglu, Daron. 2002*b*. "Technical Change, Inequality, and the Labor Market." *Journal of Economic Literature*, 40(1): 7–72.

Acemoglu, Daron, and David H. Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics.* . 4 ed., , ed. Orley Ashenfelter and David Card, Chapter 12, 1043–1171. Amsterdam:Elsevier North Holland.

Acemoglu, Daron, and David H. Autor. 2012. "What Does Human Capital Do? A Review of Goldin and Katz's The Race between Education and Technology." *Journal of Economic Literature*, 50(2): 426–463.

Acemoglu, Daron, and Fabrizio Zilibotti. 2001. "Productivity Differences." *The Quarterly Journal of Economics*, 116(2): 563–606.

Acemoglu, Daron, and Joshua Linn. 2004. "Market Size in Innovation: Theory and Evidence From the Pharmaceutical Industry." *The Quarterly Journal of Economics*, 119(3): 1049–1090.

Acemoglu, Daron, Philippe Aghion, and Giovanni L. Violante. 2001. "Deunionization, Technical Change and Inequality." *Carnegie-Rochester Conference Series on Public Policy*, 55: 229–264.

Adams, Douglas. 2002. *The Salmon of Doubt: Hitchhiking the Galaxy One Last Time.* New York:Random House.

Aghion, Philippe, and Peter Howitt. 1992. "A Model of Growth Through Creative Destruction." *Econometrica*, 60(2): 323–351.

Aghion, Philippe, Eve Caroli, and Cecilia García-Peñalosa. 1999. "Inequality and Economic Growth: The Perspective of the New Growth Theories." *Journal of Economic Literature*, 37(4): 1615–1660.

Aghion, Philippe, George-Marios Angeletos, Abhijit Banerjee, and Kalina Manova. 2010. "Volatility and growth: Credit constraints and the composition of investment." *Journal of Monetary Economics*, 57(3): 246–265.

Aghion, Philippe, Philippe Askenazy, Nicolas Berman, Gilbert Cette, and Laurent Eymard. 2012. "Credit constraints and the cyclicality of R&D investment: Evidence from France." *Journal of the European Economic Association*, 10(5): 1001–1024.

Aghion, Philippe, Ufuk Akcigit, Antonin Bergeaud, Richard Blundell, and David Hemous. 2015. "Innovation and Top Income Inequality." *NBER Working Paper #21247.*

Aguiar, Mark, and Erik Hurst. 2007. "Measuring Trends in Leisure: The Allocation of Time over Five Decades." *The Quarterly Journal of Economics*, 122(3): 969–1006.

Aguiar, Mark, and Mark Bils. 2015. "Has Consumption Inequality Mirrored Income Inequality?" *American Economic Review*, 105(9): 2725–2756.

Ahituv, Avner. 2001. "Be Fruitful or Multiply: On the Interplay between Fertility and Economic Development." *Journal of Population Economics*, 14(1): 51–71.

Ai, Chunrong, and Edward C. Norton. 2003. "Interaction terms in logit and probit models." *Economics Letters*, 80(1): 123–129.

Airaudo, Marco, and Luca Bossi. 2015. "Trickle-Down Consumption, Monetary Policy, and Inequality."

Akerman, Anders, Elhanan Helpman, Oleg Itskhoki, Marc-Andreas Muendler, and Stephen Redding. 2013. "Sources of Wage Inequality." *American Economic Review (Papers and Proceedings)*, 103(3): 214–219.

Alesina, Alberto, and Dani Rodrik. 1994. "Distributive Politics and Economic Growth." *The Quarterly Journal of Economics*, 109(2): 465–490.

Alesina, Alberto, and Eliana La Ferrara. 2005. "Preferences for redistribution in the land of opportunities." *Journal of Public Economics*, 89(5-6): 897–931.

Alesina, Alberto, and George-Marios Angeletos. 2005. "Fairness and Redistribution." *American Economic Review*, 95(4): 960–980.

Alvaredo, Facundo, Anthony B. Atkinson, Thomas Piketty, and Emmanuel Saez. 2014. "The World Top Incomes Database."

Amiti, Mary, and David E. Weinstein. 2011. "Exports and Financial Shocks." *The Quarterly Journal of Economics*, 126(4): 1841–1877.

Anderson, Eric T., and Duncan I. Simester. 2003. "Effects of $9 Price Endings on Retail Sales: Evidence from Field Experiments." *Quantitative Marketing and Economics*, 1(1): 93–110.

Andreoni, James. 1989. "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence." *Journal of Political Economy*, 97(6): 1447.

Andreoni, James. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." *The Economic Journal*, 100(401): 464–477.

Angrist, Joshua D., and Alan B. Krueger. 2001. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." *Journal of Economic Perspectives*, 15(4): 69–85.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association*, 91(434): 444–455.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare. 2012. "New Trade Models, Same Old Gains?" *American Economic Review*, 102(1): 94–130.

Ashton, Lydia. 2013. "Left-digit bias and inattention in retail purchases: Evidence from a field experiment."

Atkeson, Andrew, and Ariel Tomas Burstein. 2010. "Innovation, Firm Dynamics, and International Trade." *Journal of Political Economy*, 118(3): 433–484.

Atkinson, Anthony B. 1970. "More on the measurement of inequality." *Journal of Economic Theory*, 2: 244–263.

Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez. 2011. "Top Incomes in the Long Run of History." *Journal of Economic Literature*, 49(1): 3–71.

Atkinson, John W. 1957. "Motivational Determinants of Risk-Taking Behavior." *Psychological Review*, 64(6): 359–72.

Autor, David H. 2009. "Comment." *NBER Macroeconomics Annual*, 24(1): 277–300.

Autor, David H. 2015. "Why Are There Still So Many Jobs? The History and Future of Workplace Automation." *Journal of Economic Perspectives*, 29(3): 3–30.

Autor, David H., Alan Manning, and Christopher L. Smith. 2016. "The Contribution of the Minimum Wage to U.S. Wage Inequality over Three Decades: A Reassessment." *American Economic Journal: Applied Economics*, 8(1): 58–99.

Autor, David H., and David Dorn. 2009. "This Job is 'Getting Old': Measuring Changes in Job Opportunities using Occupational Age Structure." *American Economic Review (Papers and Proceedings)*, 99(2): 45–51.

Autor, David H., and David Dorn. 2013. "The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market." *American Economic Review*, 103(5): 1553–1597.

Autor, David H., David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review*, 103(6): 2121–2168.

Autor, David H., David Dorn, and Gordon H. Hanson. 2015. "Untangling Trade and Technology: Evidence from Local Labor Markets." *The Economic Journal*, 125(584): 621–646.

Autor, David H., Frank Levy, and Richard J. Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *The Quarterly Journal of Economics*, 116(4): 1279–1333.

Autor, David H., Lawrence F. Katz, and Alan B. Krueger. 1998. "Computing Inequality: Have Computers Changed the Labor Market?" *The Quarterly Journal of Economics*, 113(4): 1169–1213.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2005. "Rising Wage Inequality: The Role of Composition and Prices." *NBER Working Paper #11628.*

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2006. "The Polarization of the U.S. Labor Market." *The American Economic Review: Papers and Proceedings*, 96(2): 189–194.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2008. "Trends in U.S. Wage Inequality: Revising the Revisionists." *Review of Economics and Statistics*, 90(2): 300–323.

Aw, Bee Yan, Mark J. Roberts, and Daniel Yi Xu. 2011. "R&D Investment, Exporting, and Productivity Dynamics." *American Economic Review*, 101(4): 1312–1344.

Aw, Bee Yan, Mark J. Roberts, and Tor Winston. 2007. "Export Market Participation, Investments in R&D and Worker Training, and the Evolution of Firm Productivity." *The World Economy*, 30(1): 83–104.

Backus, Matt, Tom Blake, and Steven Tadelis. 2015. "Cheap Talk, Round Numbers, and the Economics of Negotiation." *NBER Working Paper #21285.*

Bakija, Jon, Adam Cole, and T. Heim Bradley. 2012. "Jobs and Income Growth of Top Earners and the Causes of Changing Income Inequality: Evidence from U.S. Tax Return Data."

Banerjee, Abhijit V. 2004. "Contracting constraints, credit markets, and economic development." In *Advances in Economics and Econometrics: Theory and Applications.* , ed. M. Dewatripoint, L. Hansen and S. Turnovsky. Cambridge:Cambridge University Press.

Banerjee, Abhijit V., and Esther Duflo. 2014. "Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program." *Review of Economic Studies*, 81(2): 572–607.

Barber, Brad M., and Terrance Odean. 2008. "All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors." *Review of Financial Studies*, 21(2): 785–818.

Barnea, Amir, Henrik Cronqvist, and Stephan Siegel. 2010. "Nature or nurture: What determines investor behavior?" *Journal of Financial Economics*, 98(3): 583–604.

Barro, Robert J., and Xavier Sala-i Martin. 1992. "Convergence." *Journal of Political Economy*, 100(2): 223–251.

Bartel, Ann P., and Nachum Sicherman. 1999. "Technological Change and Wages: An Interindustry Analysis." *Journal of Political Economy*, 107(2): 285–325.

Basu, Kaushik. 1997. "Why are so many goods priced to end in nine? And why this practice hurts the producers." *Economics Letters*, 54(1): 41–44.

Bauernschuster, Stefan, Timo Hener, and Helmut Rainer. 2016. "Children of a (Policy) Revolution: The Introduction of Universal Child Care and Its Effect on Fertility." *Journal of the European Economic Association*, forthcoming.

Baumeister, Roy F. 1984. "Choking Under Pressure: Self-Consciousness and Paradoxical Effects of Incentives on Skillful Performance." *Journal of Personality and Social Psychology*, 46(3): 610–20.

Baumeister, Roy F., Ellen Bratslavsky, Mark Muraven, and Dianne M. Tice. 1998. "Ego Depletion: Is the Active Self a Limited Resource?" *Journal of Personality and Social Psychology*, 74(5): 1252–1265.

Beaudry, Paul, Mark Doms, and Ethan Lewis. 2010. "Should the Personal Computer Be Considered a Technological Revolution? Evidence from U.S. Metropolitan Areas." *Journal of Political Economy*, 118(5): 988–1036.

Bebchuk, Lucian Arye, and Jesse M. Fried. 2003. "Executive Compensation as an Agency Problem." *Journal of Economic Perspectives*, 17(3): 71–92.

Bechtel, Michael M., Jens Hainmueller, and Yotam Margalit. 2014. "Preferences for International Redistribution: The Divide over the Eurozone Bailouts." *American Journal of Political Science*, 58(4): 835–856.

Becker, Gary S. 1974. "A Theory of Social Interactions." *Journal of Political Economy*, 82(61): 1063.

Becker, Gary S., Edward L. Glaeser, and Kevin M. Murphy. 1999. "Population and Economic Growth." *American Economic Review (Papers and Proceedings)*, 89(2): 145–149.

Beerli, Andreas, and Giovanni Peri. 2015. "The Labor Market Effects of Opening the Border: New Evidence from Switzerland." *NBER Working Paper #21319.*

Beilock, Sian L. 2008. "Math Performance in Stressful Situations." *Current Directions in Psychological Science*, 17(5): 339–343.

Beilock, Sian L., and Thomas H. Carr. 2005. "When High-Powered People Fail Working Memory and "Choking Under Pressure" in Math." *Psychological science*, 16(2): 101–105.

Benabou, Roland, and Efe A. Ok. 2001. "Social Mobility and the Demand for Redistribution: The POUM Hypothesis." *The Quarterly Journal of Economics*, 116(2): 447 – 487.

Benabou, Roland, and Jean Tirole. 2006. "Incentives and Prosocial Behavior." *American Economic Review*, 96(5): 1652–1678.

Benford, Frank. 1938. "The Law of Anomalous Numbers." *Proceedings of the American Philosophical Society*, 78(4): 551–572.

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy*, 75(4): 352–365.

Berman, Eli, John Bound, and Stephen Machin. 1998. "Implications of Skill-Biased Technological Change: International Evidence." *The Quarterly Journal of Economics*, 113(4): 1245–1279.

Berman, Eli, John Bound, and Zvi Griliches. 1994. "Changes in the Demand for Skilled Labor within U.S. Manufacturing: Evidence from the Annual Survey of Manufactures." *The Quarterly Journal of Economics*, 109(2): 367–397.

Bernard, Andrew B., and J. Bradford Jensen. 1997. "Exporters, Skill Upgrading, and the Wage Gap." *Journal of International Economics*, 42: 3–31.

Bernard, Andrew B., and J. Bradford Jensen. 1999. "Exceptional exporter performance: cause, effect, or both?" *Journal of International Economics*, 47(1): 1–25.

Bernard, Andrew B., J. Bradford Jensen, and Peter K. Schott. 2006. "Trade costs, firms and productivity." *Journal of Monetary Economics*, 53(5): 917–937.

Bernard, Andrew B., J. Bradford Jensen, Stephen J. Redding, and Peter K. Schott. 2007. "Firms in International Trade." *Journal of Economic Perspectives*, 21(3): 105–130.

Berry, Steven T. 1994. "Estimating discrete-choice models of product differentiation." *The RAND Journal of Economics*, 25(2): 242–262.

Bhagwati, Jagdish N. 1971. "The Generalized Theory of Distortions and Welfare." In *Trade, balance of payments and growth.* , ed. Jagdish N. Bhagwati, Ronald Jones, Robert Mundell and Jaroslav Vanek, 69–90. Amsterdam:North Holland.

Bick, Alexander. 2016. "The Quantitative Role of Child Care for Female Labor Force Participation and Fertility." *Journal of the European Economic Association*, 14(3): 639–668.

Bils, Mark. 2009. "Do Higher Prices for New Goods Reflect Quality Growth or Inflation?" *Quarterly Journal of Economics*, 124(2): 637–675.

Bivens, Josh, and Lawrence Mishel. 2013. "The Pay of Corporate Excutives and Financial Professionals as Evidence of Rents in Top 1 Percent Incomes." *Journal of Economic Perspectives*, 27(3): 57–78.

Blanchard, Emily, and Gerald Willmann. 2013. "Trade, Education, and The Shrinking Middle Class."

Blau, Francine D., and Lawrence M. Kahn. 2007. "Changes in the Labor Supply Behavior of Married Women: 1980-2000." *Journal of Labor Economics*, 25(3): 393–438.

Blau, Francine D., and Lawrence M. Kahn. 2015. "Immigration and the Distribution of Incomes." In *Handbook of the Economics of International Migration.* , ed. Barry R. Chiswick and Paul W. Miller. Amsterdam:Elsevier.

Blinder, Alan S. 2009. "How Many US Jobs Might be Offshorable?" *World Economics*, 10(2): 41–78.

Blinder, Alan S., and Alan B. Krueger. 2013. "Alternative Measures of Offshorability: A Survey Approach." *Journal of Labor Economics*, 31(2): S97–S128.

Bloom, David E., David Canning, and Guenther Fink. 2010. "Implications of population ageing for economic growth." *Oxford Review of Economic Policy*, 26(4): 583–612.

Boehm, Michael J. 2015. "The Price of Polarization: Estimating Task Prices under Routine-Biased Technical Change."

Bongaarts, John. 1999. "Fertility Decline in the Developed World: Where Will It End?" *American Economic Review (Papers and Proceedings)*, 89(2): 256–260.

Booth, Alison L., and Patrick Nolen. 2012. "Gender Differences in Risk Behaviour: Does Nurture Matter?" *Economic Journal*, 122(1): F56–F78.

Borjas, George J. 1987. "Self-Selection and the Earnings of Immigrants." *The American Economic Review*, 77(4): 531–553.

Borjas, George J., and Valerie A. Ramey. 1994. "Time-Series Evidence on the Sources of Trends in Wage Inequality." *The American Economic Review (Papers and Proceedings)*, 84(2): 10–16.

Borjas, George J., Richard B. Freeman, and Lawrence F. Katz. 1997. "How Much Do Immigration and Trade Affect Labor Market Outcomes?" *Brookings Papers on Economic Activity*, , (1): 1–90.

Bosch, Mariano, and Marco Manacorda. 2010. "Minimum Wages and Earnings Inequality in Urban Mexico." *American Economic Journal: Applied Economics*, 2(4): 128–149.

Bound, John, and George Johnson. 1992. "Changes in the Structure of Wages in the 1980's: An Evaluation of Alternative Explanations." *The American Economic Review*, 82(3): 371–392.

Bound, John, and Harry J. Holzer. 2000. "Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s." *Journal of Labor Economics*, 18(1): 20–54.

Bratsberg, Bernt, Oddbjørn Raaum, and Knut Røed. 2014. "Immigrants, Labour Market Performance and Social Insurance." *The Economic Journal*, 124(580): F644–F683.

Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt. 2002. "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence." *The Quarterly Journal of Economics*, 117(1): 339–376.

Broda, Christian, and John Romalis. 2009. "The Welfare Implications of Rising Price Dispersion."

Brown, Jennifer. 2011. "Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars." *Journal of Political Economy*, 119(5): 982–1013.

Brown, Jennifer, Tanjim Hossain, and John Morgan. 2010. "Shrouded Attributes and Information Suppression: Evidence from the Field." *The Quarterly Journal of Economics*, 125(2): 859–876.

Buera, Francisco J., and Joseph P. Kaboski. 2012. "The Rise of the Service Economy." *American Economic Review*, 102(6): 2540–2569.

Burkhauser, Richard V., Jeff Larrimore, and Kosali I. Simon. 2012. "A Second Opinion on the Economic Health of the American Middle Class." *National Tax Journal*, 65(1): 7–32.

Bursztyn, Leonardo, and Davide Cantoni. 2016. "A Tear in the Iron Curtain: The Impact of Western Television on Consumption Behavior." *Review of Economics and Statistics*, 98(1): 25–41.

Burtless, Gary, and Lynn A. Karoly. 1995. "Demographic Change, Rising Earnings of Personal Inequality, and the Distribution of Personal Well-Being, 1959-1989." *Demography*, 32(3): 379–405.

Busse, Meghan R., Nicola Lacetera, Devin G. Pope, Jorge Silva-Risso, and Justin R. Sydnor. 2013. "Estimating the Effect of Salience in Wholesale and Retail Car Markets." *American Economic Review (Papers and Proceedings)*, 103(3): 575–579.

Bustos, Paula. 2011*a*. "The Impact of Trade Liberalization on Skill Upgrading: Evidence from Argentina."

Bustos, Paula. 2011*b*. "Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms." *American Economic Review*, 101(1): 304–340.

Butcher, Kristin F., and David Card. 1991. "Immigration and Wages: Evidence from the 1980's." *The American Economic Review*, 81(2): 292–296.

Caliendo, Lorenzo, and Esteban Rossi-Hansberg. 2012. "The Impact of Trade on Organization and Productivity." *The Quarterly Journal of Economics*, 127(3): 1393–1467.

Camerer, Colin F. 1998. "Can Asset Markets Be Manipulated? A Field Experiment with Racetrack Betting." *Journal of Political Economy*, 106(3): 457.

Card, David. 1996. "The Effect of Unions on the Structure of Wages: A Longitudinal Analysis." *Econometrica*, 64(4): 957–979.

Card, David. 2009. "Immigration and Inequality." *American Economic Review (Papers and Proceedings)*, 99(2): 1–21.

Card, David. 2012. "Comment: The Elusive Search for Negative Wage Impacts of Immigration." *Journal of the European Economic Association*, 10(1): 211–215.

Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez. 2012. "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction." *American Economic Review*, 102(6): 2981–3003.

Card, David, and John E. DiNardo. 2002. "Skill Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles." *Journal of Labor Economics*, 20(4): 733–783.

Card, David, and Thomas Lemieux. 2001. "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis." *The Quarterly Journal of Economics*, 116(2): 705–746.

Card, David, Christian Dustmann, and Ian Preston. 2012. "Immigration, Wages, and Compositional Amenities." *Journal of the European Economic Association*, 10(1): 78–119.

Carneiro, Pedro, and Sokbae Lee. 2011. "Trends in Quality Adjusted Skill Premia in the US, 1960-2000." *American Economic Review*, 101(6): 2309–2349.

Caselli, Mauro. 2012. "Does wealth inequality reduce the gains from trade?" *Review of World Economics*, 148(2): 333–356.

Caselli, Mauro. 2013. "Credit constraints, inequality and the growth gains from trade." *Economics Letters*, 121(1): 43–47.

Cesarini, David, Christopher T. Dawes, Magnus Johannesson, Paul Lichtenstein, and Björn Wallace. 2009. "Genetic Variation in Preferences for Giving and Risk Taking." *Quarterly Journal of Economics*, 124(2): 809–842.

Chamberlain, Gary. 1980. "Analysis of Covariance Data with Qualitative." *The Review of Economic Studies*, 47(1): 225–238.

Champagne, Julien, and André Kurmann. 2013. "The Great Increase in Relative Wage Volatility in the United States." *Journal of Monetary Economics*, 60(2): 166–183.

Charles, Kerwin Kofi, Erik Hurst, and Alexandra Killewald. 2013. "Marital Sorting and Parental Wealth." *Demography*, 50(1): 51–70.

Chay, Kenneth Y., and David S. Lee. 2000. "Changes in relative wages in the 1980s: Returns to observed and unobserved skills and black-white wage differentials." *Journal of Econometrics*, 99(1): 1–38.

Checchi, Daniele, Jelle Visser, and Herman G. van de Werfhorst. 2010. "Inequality and Union Membership: The Influence of Relative Earnings and Inequality Attitudes." *British Journal of Industrial Relations*, 48(1): 84–108.

Chesnokova, Tatyana. 2007. "Immiserizing deindustrialization: A dynamic trade model with credit constraints." *Journal of International Economics*, 73(2): 407–420.

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. "Salience and Taxation: Theory and Evidence." *American Economic Review*, 99(4): 1145–1177.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner. 2014. "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility." *American Economic Review (Papers and Proceedings)*, 104(5): 141–147.

Chiquiar, Daniel, and Gordon H. Hanson. 2005. "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States." *Journal of Political Economy*, 113(2): 239–281.

Clark, Gregory. 2014. *The Son Also Rises: Surnames and the History of Social Mobility.* Princeton:Princeton University Press.

Clark, Gregory, and Neil Cummins. 2015. "Intergenerational Wealth Mobility in England, 1858-2012. Surnames and Social Mobility." *The Economic Journal*, 125: 61–85.

Coibion, Olivier, Yuriy Gorodnichenko, Lorenz Kueng, and John Silvia. 2012. "Innocent Bystanders? Monetary Policy and Inequality in the U.S." *NBER Working Paper #18170.*

Conlin, Michael, Ted O'Donoghue, and Timothy J. Vogelsang. 2007. "Projection Bias in Catalog Orders." *American Economic Review*, 97(4): 1217–1249.

Connolly, Robert A., and Richard J. Rendleman. 2008. "Skill, Luck, and Streaky Play on the PGA Tour." *Journal of the American Statistical Association*, 103(481): 74–88.

Corak, Miles. 2013. "Income Inequality, Equality of Opportunity, and Intergenerational Mobility." *Journal of Economic Perspectives*, 27(3): 79–102.

Corneo, Giacomo, and Hans Peter Grüner. 2002. "Individual preferences for political redistribution." *Journal of Public Economics*, 83(1): 83–107.

Cortes, Patricia. 2008. "The Effect of Low Skilled Immigration on U.S. Prices: Evidence from CPI Data." *Journal of Political Economy*, 116(3): 381–422.

Costantini, James A., and Marc J. Melitz. 2008. "The Dynamics of Firm-Level Adjustment to Trade Liberalization." In *The Organization of Firms in a Global Economy.* , ed. Elhanan Helpman, Dalia Marin and Thierry Verdier, 107–141. Cambridge, MA:Harvard University Press.

Cozzi, Guido, and Giammario Impullitti. 2010. "Government Spending Composition, Technical Change, and Wage Inequality." *Journal of the European Economic Association*, 8(6): 1325–1358.

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

Cusack, Thomas, Torben Iversen, and Philipp Rehm. 2006. "Risks at Work: The Demand and Supply Sides of Government Redistribution." *Oxford Review of Economic Policy*, 22(3): 365–389.

Cutler, David M., James M. Poterba, Louise M. Sheiner, and Lawrence H. Summers. 1990. "An Aging Society: Opportunity or Challenge?" *Brookings Papers on Economic Activity*, 1: 1–56.

Dahlberg, Matz, Karin Edmark, and Heléne Lundqvist. 2012. "Ethnic Diversity and Preferences for Redistribution." *Journal of Political Economy*, 120(1): 41–76.

Dalton, Hugh. 1920. "The Measurement of the Inequality of Incomes." *The Economic Journal*, 30(119): 348–361.

Dancygier, Rafaela, and Michael Donnelley. 2014. "Mass Politics in Tough Times: Opinions, Votes and Protest in the Great Recession." In *Mass Politics in Tough Times.* , ed. Larry Bartels and Nancy Bermeo. Oxford:Oxford University Press.

Davies, James B., and Michael Hoy. 2002. "Flat rate taxes and inequality measurement." *Journal of Public Economics*, 84(1): 33–46.

Deaton, Angus S., and Christina H. Paxson. 1994. "Intertemporal Choice and Inequality." *Journal of Political Economy*, 102(3): 437–467.

Deaton, Angus S., and Christina H. Paxson. 1995. "Saving, Inequality and Aging: An East Asian Perspective." *Asia-Pacific Economic Review*, 1: 7–19.

Deaton, Angus S., and Christina H. Paxson. 1997. "The Effects of Economic and Population Growth on National Saving and Inequality." *Demography*, 34(1): 97–114.

Deaton, Angus S., and Christina H. Paxson. 1998. "Aging and Inequality in Income and Health." *The American Economic Review: Papers and Proceedings*, 88(2): 248–253.

Dehejia, Rajeev, and Arvind Panagariya. 2014. "Trade Liberalization in Manufacturing and Accelerated Growth in Services in India." *NBER Working Paper #19923.*

De La Croix, David, and Omar Licandro. 2013. "The Child is Father Of the Man: Implications for the Demographic Transition." *The Economic Journal*, 123(567): 236–261.

DellaVigna, Stefano. 2009. "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature*, 47(2): 315–372.

DellaVigna, Stefano, and Eliana La Ferrara. 2015. "Economic and Social Impacts of the Media." *CEPR Discussion Paper No. 10667.*

DellaVigna, Stefano, and Ethan Kaplan. 2007. "The Fox News Effect: Media Bias and Voting." *The Quarterly Journal of Economics*, 3(122): 1187–1234.

DellaVigna, Stefano, and Joshua M. Pollet. 2007. "Demographics and Industry Returns." *American Economic Review*, 97(5): 1667–1702.

DellaVigna, Stefano, John A. List, and Ulrike Malmendier. 2012. "Testing for Altruism and Social Pressure in Charitable Giving." *Quarterly Journal of Economics*, 127(1): 1–56.

Derrien, François, and Olivier Dessaint. 2015. "The Effects of Investment Bank Rankings: Evidence from M&A League Tables."

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica*, 64(5): 1001–1044.

Docquier, Frédéric, Caglar Ozden, and Giovanni Peri. 2013. "The Labour Market Effects of Immigration and Emigration in OECD Countries." *The Economic Journal*, 124(579): 1106–1145.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. 2012. "The Intergenerational Transmission of Risk and Trust Attitudes." *The Review of Economic Studies*, 79(2): 645–677.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner. 2011. "Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences." *Journal of the European Economic Association*, 9(3): 522–550.

Dohmen, Thomas J. 2008. "Do professionals choke under pressure?" *Journal of Economic Behavior & Organization*, 65(3-4): 636–653.

Doms, Mark, Timothy Dunne, and Kenneth R. Troske. 1997. "Workers, Wages, and Technology." *The Quarterly Journal of Economics*, 112(1): 253–290.

Dooley, Martin D., and Peter Gottschalk. 1984. "Earnings Inequality among Males in the United States: Trends and the Effect of Labor Force Growth." *The Journal of Political Economy*, 92(1): 59–89.

Dotti, Valerio. 2016. "The Political Economy of Immigration and Population Ageing."

Dow, Gregory K., and Clyde G. Reed. 2013. "The Origins of Inequality: Insiders, Outsiders, Elites, and Commoners." *Journal of Political Economy*, 121(3): 609–641.

Dustmann, Christian, Johannes Ludsteck, and Uta Schoenberg. 2009. "Revisiting the German Wage Structure." *The Quarterly Journal of Economics*, 124(2): 809–842.

Dustmann, Christian, Tommaso Frattini, and Ian P. Preston. 2013. "The Effect of Immigration along the Distribution of Wages." *The Review of Economic Studies*, 80(1): 145–173.

Eaton, Jonathan, and Samuel Kortum. 2002. "Technology, Geography, and Trade." *Econometrica*, 70(5): 1741–1779.

Eaton, Jonathan, and Samuel Kortum. 2012. "Putting Ricardo to Work." *The Journal of Economic Perspectives*, 26(2): 65–90.

Ebenstein, Avraham, Ann Harrison, Margaret McMillan, and Shannon Philipps. 2014. "Estimating the Impact of Trade and Offshoring on American Workers Using the Current Population Surveys." *Review of Economics and Statistics*, 96(3): 581–595.

Egger, Hartmut, and Daniel Etzel. 2012. "The impact of trade on employment, welfare, and income distribution in unionized general oligopolistic equilibrium." *European Economic Review*, 56(6): 1119–1135.

Egger, Hartmut, and Udo Kreickemeier. 2012. "Fairness, Trade, and Inequality." *Journal of International Economics*, 86(2): 184–196.

Ehrenberg, Ronald G., and Michael L. Bognanno. 1990. "Do Tournaments Have Incentive Effects?" *The Journal of Political Economy*, 98(6): 1307–1324.

Eika, Lasse, Magne Mogstad, and Basit Zafar. 2014. "Educational Assortative Mating and Household Income Inequality." *NBER Working Paper #20271.*

Eisensee, Thomas, and David Strömberg. 2007. "News Droughts, News Floods, and U.S. Disaster Relief." *Quarterly Journal of Economics*, 122(2): 693–728.

Englmaier, Florian, Arno Schmoeller, and Till Stowasser. 2013. "Price Discontinuities in an Online Used Car Market."

Epifani, Paolo, and Gino Gancia. 2011. "Trade, markup heterogeneity and misallocations." *Journal of International Economics*, 83(1): 1–13.

Epstein, Robert, and Ronald E. Robertson. 2015. "The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections." *Proceedings of the National Academy of Sciences*, 112(33): E4512–E4521.

Ermisch, John, Marco Francesconi, and Thomas Siedler. 2006. "Intergenerational Mobility and Marital Sorting." *The Economic Journal*, 116(1999): 659–679.

Eugster, Beatrix, Rafael Lalive, Andreas Steinhauer, and Josef Zweimüller. 2011. "The Demand for Social Insurance: Does Culture Matter?" *The Economic Journal*, 121(556): F413–F448.

Facchini, Giovanni, and Anna Maria Mayda. 2008. "From Individual Attitudes towards Migrants to Migration Policy Outcomes: Theory and Evidence." *Economic Policy*, 23: 651–713.

Facchini, Giovanni, and Anna Maria Mayda. 2009. "Does the Welfare State Affect Individual Attitudes toward Immigrants? Evidence across Countries." *Review of Economics and Statistics*, 91(2): 295–314.

Falkinger, Josef. 2007. "Attention economies." *Journal of Economic Theory*, 133(1): 266–294.

Falkinger, Josef. 2008. "Limited Attention as a Scarce Resource in Information-Rich Economies." *The Economic Journal*, 118(532): 1596–1620.

Fan, Haichao, Yao Amber Li, and Stephen R. Yeaple. 2015. "Trade Liberalization, Quality, and Export Prices." *Review of Economics and Statistics*, 97(5): 1033–1051.

Farber, Henry S., and Robert Gibbons. 1996. "Learning and Wage Dynamics." *The Quarterly Journal of Economics*, 111(4): 1007–1047.

Feenberg, Daniel R, Ina Ganguli, Patrick Gaule, and Jonathan Gruber. 2015. "It's Good to be First: Order Bias in Reading and Citing NBER Working Papers." *NBER Working Paper #21141.*

Feenstra, Robert C., and Gordon H. Hanson. 1996. "Globalization, Outsourcing, and Wage Inequality." *The American Economic Review (Papers and Proceedings)*, 86(2): 240–245.

Feenstra, Robert C., and Gordon H. Hanson. 2003. "Global Production Sharing and Rising Inequality: A Survey of Trade and Wages." In *Handbook of International Trade.* , ed. E. Kwan Choi and James Harrigan, 146–185. Oxford:Blackwell Publishing.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review*, 105(10): 3150–3182.

Feng, Lei, and Mark S. Seasholes. 2005. “Do Investor Sophistication and Trading Experience Eliminate Behavioral Biases in Financial Markets?” *Review of Finance*, 9(3): 305–351.

Feng, Shuaizhang, Richard V. Burkhauser, and J. S. Butler. 2006. “Levels and Long-Term Trends in Earnings Inequality.” *Journal of Business & Economic Statistics*, 24(1): 57–62.

Fernandes, Ana M., and Caroline Paunov. 2013. “Does trade stimulate product quality upgrading?” *Canadian Journal of Economics*, 46(4): 1232–1264.

Fernandes, Ana M., and Caroline Paunov. 2015. “The Risks of Innovation: Are Innovating Firms Less Likely to Die?” *Review of Economics and Statistics*, 97(3): 638–653.

Fernández, Raquel. 2002. “Education, Segregation and Marital Sorting: Theory and an Application to the UK.” *European Economic Review*, 46: 993–1022.

Fernández, Raquel, and Richard Rogerson. 2001. “Sorting and Long-Run Inequality.” *The Quarterly Journal of Economics*, 116(4): 1305–1341.

Fieler, Ana Cecilia, Marcela Eslava, and Daniel Xu. 2014. “Trade, Skills, and Quality Upgrading: A Theory with Evidence from Colombia.” *NBER Working Paper #19992.*

Finkelstein, Amy. 2009. “Z-Tax: Tax Salience and Tax Rates.” *The Quarterly Journal of Economics*, 124(3): 969–1010.

FIS - International Ski Federation. 2012. *Audi FIS Alpine Skiing World Cup Prize Money Ranking Season 2012/2013.* Oberhofen, Switzerland.

Foged, Mette, and Giovanni Peri. 2016. “Immigrants’ Effect on Native Workers: New Analysis on Longitudional Data.” *American Economic Journal: Applied Economics*, 8(2): 1–34.

Foley, C. Fritz, and Kalina Manova. 2015. “International Trade, Multinational Activity, and Corporate Finance.” *Annual Review of Economics*, 7: 119–146.

Föllmi, Reto, and Isabel Martínez. 2016. “Volatile Top Income Shares in Switzerland? Reassessing the Evolution Between 1981 and 2010.” *Review of Economics and Statistics*, forthcoming.

Föllmi, Reto, and Josef Zweimüller. 2006. “Income Distribution and Demand-Induced Innovations.” *Review of Economic Studies*, 73(4): 941–960.

Föllmi, Reto, and Manuel Oechslin. 2010. “Market Imperfections, Wealth Inequality, and the Distribution of Trade Gains.” *Journal of International Economics*, 81(1): 15–25.

Föllmi, Reto, and Manuel Oechslin. 2014. “Globalization and Productivity in the Developing World.”

Föllmi, Reto, Stefan Legge, and Lukas Schmid. 2016. “Do Professionals Get It Right? Limited Attention and Risk-Taking Behaviour.” *The Economic Journal*, 126(592): 724–755.

Föllmi, Reto, Tobias Würgler, and Josef Zweimüller. 2014. “The Macroeconomics of Model T.” *Journal of Economic Theory*, 153: 617–647.

Fong, Christina. 2001. “Social preferences, self-interest, and the demand for redistribution.” *Journal of Public Economics*, 82(2): 225–246.

Formby, John P., Terry G. Seaks, and W. James Smith. 1989. "On the Measurement and Trend of Inequality: A Reconsideration." *The American Economic Review*, 79(1): 256–264.

Frangakis, Constantine E., and Donald B. Rubin. 2002. "Principal Stratification in Causal Inference." *Biometrics*, 58(1): 21–29.

Freeman, Richard B. 1975. "Overinvestment in College Training?" *Journal of Human Resources*, 10(3): 287–311.

Freeman, Richard B. 1976. *The Overeducated American.* New York:Academic Press.

Freeman, Richard B. 1979. "The Effect of Demographic Factors on Age-Earnings Profiles." *The Journal of Human Resources*, 14(3): 289–318.

Friedman, Milton, and L. J. Savage. 1948. "The Utility Analysis of Choices Involving Risk." *Journal of Political Economy*, 56(4): 279–304.

Friesen, Jane, and Brian Krauth. 2007. "Sorting and Inequality in Canadian Schools." *Journal of Public Economics*, 91(11-12): 2185–2212.

Frölich, Markus. 2007. "Regression Discontinuity Design with Covariates." *IZA Discussion Paper*, No.3024.

Frydman, Carola, and Raven E. Saks. 2010. "Executive compensation: A new view from a long-term perspective, 1936-2005." *Review of Financial Studies*, 23(5): 2099–2138.

Gabaix, Xavi, and David Laibson. 2006. "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *The Quarterly Journal of Economics*, 121(2): 505–540.

Gabaix, Xavier, and Augustin Landier. 2008. "Why has CEO pay increased so much?" *The Quarterly Journal of Economics*, 123(1): 49–100.

Galor, Oded, and Andrew Mountford. 2008. "Trading Population for Productivity: Theory and Evidence." *The Review of Economic Studies*, 75: 1143–1179.

Galor, Oded, and Daniel Tsiddon. 1997. "Technological Progress, Mobility, and Economic Growth." *The American Economic Review*, 87(3): 363–382.

Galor, Oded, and Joseph Zeira. 1993. "Income Distribution and Macroeconomics." *The Review of Economic Studies*, 60(1): 35–52.

Galor, Oded, and Omer Moav. 2000. "Ability-biased technological transition, wage inequality, and economic growth." *The Quarterly Journal of Economics*, 112(2): 469–497.

Gardner, Margo, and Laurence Steinberg. 2005. "Peer Influence on Risk Taking, Risk Preference, and Risky Decision Making in Adolescence and Adulthood: An Experimental Study." *Developmental Psychology*, 41(4): 625–35.

Garicano, Luis, and Esteban Rossi-Hansberg. 2015. "Knowledge-based Hierarchies: Using Organizations to Understand the Economy." *The Annual Review of Economics*, 7: 1–30.

Gelman, Andrew, and Guido Imbens. 2014. "Why High-Order Polynomials Should Not be Used in Regression Discontinuity Designs." *NBER Working Paper #20405.*

Genakos, Christos, and Mario Pagliero. 2012. "Interim Rank, Risk Taking, and Performance in Dynamic Tournaments." *Journal of Political Economy*, 120(4): 782–813.

Gentzkow, Matthew, and Jesse M. Shapiro. 2010. "What Drives Media Slant? Evidence From U.S. Daily Newspapers." *Econometrica*, 78(1): 35–71.

Gill, David, and Victoria Prowse. 2012. "A Structural Analysis of Disappointment Aversion in a Real Effort Competition." *American Economic Review*, 102(1): 469–503.

Gill, David, Zdenka Kissova, Jaesun Lee, and Victoria Prowse. 2016. "First-place loving and last-place loathing: How rank in the distribution of performance affects effort provision." *University of Oxford Discussion Paper 783*.

Gilovich, Thomas, Robert Vallone, and Amos Tversky. 1985. "The Hot Hand in Basketball: On the Misperception of Random Sequences." *Cognitive Psychology*, 17(3): 295–314.

Goldberg, Pinelopi K., and Nina Pavcnik. 2007. "Distributional Effects of Globalization in Developing Countries." *Journal of Economic Literature*, 45(1): 39–82.

Goldin, Claudia, and Lawrence F. Katz. 2008. *The Race between Education and Technology.* Cambridge, Mass.:Harvard University Press.

Goos, Maarten, Alan Manning, and Anna Salomons. 2009. "Job Polarization in Europe." *American Economic Review (Papers and Proceedings)*, 99(2): 58–63.

Goos, Maarten, Alan Manning, and Anna Salomons. 2014. "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring." *American Economic Review*, 104(8): 2509–2526.

Goos, Maarten, and Alan Manning. 2003. "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain."

Goos, Maarten, and Alan Manning. 2007. "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain." *The Review of Economics and Statistics*, 89(1): 118–133.

Gorodnichenko, Yuriy, and Monika Schnitzer. 2013. "Financial Constraints and Innovation: Why Poor Countries Don't Catch Up." *Journal of the European Economic Association*, 11(5): 1115–1152.

Greenaway, David, Wyn Morgan, and Peter Wright. 2002. "Trade liberalisation and growth in developing countries." *Journal of Development Economics*, 67(1): 229–244.

Green, Brett S., and Jeffrey Zwiebel. 2015. "The Hot-Hand Fallacy: Cognitive Mistakes or Equilibrium Adjustments? Evidence from Major League Baseball."

Greenwood, Jeremy, Nezih Guner, Georgi Kocharkov, and Cezar Santos. 2014. "Marry Your Like: Assortative Mating and Income Inequality." *American Economic Review (Papers and Proceedings)*, 104(5): 348–353.

Grossman, Gene M., and Elhanan Helpman. 2014. "Growth, Trade, and Inequality." *CEPR Discussion Paper No. 10235*.

Grossman, Gene M., and Esteban Rossi-Hansberg. 2008. "Trading Tasks: A Simple Theory of Offshoring." *American Economic Review*, 95(5): 1978–1997.

Guillaud, Elvire. 2013. "Preferences for redistribution: An empirical analysis over 33 countries." *Journal of Economic Inequality*, 11(1): 57–78.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2006. "Does Culture Affect Economic Outcomes?" *Journal of Economic Perspectives*, 20(2): 23–48.

Guvenen, Fatih, and Burhanettin Kuruscu. 2009. "A Quantitative Analysis of the Evolution of the U.S. Wage Distribution, 1970-2000." *NBER Macroeconomics Annual*, 24(1): 227–276.

Guvenen, Fatih, and Burhanettin Kuruscu. 2012. "Understanding the Evolution of the US Wage Distribution: A Theoretical Analysis." *Journal of the European Economic Association*, 10(3): 482–517.

Guvenen, Fatih, Burhanettin Kuruscu, and Serdar Ozkan. 2014. "Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis." *Review of Economic Studies*, 81: 818–850.

Hahn, Jinyong, Petra Todd, and Wilbert van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica*, 69(1): 201–209.

Hainmueller, Jens, and Michael J. Hiscox. 2007. "Educated Preferences: Explaining Attitudes Toward Immigration in Europe." *International Organization*, 61(02): 399–442.

Hainmueller, Jens, and Michael J. Hiscox. 2010. "Attitudes toward Highly Skilled and Low-skilled Immigration: Evidence from a Survey Experiment." *American Political Science Review*, 104(01): 61–84.

Hainmueller, Jens, Michael J. Hiscox, and Yotam Margalit. 2015. "Do Concerns About Labor Market Competition Shape Attitudes Toward Immigration? New Evidence." *Journal of International Economics*, 97(1): 193–207.

Halla, Martin, Alexander F. Wagner, and Josef Zweimüller. 2015. "Immigration and Voting for the Far Right."

Halter, Daniel, Manuel Oechslin, and Josef Zweimüller. 2014. "Inequality and growth: The neglected time dimension." *Journal of Economic Growth*, 19(1): 81–104.

Hamermesh, Daniel S. 1999. "Changing Inequality in Markets for Workplace Amenities." *The Quarterly Journal of Economics*, 114(4): 1085–1123.

Hamlen, William A. 1991. "Superstardom in Popular Music: Empirical Evidence." *The Review of Economics and Statistics*, 73(4): 729–733.

Hanson, Robin, Ryan Oprea, and David Porter. 2006. "Information aggregation and manipulation in an experimental market." *Journal of Economic Behavior and Organization*, 60(4): 449–459.

Harbaugh, William T. 1998. "What do donations buy? A model of philanthropy based on prestige and warm glow." *Journal of Public Economics*, 67(2): 269–284.

Harrison, Ann, John McLaren, and Margaret McMillan. 2011. "Recent Perspectives on Trade and Inequality." *Annual Review of Economics*, 3(1): 261–289.

Hart, Sergiu. 2005. "Adaptive Heuristics." *Econometrica*, 73(5): 1401–1430.

Haskel, Jonathan, Robert Z. Lawrence, Edward E. Leamer, and Matthew J. Slaughter. 2012. "Globalization and US Wages: Modifying Classic Theory to Explain Recent Facts." *The Journal of Economic Perspectives*, 26(2): 119–140.

Hatton, Timothy J. 2014. "The economics of international migration: A short history of the debate." *Labour Economics*, 30: 43–50.

Heckscher, Eli Filip. 1919. "The effect of foreign trade on the distribution of income." *Ekonomisk Tidskrif*, 21(2): 1–32.

Helpman, Elhanan, Oleg Itskhoki, and Stephen Redding. 2010. "Unequal Effects of Trade on Workers with Different Abilities." *Journal of the European Economic Association*, 8(2-3): 421–433.

Helpman, Elhanan, Oleg Itskhoki, Marc-Andreas Muendler, and Stephen Redding. 2015. "Trade and Inequality: From Theory to Estimation."

Higgins, Matthew, and Jeffrey G. Williamson. 2002. "Explaining Inequality the World Round: Cohort Size, Kuznets Curves, and Openness." *Southeast Asian Studies*, 40(3): 268–302.

Hirschl, Thomas A., and Mark R. Rank. 2015. "The Life Course Dynamics of Affluence." *Plos One*, 10(1).

Hobbs, Frank, and Nicole Stoops. 2002. *Demographic Trends in the 20th Century.* Washington DC:U.S. Government Printing Office.

Hoelzl, Erik, and Aldo Rustichini. 2005. "Overconfident: Do You Put your Money on it?" *The Economic Journal*, 115(503): 305–318.

Hoffman, Mitchell. 2011. "Does Higher Income Make You More Altruistic? Evidence from the Holocaust." *Review of Economics and Statistics*, 93(3): 876–887.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *The Quarterly Journal of Economics*, 124(4): 1403–1448.

Jaimovich, Nir, and Henry E. Siu. 2009. "The Young, the Old, and the Restless: Demographics and Business Cycle Volatility." *American Economic Review*, 99(3): 804–826.

Jensen, Robert, and Emily Oster. 2009. "The Power of TV: Cable Television and Women's Status in India." *The Quarterly Journal of Economics*, 124(3): 1057–1094.

Jerome, Harry. 1934. *Mechanization in History.* New York:NBER.

Jones, Larry E., and Michele Tertilt. 2008. "An Economic History of Fertility in the U.S.: 1826-1960." *NBER Working Paper #12796.*

Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce. 1993. "Wage Inequality and the Rise in Returns to Skill." *The Journal of Political Economy*, 101(3): 410–442.

Kagan, Robert. 2003. *Of Paradise and Power - America and Europe in the New World Order.* New York:Random House.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow.* New York:Farrar, Straus and Giroux.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 47(2): 263–292.

Kahn, Lawrence M. 2000. "The Sports Business as a Labor Market Laboratory." *Journal of Economic Perspectives*, 14(3): 75–94.

Kaldor, Nicholas. 1955. "Alternative Theories of Distribution." *The Review of Economic Studies*, 23(2): 83–100.

Kaplan, Steven N., and Joshua Rauh. 2013. "It's the Market: The Broad-Based Rise in the Return to Top Talent." *Journal of Economic Perspectives*, 27(3): 35–56.

Katz, Lawrence F., and David H. Autor. 1999. "Changes in the Wage Structure and Earnings Inequality." In *Handbook of Labor Economics*. . 3A ed., , ed. Orley Ashenfelter and David Card, Chapter 6, 1463–1555. Amsterdam:Elsevier.

Katz, Lawrence F., and Kevin M. Murphy. 1992. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *The Quarterly Journal of Economics*, 107(1): 35–78.

Klaassen, Franc J.G.M., and Jan R. Magnus. 2009. "The efficiency of top agents: An analysis through service strategy in tennis." *Journal of Econometrics*, 148(1): 72–85.

Kleibergen, Frank, and Richard Paap. 2006. "Generalized reduced rank tests using the singular value decomposition." *Journal of Econometrics*, 133(1): 97–126.

Knack, Stephen, and Philip Keefer. 1997. "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation." *The Quarterly Journal of Economics*, 112(4): 1251–1288.

Kopczuk, Wojciech, Emmanuel Saez, and Jae Song. 2010. "Earnings Inequality and Mobility in the United States: Evidence from Social Security Data since 1937." *The Quarterly Journal of Economics*, 125(1): 91–128.

Korvorst, Marjolein, and Markus F. Damian. 2008. "The differential influence of decades and units on multidigit number comparison." *The Quarterly Journal of Experimental Psychology*, 61(8): 1250–1264.

Köszegi, Botond, and Adam Szeidl. 2013. "A Model of Focusing in Economic Choice." *Quarterly Journal of Economics*, 128(1): 53–104.

Köszegi, Botond, and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics*, 121(4): 1133–1165.

Kremer, Michael. 1997. "How Much Does Sorting Increase Inequality?" *The Quarterly Journal of Economics*, 112(1): 115–139.

Krueger, Alan B. 1993. "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989." *The Quarterly Journal of Economics*, 108(1): 33–60.

Krueger, Dirk, and Alexander Ludwig. 2007. "On the consequences of demographic change for rates of returns to capital, and the distribution of wealth and welfare." *Journal of Monetary Economics*, 54(1): 49–87.

Kuhnen, Camelia M., and Agneieszka Tymula. 2012. "Feedback, Self-Esteem, and Performance in Organizations." *Management Science*, 58(1): 94–113.

Kuhnen, Camelia M., and Jeffrey H. Zwiebel. 2009. "Executive Pay, Hidden Compensation and Managerial Entrenchment."

Kuziemko, Iliyana, Michael I. Norton, Emmanuel Saez, and Stefanie Stantcheva. 2015. "How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments." *American Economic Review*, 105(4): 1478–1508.

Kuznets, Simon. 1955. "Economic Growth and Income Inequality." *The American Economic Review*, 45(1): 1–28.

Lacetera, Nicola, Devin G. Pope, and Justin R. Sydnor. 2012. "Heuristic Thinking and Limited Attention in the Car Market." *American Economic Review*, 102(5): 2206–2236.

La Ferrara, Eliana, Alberto Chong, and Suzanne Duryea. 2012. "Soap Operas and Fertility: Evidence from Brazil." *American Economic Journal: Applied Economics*, 4(4): 1–31.

Lam, David. 1997. "Demographic Variables and Income Inequality." In *Handbook of Population and Family Economics.* 1015–1059.

Last, Jonathan V. 2013. *What to Expect When No One's Expecting: America's Coming Demographic Disaster.* New York:Encounter Books.

Lee, Choong Ki, Ki Joon Back, David C. Hodgins, and Tae Kyung Lee. 2013. "Examining Antecedents and Consequences of Gambling Passion: The Case of Gambling on Horse Races." *Psychiatry Investigation*, 10(4): 365–372.

Lee, David S. 1999. "Wage Inequality in the U.S. during the 1980s: Rising Dispersion or Falling Minimum Wage?" *The Quarterly Journal of Economics*, 114(3): 977–1023.

Lee, David S., and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48(June): 281–355.

Lee, Ronald, and Andrew Mason. 2014. "Is Low Fertility Really a Problem? Population Aging, Dependency, and Consumption." *Science*, 346(6206): 229–234.

Legge, Stefan, and Lukas Schmid. 2016. "Media Attention and Betting Markets." *European Economic Review*, 87: 304–333.

Lemieux, Thomas. 2006. "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *The American Economic Review*, 96(3): 461–498.

Lemieux, Thomas. 2008. "The Changing Nature of Wage Inequality." *Journal of Population Economics*, 21: 21–48.

Lemieux, Thomas, W. Bentley MacLeod, and Daniel Parent. 2009. "Performance Pay and Wage Inequality." *The Quarterly Journal of Economics*, 124(1): 1–49.

Leonardi, Marco. 2015. "The Effect of Product Demand on Inequality: Evidence from the United States and the United Kingdom." *American Economic Journal: Applied Economics*, 7(3): 221–247.

Levine, David K. 1998. "Modeling Altruism and Spitefulness in Experiments." *Review of Economic Dynamics*, 1(3): 593–622.

Levitt, Steven D. 2004. "Why are gambling markets organized so differently from financial markets?" *The Economic Journal*, 114(3): 223–246.

Levitt, Steven D., and John A. List. 2007. "What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?" *The Journal of Economic Perspectives*, 21(2): 153–174.

Levitt, Steven D., and John A. List. 2008. "Homo Economicus Evolves." *Science*, 319(5865): 909–910.

Levy, Frank, and Peter Temin. 2007. "Inequality and Institutions in 20th Century America." *NBER Working Paper #13106.*

Levy, Frank, and Richard J. Murnane. 1992. "U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations." *Journal of Economic Literature*, 30(3): 1333–1381.

Lewis, Ethan. 2011. "Immigration, Skill Mix, and Capital Skill Complementarity." *The Quarterly Journal of Economics*, 126(2): 1029–1069.

Lileeva, Alla, and Daniel Trefler. 2010. "Improved Access to Foreign Markets Raises Plant-level Productivity...For Some Plants." *The Quarterly Journal of Economics*, 125(3): 1051–1099.

Lindley, Joanne, and Stephen Machin. 2011. "Rising Wage Inequality and Postgraduate Education."

Lindsey, Brink. 2009. "Paul Krugman's Nostalgianomics: Economic Policies, Social Norms, and Income Inequality." *The Cato Institute*.

Lise, Jeremy, and Shannon Seitz. 2011. "Consumption Inequality and Intra-household Allocations." *Review of Economic Studies*, 78(1): 328–355.

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

List, John A. 2004. "Neoclassical Theory Versus Prospect Theory: Evidence from the Marketplace." *Econometrica*, 72(2): 615–625.

List, John A. 2011. "Does Market Experience Eliminate Market Anomalies? The Case of Exogenous Market Experience." *American Economic Review (Papers and Proceedings)*, 101(3): 313–317.

Lockwood, Benjamin B., and Matthew Weinzierl. 2015. "De Gustibus non est Taxandum: Heterogeneity in Preferences and Optimal Redistribution." *Journal of Public Economics*, 124: 74–80.

Long, Jason, and Joseph Ferrie. 2013. "Intergenerational Occupational Mobility in Britain and the U.S. Since 1850." *The American Economic Review*, 103(5): 1109–1137.

Lubbers, Marcel, Mérove Gijsberts, and Peer Scheepers. 2002. "Extreme right-wing voting in Western Europe." *European Journal of Political Research*, 41(3): 345–378.

Luca, Michael. 2016. "Reviews, Reputation, and Revenue: The Case of Yelp.com." *Harvard Business School Working Paper 12-016*.

Luttmer, Erzo F. P. 2001. "Group Loyalty and the Taste for Redistribution." *Journal of Political Economy*, 109(3): 500–528.

Luttmer, Erzo F. P. 2005. "Neighbors as Negatives: Relative Earnings and Well-Being." *Quarterly Journal of Economics*, 120(3): 963–1002.

Machin, Stephen, and John Van Reenen. 1998. "Technology and Changes in Skill Structure: Evidence from Seven OECD Countries." *The Quarterly Journal of Economics*, 113(4): 1215–1244.

Macunovich, Diane J. 1998. "Relative Cohort Size and Inequality in the United States." *The American Economic Review: Papers and Proceedings of the Hundred and Tenth Annual Meeting of the American Economic Association*, 88(2): 259–264.

Macunovich, Diane J. 1999. "The Fortunes of One's Birth: Relative Cohort Size and the Youth Labor Market in the U.S." *Journal of Population Economics*, 12(2): 215–272.

MaCurdy, Thomas. 2015. "How Effective Is the Minimum Wage at Supporting the Poor?" *Journal of Political Economy*, 123(2): 497–545.

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

Manacorda, Marco, Alan Manning, and Jonathan Wadsworth. 2012. "The Impact of Immigration on the Structure of Wages: Theory and Evidence From Britain." *Journal of the European Economic Association*, 10(1): 120–151.

Mankiw, N. Gregory. 2013. "Defending the One Percent." *Journal of Economic Perspectives*, 27(3): 21–34.

Manova, Kalina. 2013. "Credit Constraints, Heterogeneous Firms, and International Trade." *The Review of Economic Studies*, 80(2): 711–744.

Manova, Kalina, Shang-Jin Wei, and Zhiwei Zhang. 2015. "Firm Exports and Multinational Activity Under Credit Constraints." *Review of Economics and Statistics*, 97(3): 574–588.

Mastrorocco, Nicola, and Luigi Minale. 2016. "Information and Crime Perceptions: Evidence from a Natural Experiment." *CReAM Discussion Paper CPD 01/16.*

Mazzolari, Francesca, and Giuseppe Ragusa. 2013. "Spillovers from High-Skill Consumption to Low-Skill Labor Markets." *Review of Economics and Statistics*, 95(1): 74–86.

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics*, 142(2): 698–714.

Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, 71(6): 1695–1725.

Melitz, Marc J., and Gianmarco I. P. Ottaviano. 2008. "Market Size, Trade, and Productivity." *Review of Economic Studies*, 75(1): 295–316.

Melitz, Marc J., and Stephen J. Redding. 2014. "Missing Gains from Trade?" *American Economic Review (Papers and Proceedings)*, 104(5): 317–321.

Meltzer, Allan H., and Scott F. Richard. 1981. "A Rational Theory of Government." *Journal of Political Economy*, 89(5): 914–927.

Meng, Kyle C. 2016. "Using a Free Permit Rule to Forecast the Marginal Abatement Cost of Proposed Climate Policy."

Meyer, Bruce D., and James X. Sullivan. 2011. "The Material Well-Being of the Poor and the Middle Class Since 1980."

Meyer, Bruce D., and James X. Sullivan. 2013. "Consumption and Income Inequality and the Great Recession." *American Economic Review (Papers and Proceedings)*, 103(3): 178–183.

Mian, Atif, Amir Sufi, and Francesco Trebbi. 2014. "Resolving Debt Overhang: Political Constraints in the Aftermath of Financial Crises." *American Economic Journal: Macroeconomics*, 6(2): 1–28.

Miao, Qing, and David Popp. 2013. "Necessity as the Mother of Invention: Innovative Responses to Natural Disasters." *NBER Working Paper #19223.*

Miller, Joshua B, and Adam Sanjurjo. 2015. "Surprised by the Gambler's and Hot Hand Fallacies? A Truth in the Law of Small Numbers." *IGIER Working Paper #552.*

Mincer, Jacob. 1974. *Schooling, Experience, and Earnings.* Vol. I, New York:Columbia University Press.

Mincer, Jacob. 1997. “Changes in Wage Inequality, 1970-1990.” *Research in Labor Economics*, 16: 1–18.

Monte, Ferdinando. 2011. “Skill Bias, Trade, and Wage Dispersion.” *Journal of International Economics*, 83(2): 202–218.

Moretti, Enrico. 2013. “Real Wage Inequality.” *American Economic Journal: Applied Economics*, 5(1): 65–103.

Mudde, Cas. 2007. *Populist Radical Right Parties in Europe.* Cambridge:Cambridge University Press.

Mullainathan, Sendhil, Joshua Schwartzstein, and Andrei Shleifer. 2008. “Coarse thinking and persuasion.” *The Quarterly Journal of Economics*, 123(2): 577–619.

Mulligan, Casey B., and Yona Rubinstein. 2008. “Selection, investment, and women’s relative wages over time.” *The Quarterly Journal of Economics*, 123(3): 1061–1110.

Murphy, Kevin M., and Finis Welch. 1992. “The Structure of Wages.” *The Quarterly Journal of Economics*, 107(1): 285–326.

Muûls, Mirabelle. 2015. “Exporters, Importers and Credit Constraints.” *Journal of International Economics*, 95(2): 333–343.

Nordhaus, William D. 2007. “Two Centuries of Productivity Growth in Computing.” *Journal of Economic History*, 67(1): 128–159.

Norton, Edrvard C., Hua Wang, and Chunrong Ai. 2004. “Computing interaction effects and standard errors in logit and Probit models.” *The Stata Journal*, 4(2): 154–167.

Ohlin, Bertil G. 1933. *Interregional and international trade.* Cambridge, MA:Cambridge University Press.

Okazawa, Ryosuke. 2013. “Skill-biased technical change, educational choice, and labor market polarization: the U.S. versus Europe.” *The Journal of Economic Inequality*, 11: 321–342.

Olsen, Mancur. 1965. *The Logic of Collective Action.* Harvard:Harvard University Press.

Oprea, Ryan, David Porter, Chris Hibbert, Robin D. Hanson, and Dorina Tila. 2008. “Can Manipulators Mislead Market Observers?”

O’Rourke, Kevin H., and Richard Sinnott. 2006. “The determinants of individual attitudes towards immigration.” *European Journal of Political Economy*, 22(4): 838–861.

Ottaviano, Gianmarco I. P., and Giovanni Peri. 2012. “Rethinking the Effect of Immigration on Wages.” *Journal of the European Economic Association*, 10(1): 152–197.

Paglin, Morton. 1975. “The Measurement and Trend of Inequality: A Basic Revision.” *The American Economic Review*, 65(4): 598–609.

Persson, Torsten, and Guido Tabellini. 1994. “Is Inequality Harmful for Growth?” *The American Economic Review*, 84(3): 600–621.

Peters, Katrin, and Monika Schnitzer. 2015. “Trade liberalization and credit constraints: Why opening up may fail to promote convergence.” *Canadian Journal of Economics*, 48(5).

Pierce, Brooks. 2001. "Compensation Inequality." *The Quarterly Journal of Economics*, 116(4): 1493–1525.

Piketty, Thomas, and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913-1998." *The Quarterly Journal of Economics*, 118(1): 1–39.

Piketty, Thomas, Emmanuel Saez, and Stefanie Stantcheva. 2014. "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities." *American Economic Journal: Economic Policy*, 6(1): 230–271.

Piven, Frances Fox, and Richard Cloward. 1971. *Regulating the Poor: The Functions of Public Welfare.* New York:Random House.

Pope, Devin G., and Maurice E. Schweitzer. 2011. "Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes." *American Economic Review*, 101(1): 129–157.

Poterba, James M. 2014. "Retirement Security in an Aging Society." *NBER Working Paper #19930.*

Poutvaara, Panu, and Max Friedrich Steinhardt. 2015. "Bitterness in Life and Attitudes Towards Immigration." *CESifo Working Paper No. 5661.*

Prettner, Klaus. 2013. "Population aging and endogenous economic growth." *Journal of Population Economics*, 26(2): 811–834.

Rajan, Raghuram G., and Luigi Zingales. 1998. "Financial Dependence and Growth." *The American Economic Review*, 88(3): 559–586.

Razin, Assaf, Efraim Sadka, and Phillip Swagel. 2002*a*. "Tax burden and migration: a political economy theory and evidence." *Journal of Public Economics*, 85(2): 167–190.

Razin, Assaf, Efraim Sadka, and Phillip Swagel. 2002*b*. "The Aging Population and the Size of the Welfare State." *Journal of Political Economy*, 110(4): 900–918.

Reuter-Lorenz, Patricia A., and Denise C. Park. 2014. "How Does it STAC Up? Revisiting the Scaffolding Theory of Aging and Cognition." *Neuropsychology Review*, 24(3): 355–370.

Rhode, Paul W., and Koleman S. Strumpf. 2004. "Historical Presidential Betting Markets." *The Journal of Economic Perspectives*, 18(2): 127–141.

Rhode, Paul W., and Koleman S. Strumpf. 2008. "Manipulating Political Stock Markets: A Field Experiment and a Century of Observational Data."

Ricardo, David. 1817. "On the Principles of Political Economy and Taxation." In *The Works and Correspondence of David Ricardo.* , ed. Piero Sraffa. Cambridge:Cambridge University Press, 1951.

Rojas, Juan A. 2005. "Life-cycle earnings, cohort size effects and social security: A quantitative exploration." *Journal of Public Economics*, 89(2-3): 465–485.

Rosen, Sherwin. 1981. "The Economics of Superstars." *The American Economic Review*, 71(5): 845–858.

Rothschild, David M., and Rajiv Sethi. 2015. "Trading Strategies and Market Microstructure: Evidence from a Prediction Market."

Salganik, Matthew J., Peter Sheridan Dodds, and Duncan J. Watts. 2006. "Experimental Study of Inequality and Cultural Market." *Science*, 311(5762): 854–856.

Sánchez-Pagés, Santiago, and Ángel Solano García. 2015. "Immigration, Conflict and Redistribution." *The Scandinavian Journal of Economics*, forthcoming.

Scheve, Kenneth F., and Matthew J. Slaughter. 2001. "Labor Market Competition and Individual Preferences Over Immigration Policy." *Review of Economics and Statistics*, 83(1): 133–145.

Schindler, Robert M., and Patrick N. Kirby. 1997. "Patterns of Rightmost Digits Used in Advertised Prices: Implications for Nine-Ending Effects." *Journal of Consumer Research*, 24(September): 192–201.

Schmookler, Jacob. 1966. *Invention and Economic Growth.* Cambridge:Harvard University Press.

Simon, Herbert A. 1955. "A Behavioral Model of Rational Choice." *The Quarterly Journal of Economics*, 69: 99–118.

Ski World Cup Wengen. 2012. *Fakten und Zahlen.* Wengen.

Snowberg, Erik, and Justin Wolfers. 2010. "Explaining the Favorite-Long Shot Bias: Is It Risk-Love or Misperceptions?" *The Journal of Political Economy*, 118(4): 723–746.

Solt, Frederick. 2009. "Standardizing the World Income Inequality Database." *Social Science Quarterly*, 90(2): 231–242, SWIID Version 4.0, September 2013.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti. 2011. "Growing Like China." *The American Economic Review*, 101(2): 196–233.

Sowell, Thomas. 2011. *The Thomas Sowell Reader.* New York:Basic Books.

Stango, Victor, and Jonathan Zinman. 2014. "Limited and Varying Consumer Attention: Evidence from Shocks to the Salience of Bank Overdraft Fees." *Review of Financial Studies*, 27(4): 990–1030.

Stanovich, Keith E., and Richard F. West. 2000. "Individual differences in reasoning: implications for the rationality debate?" *The Behavioral and Brain Sciences*, 23(5): 645–665.

Stolper, Wolfgang F., and Paul A. Samuelson. 1941. "Protection and Real Wages." *The Review of Economic Studies*, 9(1): 58–73.

Storesletten, Kjetil. 2000. "Sustaining Fiscal Policy Through Immigration." *Journal of Political Economy*, 108(2): 300.

Storesletten, Kjetil. 2003. "Fiscal Implications of Immigration: A Net Present Value Calculation." *The Scandinavian Journal of Economics*, 105(3): 487–506.

Strömberg, David. 2015. "Media Coverage and Political Accountability: Theory and Evidence." In *Handbook of Media Economics.* , ed. Simon Anderson, David Strömberg and Joel Waldfogel, 1–26. Amsterdam:North-Holland.

Strulik, Holger, and Katharina Werner. 2016. "50 Is the New 30 - Long-Run Trends of Schooling and Retirement Explained By Human Aging." *Journal of Economic Growth*, 21: 165–187.

Strulik, Holger, Klaus Prettner, and Alexia Prskawetz. 2013. "The past and future of knowledge-based growth." *Journal of Economic Growth*, 18(4): 411–437.

Taleb, Nassim Nicholas. 2015. "How to (Not) Estimate Gini Coefficients for Fat Tailed Variables."

Taylor, Paul, Rick Fry, and Russ Oates. 2014. "The Rising Cost of Not Going to College." Pew Research Center, Washington D.C.

Thaler, Richard H., and William T. Ziemba. 1988. "Anomalies: Parimutuel Betting Markets: Racetracks and Lotteries." *The Journal of Economic Perspectives*, 2(2): 161–174.

Thomson, Russell. 2015. "The Effectiveness of R&D Tax Credits." *The Review of Economics and Statistics*, forthcoming.

Tinbergen, Jan. 1975. *Income Distribution: Analysis and Policies.* Amsterdam:North Holland.

Topel, Robert H. 1994. "Regional Labor Markets and the Determinants of Wage Inequality." *The American Economic Review*, 84(2): 17–22.

Topel, Robert H. 1997. "Factor Proportions and Relative Wages: The Supply-Side Determinants of Wage Inequality." *The Journal of Economic Perspectives*, 11(2): 55–74.

Tran, Anh, and Richard Zeckhauser. 2012. "Rank as an inherent incentive: Evidence from a field experiment." *Journal of Public Economics*, 96(9-10): 645–650.

Tversky, Amos, and Daniel Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science*, 185(4157): 1124–31.

Verhoogen, Eric A. 2008. "Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector." *The Quarterly Journal of Economics*, 123(2): 489–530.

von Weizsäcker, Robert K. 1996. "Distributive implications of an aging society." *European Economic Review*, 40(3-5): 729–746.

Welch, Finis. 1979. "Effects of cohort size on earnings: The baby boom babies' financial bust." *The Journal of Political Economy*, 87(5): S65–S97.

Wilkes, Rima, and Catherine Corrigall-Brown. 2011. "Explaining time trends in public opinion: Attitudes towards immigration and immigrants." *International Journal of Comparative Sociology*, 52(1-2): 79–99.

Williamson, Jeffrey G. 1998. "Growth, Distribution, and Demography: Some Lessons from History." *Explorations in Economic History*, 35(3): 241–271.

Wolfers, Justin, and Eric Zitzewitz. 2006. "Prediction Markets in Theory and Practice." *NBER Working Paper #12083.*

Wolff, Edward Nathan. 1996. "The Growth of Information Workers in the U.S. Economy, 1950-1990: The Role of Technological Change, Computerization, and Structural Change." *C.V. Starr Center for Applied Economics, New York University*, 96-41.

Wood, Adrian. 1995. "How Trade Hurt Unskilled Workers." *The Journal of Economic Perspectives*, 9(3): 57–80.

Curriculum Vitae

Stefan Legge

Born October 22, 1986 in Siegen (Germany)

EDUCATION

Ph.D. in Economics and Finance, University of St.Gallen, 09/2016
M.A. in Economics, University of St.Gallen, 09/2011
B.Sc. in Economics, University of Mannheim, 03/2009

EXTENDED VISITS

Visiting Scholar, Princeton University, 07/2015 - 01/2016
Visiting Scholar, UC Berkeley, 09/2014 - 06/2015
Visiting Student, UC San Diego, 09/2010 - 12/2010

EXPERIENCE

Research Assistant, SIAW-HSG, University of St.Gallen, 09/2011 - present
Research Assistant, SEW-HSG, University of St.Gallen, 10/2009 - 07/2010
Student Assistant, University of Mannheim, 06/2007 - 01/2009
Intern, BMW Group AG, Munich, 03/2009 - 08/2009
Intern, ALDI Süd, Frankfurt, 08/2008

SCHOLARSHIPS AND AWARDS

Prize for the Best Dissertation in Economics, University of St.Gallen, 2016
Young Economist Award, Swiss Society of Economics and Statistics, 2014
Scholarship for Research at UC Berkeley, Swiss National Science Foundation, 2014
Presentation Grant, Verein für Socialpolitik, 2013 and 2014
Best Paper Prize, Warsaw International Economic Meeting, 2013
Scholarship for Outstanding Doctoral Candidates, e-fellows, 2009 - 2016

PRESENTATIONS

2016: Royal Economic Society, SSES, University of St.Gallen

2015: European Winter Meeting of the Econometric Society, Princeton University, Midwest Macro Conference, Northeastern Political Science Association

2014: Midwest International Trade Conference, ESEM, EEA, International Economic Association, CEPR/Sinergia Cenference, Warsaw International Economic Meeting, Spring Meeting of Young Economists, Ruhr Graduate School, University of St.Gallen

2013: NBER Summer Institute, EEA, Warsaw International Economic Meeting, Warwick University, Spring Meeting of Young Economists

2012: University of St.Gallen, University of Zurich, Sinergia Workshop