

Matti Sarvimäki

Essays on Migration

HELSINGIN KAUPPAKORKEAKOULU
HELSINKI SCHOOL OF ECONOMICS

A-341



Matti Sarvimäki

Essays on Migration

HELSINKI SCHOOL OF ECONOMICS

ACTA UNIVERSITATIS OECONOMICAE HELSINGIENSIS

A-341

© Matti Sarvimäki and
Helsinki School of Economics

ISSN 1237-556X
ISBN 978-952-488-305-4

E-version:
ISBN 978-952-488-306-1

Helsinki School of Economics -
HSE Print 2009

Author	Matti Sarvimäki Research Economist, Spatial Economics Research Centre London School of Economics and Political Science Houghton Street, London WC2A 2AE, United Kingdom m.sarvimaki@lse.ac.uk
Title	Essays on Migration
Supervisors	Roope Uusitalo Research Director Government Institute for Economic Research Pertti Haaparanta Professor, Helsinki School of Economics
Preliminary Examiners	Kristiina Huttunen Senior Researcher Labour Institute for Economic Research Olof Åslund Associate Professor IFAU and Uppsala University
Custos (Chairperson) of the defence	Professor Pertti Haaparanta
Opponent	Professor David Card University of California, Berkeley
Type of Research	Based on essays
Essays	Matti Sarvimäki and Aki Kangasharju, “Temporary Shocks and the Geography of Economic Activity: Evidence from a Large-Scale Resettlement Policy”, <i>Unpublished</i> Matti Sarvimäki, Roope Uusitalo and Markus Jäntti, “Long-Term Effects of Forced Migration”, <i>Unpublished</i> Matti Sarvimäki, “Assimilation to a Welfare State: Labor Market Performance and Use of Social Benefits by Immigrants to Finland”, <i>Unpublished</i> Matti Sarvimäki and Kari Hämäläinen, “Moving Immigrants from Welfare to Work: The Impact of an Integration Program”, <i>Unpublished</i>

Abstract

This dissertation consists of four empirical essays and an introductory chapter on the economics of migration. The first two essays study the long-term effects of relocating 11% of the Finnish population after the World War II. The first essay suggests that this shock set off a self-reinforcing process, where municipalities receiving the largest flows of displaced persons grew faster also in later periods. The second essay shows that being displaced increased long-term income of men, but had no effect on income of women. A large part of the effect is attributed to faster transition from traditional (rural) to modern (urban) occupations among the displaced. The remaining essays focus on more recent immigrants to Finland. The third essay finds that upon arrival immigrants earn substantially less than comparable natives. Only men from OECD countries converge to natives' earnings over time. The last essay suggests that the a change in the integration policy in 1999 increased immigrants' employment and earnings and decreased their dependency on social benefits.

Keywords: Migration, labor economics, economic geography

JEL Codes: J61, J31, F12, F22, N94, O15, R12, R23, H43

Acknowledgments

I have been privileged to write these essays while working in several outstanding academic institutions and being surrounded by brilliant colleagues. The Government Institute for Economic Research (VATT) has kindly provided me an office, data, computing facilities and often even a salary ever since I joined them to work on my masters thesis in fall 2002. It is clear that this dissertation would not exist without this support. I have also greatly benefited from the opportunity to visit the PhD Program in Economics at Universidad Carlos III de Madrid for the fall term 2003 and the Centre for Research and Analysis on Migration at the Department of Economics, University College London for the academic year 2005–2006. Financial support from VATT, Yrjö Jahnsson Foundation, Jenny and Antti Wihuri Foundation and Palkansaaajasäätiö as well as the Graduate School Fellow position at Helsinki School of Economics / Finnish Doctoral Programme in Economics in 2006–2007 allowed me to work full-time on my studies. Furthermore, it is hard to imagine a better place to finish this dissertation than the Spatial Economics Research Centre at the London School of Economics and Political Science, where I have worked since fall 2008.

I have been very fortunate to have Roope Uusitalo as my supervisor. He used a vast amount of time to give first-class advice and support, and provided an outstanding input as a co-author of one of the essays. It has also been a privilege to work with Kari Hämäläinen, Markus Jäntti and Aki Kangasharju, who have played important roles both as co-authors and mentors. Pertti Haaparanta has done a great job as my second supervisor at HSE. Christian Dustmann supervised me at UCL and profoundly influenced the way I have approached research ever since. Kristiina Huttunen and Olof Åslund were excellent external examiners, whose insightful comments and suggestions improved the essays considerably. I also thank Reino Hjerppe, Seija Ilmakunnas and Jaakko Kiander for the opportunity to pursue this research at VATT.

It has been a great honor and joy to join the research communities in Helsinki and London. In addition to those already mentioned above, a large number of people have influenced my work, helped in various ways or just made life fun. My special thanks go to Teemu Lyytikäinen and Jouko Verho, who have proved particularly distinguished on all these accounts. I would also like to thank Essi Eerola, Elias Einiö, Tommaso Frattini, Albrecht Glitz, Raimo Hintikka, Pekka Ilmakunnas, Seppo Kari, Ossi Korkeamäki, Tuomas Kosonen, Tomi Kyyrä, Josep Mestres, Ville Mälkönen, Sari Pekkala Kerr, Tuomas Pekkarinen, Torsten Santavirta, Daniel Sturm, Otto Toivanen, Juha Tuomala, Janne Tukiainen and Juuso Välimäki. The list could, and probably should, be much longer. In particular, I am omitting many great friends who are not academic economists.

I am also very much indebted to my family. My parents, Ilkka and Pirjo, have first and foremost provided a home filled with love and protection, but also an environment where intellectual curiosity and education were always highly valued. My sister Riikka has been a truly great little sister. I have been exceptionally lucky in terms of in-laws, too. In fact, Pertti and Terttu Ollikainen, as well as my brother-in-laws Juha and Klaus Ollikainen, should have been mentioned already at the beginning of these acknowledgments, as a substantial part of this dissertation was written while enjoying their hospitality in Lapinlahti. Of course, my greatest gratitude goes to my wife Assi and our son Leo, who make me a very happy man every day.

London, February 2009

Matti Sarvimäki

Contents

Abstract	5
Acknowledgments	7
Chapter 1. Introduction	11
1.1. Migration and Trade	11
1.2. Migration and the Labor Market	14
References	23
Chapter 2. Temporary Shocks and the Geography of Economic Activity: Evidence from a Large-Scale Resettlement Policy	27
2.1. Introduction	27
2.2. Post-WWII Resettlement in Finland	29
2.3. Data	33
2.4. Empirical Model	36
2.5. Results	38
2.6. Conclusions	45
References	47
Chapter 3. Long-Term Effects of Forced Migration	49
3.1. Introduction	49
3.2. Historical Background	51
3.3. Data	55
3.4. Empirical Approach	59
3.5. Results	61
3.6. Robustness Checks	73
3.7. Discussion	78
3.8. Conclusions	80
References	81
Appendix	82

Chapter 4. Assimilation to a Welfare State: Labor Market Performance and Use of Social Benefits by Immigrants to Finland	85
4.1. Introduction	85
4.2. Background	87
4.3. Data	90
4.4. Empirical Framework	92
4.5. Results	95
4.6. Conclusions	103
References	105
Appendix	107
Chapter 5. Moving Immigrants from Welfare to Work: The Impact of an Integration Program	109
5.1. Introduction	109
5.2. Immigration to Finland	111
5.3. The Integration Act	112
5.4. Empirical Strategy	114
5.5. Data and Results	117
5.6. Robustness Checks	125
5.7. Conclusions	132
References	134
Appendix	136

CHAPTER 1

Introduction

This dissertation consists of four empirical essays on economics of migration. While migration is a topic that has been widely studied both in economics and in virtually all fields of social sciences, it seems fair to argue that these essays provide interesting new results. The aim of this introductory chapter is to back up this claim by providing a review of the literature and by introducing the contributions of the essays. I will also discuss briefly the data and the methodological approach.

The chapter is organized as follows. The next section reviews the work on migration done by trade economists. As discussed in more detail below, early theoretical contributions on migration were extensions of the standard model for international trade. These insight are still very influential. Thus I review them in some detail before focusing on the more recent literature on new economic geography and the contributions of the first essay. Section 1.2 turns to the work on migration in the field of labor economics. I begin by briefly discussing the empirical results on one of the most important question with regard to the policy debate: the impact of immigration on native wages and employment. This literature also provides a good framework to illustrate the “experimentalists” approach used in the essays. The last three subsections review literature on the immigrant experience and discuss the contributions of the three remaining essays.

1.1. Migration and Trade

1.1.1. The Heckscher-Ohlin Model. The early studies on economics of migration grew from the analysis of international trade. The influence of this work is illustrated by the fact that the curricula of virtually all present economics undergraduate programs include the Heckscher-Ohlin model (henceforth the HO-model). As a consequence, most economics students will learn that trade and migration are substitutes. The intuition behind this argument, first put forth by Mundell (1957), is that whether labor inputs are imported in the form of immigrants or goods is almost irrelevant:

if factors of production move easily between industries and if transportation costs are unimportant, both mechanisms lead to factor price equalization. This will lead to an increase in world output and a transfer of income from the previously scarce to the previously abundant resources. In other words, both trade and immigration would increase the income of capital owners and high-skilled labor in the rich countries, while hurting the low-skilled workers. The consequences for poor countries would be the opposite.

The prediction that immigration hurts low-skilled natives and benefits the others seems to be widely accepted even among people who never attended an economics class. Interestingly, however, researchers have struggled to establish empirical evidence on the adverse effect of immigration on native wages and employment (see below). At first sight, this seems to contradict the basic assumption of conventional economic theory, namely that increases in labor supply should lead to lower wages or higher unemployment. However, this apparent paradox is also explained by the HO-model. As first noted by Rybczynski (1955), labor flows or trade affect wages in the HO-model only relative to the situation where both factor mobility and trade are restricted. If trade is free and transportation cost can be neglected, migration affects only the structure of production.

While the HO-model provides several important insights, it also omits many important factors. Of course, this is what a useful theory needs to do in order to clearly illustrate one particular mechanism. Nevertheless, this dissertation focuses on some of these mechanisms assumed away in the HO-model. I will next discuss the background and contributions of each essay in more detail.

1.1.2. (New) Economic Geography. At the core of the HO-model are two assumptions: constant returns to scale and differences in factor endowments across locations. An important implication of these assumptions is that the spatial distribution of economic activity should be stable. In the language of economic geographers, “first nature” or “fundamentals” alone would determine the regional structure.

There is a long tradition both in economics and geography to challenge this prediction. In economics, the idea that something beyond the fundamentals might play an important role date back to at least Marshall (1890). However, for a long time economists did not know how to express these ideas

formally, and thus their influence was limited. More recently, a class of formal models, typically referred to as the “new economic geography” (NEG), has popularized these insights among mainstream economists.

The NEG theories have their roots in the “new trade theory” following Krugman (1979), which studied the implications of imperfect competition, economies of scale and transportation costs on international trade (see Helpman and Krugman, 1985, for a review). The key extension leading to NEG, usually attributed to Krugman (1991), was to make the location decisions of firms and workers endogenous. This allowed the theory to explain how circular causation may explain the tendency of economic activity to agglomerate. The NEG theories put geographical mobility of workers and firms at the center of the explanation. The intuition is the following: Firms want to concentrate their production in order to exploit economies of scale. Firms also want to be close to consumers in order to save in transportation costs. On the other hand, workers want to be where the firms are, in order to benefit from well-paid jobs. Workers are also consumers, who want to be close to firms in order to benefit from a more competitive market and a larger variety of products. That is, everyone want to be where everyone else are, simply because the others are there. A consequence of such circular causation is that there are several possible locations where economic activity could be concentrated in and which of these potential equilibria materializes depends on initial conditions, history of shocks and expectations (see Fujita et al., 1999; Baldwin et al., 2003, for reviews of the theoretical literature).

These theories are highly relevant for policy makers as they suggest that a sufficiently large intervention could set off a self-reinforcing process. For instance, one could argue that subsidizing a certain region would allow it to reach a critical mass, after which it would continue to grow without further expense to the tax-payer.

However, there is little evidence on such process actually taking place. In fact, Head and Mayer (2004, p. 2662) argue in their survey that the current empirical evidence “suggests that those two celebrated characteristics of NEG models [long-term effects of temporary shocks and the role of historical accidents] should perhaps be considered more as fascinating theoretical ‘exotica’ rather than as robust elements of economic geography”. To a large extent, this conclusion is due to influential work by Davis and Weinstein (2002), who show that the city size distribution of Japan appears to

be remarkably stable, even to the massive bombings during the World War II.

The first essay, written jointly with Aki Kangasharju, suggests that the assertion of empirical irrelevance of the NEG may have been premature. We use municipality level data for the years 1930–2000 and a similar approach as Davis and Weinstein (2002) to study a somewhat different situation. We focus on migrations that took place during and after the World War II, when Finland ceded a tenth of its territory to the Soviet Union and resettled 11 percent of the population to the remaining parts of country. Characteristics of the settlement policy created a quasi-experimental setting, which allows us to identify the persistence of temporary shocks on the spatial concentration of rural population.

We find that the resettlement had a permanent effect. Furthermore, the results suggest that it set off a self-reinforcing process, where rural municipalities receiving the largest flows of displaced persons grew faster (or declined slower) also in later periods. To the best of our knowledge, these are among the first quasi-experimental results suggesting that a temporary shock has set off a self-reinforcing process pushing the distribution of population from one equilibrium to another.

1.2. Migration and the Labor Market

While trade theory has provided some important insights on the understanding of migration, a large share of the current literature is written by labor economists. In contrast to the trade literature, this work has been predominantly empirical. While the questions asked by empirical labor economist range from attitudes towards immigrants (Scheve and Slaughter, 2001; Dustmann and Preston, 2007) to the impact of immigration on prices (Lach, 2007; Cortes, 2008), public finances (Smith and Edmonston, 1997; Storesletten, 2000) and technology adaption (Lewis, 2003, 2005), just to name a few, it nevertheless seems reasonable to limit this review to two questions that have dominated the literature: (1) the impact of immigration on native wages and employment and (2) labor market assimilation of immigrants. The next subsection shortly reviews the literature assessing the first question. However, the idea is to use this literature as an example in order to illustrate the methodological approach, often referred to as “experimentalism”, used in the following essays. The remaining subsections review studies on the immigrant experience and discuss the contributions of the remaining essays.

1.2.1. Impact of Immigration on Native Wages: Structuralists vs. Experimentalists Approach. The question of how much immigrants hurt the labor market prospects of natives is not new (see Hatton and Williamson, 2006, for a review for the past two centuries). However, it gained new prominence in the public and academic debate following an influential paper by Card (1990). For non-economists, the issue was the surprising result: immigration did not appear to affect the labor market outcomes of natives at all. For economists, the interest was both on the result and on the empirical approach. With respect to the former, the paper led to an outburst of studies aiming to assess whether Card's results were correct and if they were, why (see Dustmann and Glitz, 2005, for a recent review). This debate is still active and there is currently no consensus on whether the impact of immigration on wages of low-skilled natives is negligible (Card, 1990, 2001, 2005) or important (Borjas et al., 1997; Borjas, 2003).

More importantly, however, Card's (1990) paper became one of the best known examples of using a "natural experiment" to identify a causal relationship in observational data. Thus it had a large impact on arguably the most important methodological debate in empirical economics: the one between "structuralists" and "experimentalists" (for expositions of both sides see Keane, 2006; Angrist and Pischke, 2009). Fully reviewing this debate is beyond the scope of this chapter. Yet, it is illustrative to briefly look at two alternative ways that economists have used to assess the impact of immigration on native wages.

The fundamental question in all causal empirical work concerns the construction of a counterfactual. For instance, when we ask what is the causal effect of immigration to native wages, we want to know how native wages in the presence of certain amount of immigration compare to the counterfactual state, where no immigrants had arrived. The challenge is that only one state of the world – the one that actually took place – is observed and thus the counterfactual has to be produced somehow. The essential difference in empirical approaches is in the way this task is accomplished.

The "structuralists" construct the counterfactual by using an economic model. One would typically start from a formal theoretical model describing what one believes to be the essential mechanisms of the real world. The form of these mechanisms (the structure) is assumed *a priori*, but their magnitudes (parameter values) are usually estimated from data. Alternatively some or

all parameter values are plugged in to the model, typically based on estimates presented in previous literature.

Borjas et al. (1997) provide an example of using a formal model and outside information for the parameter values to study the impact of trade and immigration to the wages of low-skilled workers in the United States.¹ They model the aggregate technology by CES production function with two inputs (skilled and unskilled labor) and plug in the key parameter values (wage elasticities) from two previous studies (Borjas et al., 1992; Katz and Murphy, 1992). This allows them to construct a counterfactual of what would have happened to the wages of low-skilled natives in the absence of immigration. Comparing this number to the observed wages leads to a conclusion that immigration to the United States has had a marked adverse impact on the economic status of native high-school dropouts.

By contrast, the experimentalist approach attempts to minimize assumptions on the underlying mechanisms. To do this, the counterfactual is constructed by seeking a plausible control group. In the best case, one is able to run an experiment and allocate units to a treatment and a control group randomly. If such an experiment is successfully implemented, the causal effect is identified by simply comparing means of the two groups. No assumptions on the mechanisms producing the effect are required (see Duflo et al., 2008, for discussion).

The problem is that answering many important questions would require experiments that cannot be executed for practical or ethical reasons. For instance, the challenge of estimating the impact of immigration on native wages is that immigrants tend to move to areas where wages are growing. Hence, it is difficult to separate the impact of immigration on wages from the impact of wages on immigration. This problem could be solved by running an experiment where some locations were randomly allocated to become

¹While Borjas et al. (1997) is the leading example of using a formal model to simulate the impact of immigration on native wages, I do not claim that it would be a fair example to describe structural econometrics in general or that the authors should be considered as “structuralists”. For a more representative review of structural work, see e.g. Belzil (2007); Pakes (2008). It should also be stressed that while some discussant tend to frame the debate as a conflict between two opposing views, most empirical economists see the approaches as complements. Most “experimentalists” do not advocate abandoning theory; just testing it with minimum *a priori* assumptions. Similarly, most “structuralists” do not claim that experiments would be useless; just that experiments do not answer all important questions. Furthermore, many economists would be reluctant to categorize themselves as an “experimentlists” or a “structuralists”. Indeed, some of the most interesting recent work combine both approaches (see e.g. Bandiera et al., 2005; Todd and Wolpin, 2006).

a treatment area, which would receive large immigrant flows, while other locations were left as control areas, which would ban immigration. Clearly, such an experiment cannot be executed.

However, sometimes situations *resembling* an experiment occur. The innovation of Card (1990) was to use a sudden large-scale migration from Cuba to Miami, known as the Mariel Boatlift, as such a “natural experiment”. This migration took place after a peculiar sequence of events led Fidel Castro to declare that Cubans wishing to leave the country were free to do so. Before the policy was suspended, 125,000 Cubans had moved to the United States. Half of them settled to the nearest and most familiar city, Miami. As a consequence, the Miami labor force increased by roughly 7 percent between May and September of 1980. Importantly, this occurred for reasons that had nothing to do with the labor market situation. Thus comparing the wage and employment growth of low-skilled workers in Miami to that of low-skilled workers in similar cities, which did not experience sudden immigrant influx should be informative about the causal effect of immigration on native labor market prospects. Card (1990) showed that wages and employment in Miami evolved in a similar way as those in the comparison cities. Under the critical assumption that the comparison cities provide a good approximation of what would have happened in Miami without the Mariel Boatlift, one can conclude that at least this particular immigration flow did not affect native wages or employment.²

While none of the essays in this dissertation studies the impact of immigration on natives, they are nevertheless heavily influenced by Card (1990). Three use a quasi-experimental approach.³ The “experiment” in the first two essays – displacement of 11% of the Finnish population – closely resembles the Mariel Boatlift. Note that this quasi-experiment allows us to study questions that cannot be addressed using a real experiment. Displacing a large number of people in the name of science is not likely to be in the agenda of any government, and even if it was, one would need to wait decades before

²See Borjas (2003) for a critique of this identifying assumption.

³The remaining essay on labor market assimilation (Chapter 4) should probably be considered as a descriptive study. However, this is not due to the typical reason of omitted variables biasing the parameter of interests. Instead, the topic of the essay – association of time spent in the host country and labor market outcomes – cannot be a cause in a sense of Holland’s (1986, p. 954) definition that “causes are only those things that could, in principle, be treatments in experiments”. That is, it is hard to imagine a treatment that would manipulate the time an immigrant has spent in the host country. Thus years-since-migration is perhaps best understood as an attribute.

observing the outcomes we study. Thus the questions we ask in the first two essays can, in practice, only be answered in a quasi-experimental setting.

By contrast, the last essay uses a quasi-experiment to study a question that would be best assessed with a randomized field experiment. As discussed in more detail below, this essay estimates the impact of an integration program to the labor market outcomes of immigrants using a quasi-experiment created by a legislative date rule determining eligibility to the program. The reason we take a quasi-experimental approach is that there is no real experiment to use. While having a quasi-experiment is certainly better than nothing, it comes with caveats that substantially decrease our ability to provide informative policy advice. Thus both policy and science would be considerably improved if political will to test this type of policies using genuine field experiments would increase. Such experiments would also provide a good opportunity to study whether current structural models are sufficient for providing reliable *ex ante* evaluation of a proposed policy.

1.2.2. Forced Migration. I now turn to the literature studying the immigrant experience. Ever since Sjaastad's (1962) seminal paper, migration decision has been placed in the human capital framework, where migration is an investment that includes costs and benefits for the potential migrant. If the benefits outweigh the costs, the person moves. However, sometimes migrations are not voluntary. While one can always fit the "decision" of being displaced to the human capital framework by reframing the choice set (the alternative to migration being persecution or death), these situations clearly call for different policy response than voluntary migrations. However, little is known either on the impact of forced migration on those who are displaced or on the appropriate policy response.

The second essay, written jointly with Roope Uusitalo and Markus Jäntti, studies one large-scale displacement. As discussed above, one of the outcomes of World War II was ceding a tenth of the Finnish territory and resettling the entire the population living in these areas to the remaining parts of the country. Organizing the resettlement was an immense challenge. The war had left the country of four million with roughly 92,000 dead and 228,000 injured. Much of the production capacity was destroyed and there were large war reparations to be paid. On top of this, 11 percent of the population was displaced and their former homes were now part of the Soviet Union.

The policy response was bold. The Parliament decided that the displaced would be compensated for their losses and these compensation would be financed by levying a massive tax on capital. In particular, the displaced farmers were given agricultural land, half of which was expropriated from private farmers living in their farms.

In the essay, we use individual level panel-data for the years 1939–1990 to study the long-term effects of the displacement on economic outcomes of the displaced. Surprisingly, we find that being displaced *increased* long-term earnings: more than three decades after the war started, displaced men had substantially higher earnings than otherwise similar men who were not forced to move. We are able to explain a large part of this effect by faster mobility away from the traditional rural occupations to modern urban ones.

In my view, the implications of these results go beyond demonstrating that this particular policy turned out to help the displaced. In particular, our findings provide indirect evidence on the importance of non-monetary costs associated with migration. That is, the opportunity to increase one's earnings through migration was available for everyone. Yet, the displaced were far more likely to exploit these opportunities than the others. This finding is consistent with the existence of large (non-monetary) costs to migration. In essence, being displaced reduced these costs in a brutal way. If one is willing to extrapolate from these results, they suggest that reducing migration costs in more pleasant ways – e.g. by subsidizing moves – would improve economic efficiency.

1.2.3. Assimilation of Immigrants. The remaining two essays focus on immigration of foreign-born persons. One of the core question in this literature concerns assimilation of immigrants to the host country's labor market. This branch of literature was launched by Chiswick (1978), who asked how immigrants assimilated to the labor market of the United States. The choice of the word “assimilation” – or Chiswick's original “Americanization” – is perhaps misfortune as, for many, it carries normative weight from the debate on whether immigrants should assimilate in the sense of adopting host country's culture. However, in the context of economics, the question is simply and *only* about whether immigrants' employment, earnings and social benefits converge to those of comparable natives and at what pace. The answer is not entirely clear even in the case of the United States, which has been the subject of the most extensive study. While there is a consensus

on that immigrants tend to have lower earnings than comparable natives upon arrival, it is harder to assess whether immigrants ever reach the natives. When one includes other countries, the results become even more mixed (see Borjas, 1994; Boeri et al., 2002, for surveys).

The third essay adds the case of Finland to this literature. I use individual-level panel data derived by linking information from several administrative registers. I find that the initial immigrant-native earnings gap is large. While immigrants' earnings grow rapidly as they spend more time in Finland, only men from OECD countries close the gap to comparable natives. Women from non-OECD countries do not reach even the low-skilled natives. The results also reveal some more surprising patterns. In particular, they are consistent with the idea, discussed by e.g. Borjas and Hilton (1996), that one form of assimilation may be learning to use the welfare system. More precisely, I find that among non-OECD immigrants, the use of social assistance increases over the first five years in Finland, even though their earnings simultaneously increase substantially. These findings suggest that immigrants initially under-utilize the welfare system.

The essay also contributes to the methodological discussion on how assimilation profiles should be estimated. Earlier studies have shown that the assimilation estimates may be biased due to changes in cohort "quality" (Borjas, 1985), selective emigration (Edin et al., 2000; Lubotsky, 2007), heterogeneous assimilation profiles (Crossley et al., 2001) and trends in labour market conditions (Bratsberg et al., 2006). My essay provides further evidence that neglecting non-random return migration is likely to create some bias. However, the second methodological point is more important, though obvious. I show that neglecting non-random selection into employment – as the current practice tends to be – may lead to severely distorted conclusions. I also discuss a simple way to avoid this bias.

1.2.4. Integration Policy. Most of the literature on the economics of migration – including the three essays of this dissertation discussed above – have attempted to shed light on the underlying mechanisms that affect migration and its consequences. While this research is essential for providing the background information for setting up policy, one would also want to have direct evidence on which policies are efficient and which are not. In particular, there appears to be large demand for policies that would reduce immigrants' reliance on public assistance programs.

Such policies have taken three forms. The first approach has been adopting a skill-based immigration policy such as the point systems introduced by Canada in the late 1960s and by Australia in the early 1970s. If successful, these policies would lead to admitting only those who are the least likely to collect public benefits and the most likely to contribute to tax revenues. However, the evidence on efficiency of point systems is mixed. The differences in the average skill-level of immigrants to Canada, Australia and the United States appear to be entirely due to differences in the mix of source countries (Borjas, 1993; Antecol et al., 2003). Yet, it is not clear whether these differences follow from variation in immigration policies or from geography, history and the share of illegal immigrants. In any case, even if skill-based admission could be successfully implemented, it would carry the cost of lowering global economic efficiency, hurting the sending countries and eliminating the benefits from low-skilled migration to consumers and employers in the receiving countries (see Hanson, 2005, for discussion).

The second approach is to allow low-skilled immigrants to enter, but to limit their rights for public benefits. To a large extent, this is the policy that the United States adopted in the 1996 welfare reform by denying most types of means-tested assistance to non-citizens. Again, it is not clear how large a difference the reform made. While participation rates of immigrants declined faster than participation rates of natives in California, the reform was offset by changes in the provision of state-funded programs and increasing naturalizations elsewhere (Borjas, 2002; Hanson, 2005).

The third approach is to design policies that increase the propensity of low-skilled immigrants to be employed. The leading examples are the integration programs launched by many European countries. These programs are typically framed as a mutual contract between the state and the immigrant. The state commits to provide language training, civic courses, labor market orientation and sometimes vocational training. The immigrant commits to participate and this commitment is typically “encouraged” by making eligibility to welfare benefits and/or gaining permanent residence permit conditional on participation. The evidence on the impacts of these programs is currently largely absent.

The last essay, written jointly with Kari Hämäläinen, studies one European integration program. More precisely, we use individual-level administrative data and focus on integration plans introduced as a part of the 1999 Act on the Integration of Immigrants and Reception of Asylum Seekers. These

plans are an individualized paths of active labor market policy measures combined with a reduction of welfare benefits in the case on non-compliance. The quasi-experimental setting we exploit was created by making participation obligatory to only those who had arrived after May 1st, 1997. This creates a situation where those arriving just before the threshold date serve as a control group for those who arrived just after. Comparison of these two groups allow us to estimate the causal effect of getting an integration plan for a subpopulation of immigrants. Our results suggest that the integration plans have been remarkably efficient in improving immigrants' mid-term employment and earnings and in reducing their dependency on welfare benefits.

References

- ANGRIST, J. D. AND J.-S. PISCHKE (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.
- ANTECOL, H., P. KUHN, AND S. TREJO (2003): "Immigration Policy and the Skills of Immigrants to Australia, Canada, and the United States," *Journal of Human Resources*, 38, 192–218.
- BALDWIN, R., R. FORSLID, P. MARTIN, G. OTTAVIANO, AND F. ROBERT-NICOUD (2003): *Economic Geography and Public Policy*, Princeton University Press.
- BANDIERA, O., I. BARANKAY, AND I. RASUL (2005): "Social Preferences and the Response to Incentives: Evidence from Personnel Data," *Quarterly Journal of Economics*, 120, 917–962.
- BELZIL, C. (2007): "The Return to Schooling in Structural Dynamic Models: a Survey," *European Economic Review*, 51, 1059–1105.
- BOERI, T., G. HANSON, AND B. MCCORMICK (2002): *Immigration Policy and the Welfare System*, Oxford: Oxford University Press.
- BORJAS, G. (1985): "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants," *Journal of Labor Economics*, 3, 463–489.
- (1993): "Immigration Policy, National Origin, and Immigrant Skills: A Comparison of Canada and the United States," in *Small Differences That Matter: Labor Markets and Income Maintenance in Canada and the United States*, ed. by D. Card and R. Freeman, University Of Chicago Press.
- (1994): "The Economics of Immigration," *Journal of Economic Literature*, 32, 1667–1717.
- (2002): "Welfare Reform and Immigrant Participation in Welfare Programs," *International Migration Review*, 36, 1093–1123.
- (2003): "The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market," *Quarterly Journal of Economics*, 118, 1335–1374.
- BORJAS, G., R. FREEMAN, AND L. KATZ (1992): "On the labor market effects of immigration and trade," in *Immigration and the work force*, ed. by G. Borjas and R. Freeman, Chicago: University of Chicago Press, 213–244.
- (1997): "How Much Do Immigration and Trade Affect Labor Market Outcomes?" *Brookings Papers on Economic Activity: Macroeconomics*.

- BORJAS, G. AND L. HILTON (1996): "Immigration and the Welfare State: Immigrant Participation in Means-Tested Entitlement Programs," *Quarterly Journal of Economics*, 111, 575–604.
- BRATSBERG, B., E. BARTH, AND O. RAAUM (2006): "Local Unemployment and the Relative Wages of Immigrants: Evidence from the Current Population Surveys," *Review of Economics and Statistics*, 88, 243–263.
- CARD, D. (1990): "The Impact of the Mariel Boatlift on the Miami Labor Market," *Industrial and Labor Relations Review*, 43, 245–257.
- (2001): "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 19, 22–64.
- (2005): "Is the New Immigration Really so Bad?" *Economic Journal*, 115, 300–323.
- CHISWICK, B. (1978): "The Effect of Americanization on the Earnings of Foreign-born Men," *Journal of Political Economy*, 86, 897–921.
- CORTES, P. (2008): "The Effect of Low-Skilled Immigration on US Prices: Evidence from CPI Data," *Journal of Political Economy*, 116, 381–422.
- CROSSLEY, T., J. McDONALD, AND C. WORSWICK (2001): "Immigrant Benefit Receipt Revisited: Sensitivity to the Choice of Survey Years and Model Specification," *Journal of Human Resources*, 36, 379–397.
- DAVIS, D. AND D. WEINSTEIN (2002): "Bones, Bombs, and Break Points: The Geography of Economic Activity," *American Economic Review*, 92, 1269–1289.
- DUFLO, E., R. GLENNERSTER, AND M. KREMER (2008): "Using Randomization in Development Economics Research: A Toolkit," in *Handbook of Development Economics*, ed. by T. P. Schultz and J. A. Strauss, Elsevier, vol. 4, chap. 61, 3895–3962.
- DUSTMANN, C. AND A. GLITZ (2005): *Immigration, Jobs and Wages*, London: Centre for Economic Policy Research.
- DUSTMANN, C. AND I. PRESTON (2007): "Racial and Economic Factors in Attitudes to Immigration," *B.E. Journal of Economic Analysis and Policy (Advances)*, 7.
- EDIN, P., R. LALONDE, AND O. ÅSLUND (2000): "Emigration of Immigrants and Measures of Immigrant Assimilation: Evidence from Sweden," *Swedish Economic Policy Review*, 7, 163–204.

- FUJITA, M., P. KRUGMAN, AND A. VENABLES (1999): *The Spatial Economy: Cities, Regions, and International Trade*, Cambridge, Massachusetts: MIT Press.
- HANSON, G. (2005): *Why Does Immigration Divide America?: Public Finance and Political Opposition to Open Borders*, Washington, DC: Peterson Institute.
- HATTON, T. AND J. WILLIAMSON (2006): *Migration and the World Economy: Two Centuries for Policy and Performance*, Cambridge, Massachusetts: MIT Press.
- HEAD, K. AND T. MAYER (2004): "The empirics of agglomeration and trade," in *Handbook of Regional and Urban Economics*, ed. by J. V. Henderson and J. F. Thisse, Elsevier, vol. 4, chap. 59, 2609–2669.
- HELPMAN, E. AND P. KRUGMAN (1985): *Market Structure and Foreign Trade*, Cambridge, Massachusetts: MIT Press.
- HOLLAND, P. (1986): "Statistics and causal inference," *Journal of the American Statistical Association*, 81, 945–960.
- KATZ, L. AND K. MURPHY (1992): "Changes in Relative Wages, 1963–87: Supply and Demand Factors," *Quarterly Journal of Economics*, 107, 35–78.
- KEANE, M. (2006): "Structural vs. Atheoretic Approaches to Econometrics," *Forthcoming in Journal of Econometrics*.
- KRUGMAN, P. (1979): "Increasing Returns, Monopolistic Competition, and International Trade," *Journal of International Economics*, 9, 469–479.
- (1991): "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99, 483.
- LACH, S. (2007): "Immigration and Prices," *Journal of Political Economy*, 115, 548–587.
- LEWIS, E. (2003): "Local, Open Economies Within the US: How Do Industries Respond to Immigration?" Working Paper 04-01. Federal Reserve Bank of Philadelphia.
- (2005): "Immigration, skill mix, and the choice of technique," Working Paper 05-08. Federal Reserve Bank of Philadelphia.
- LUBOTSKY, D. (2007): "Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings," *Journal of Political Economy*, 115, 820–867.
- MARSHALL, A. (1890): *Principles of Economics*, London: Macmillan.
- MUNDELL, R. (1957): "International Trade and Factor Mobility," *American Economic Review*, 47, 321–335.

- PAKES, A. (2008): "Theory and Empirical Work on Imperfectly Competitive Markets," NBER Working Paper No. W14117.
- RYBCZYNSKI, T. (1955): "Factor Endowments and Relative Commodity Prices," *Economica*, 22, 336–341.
- SCHEVE, K. AND M. SLAUGHTER (2001): "Labor Market Competition and Individual Preferences Over Immigration Policy," *Review of Economics and Statistics*, 83, 133–145.
- SJAASTAD, L. (1962): "The Costs and Returns of Human Migration," *Journal of Political Economy*, 70, 80–93.
- SMITH, J. AND B. EDMONSTON (1997): *The New Americans. Economic, Demographic, and Fiscal Effects of Immigration*, Washington, DC: National Academy Press.
- STORESLETTEN, K. (2000): "Sustaining Fiscal Policy through Immigration," *Journal of Political Economy*, 108, 300–323.
- TODD, P. AND K. WOLPIN (2006): "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility," *American Economic Review*, 96, 1384–1417.

Temporary Shocks and the Geography of Economic Activity: Evidence from a Large-Scale Resettlement Policy

Abstract.¹ We study the stability of regional population structure on exogenous temporary shocks. The analysis exploits a quasi-experimental setting caused by the World War II, which led Finland to relocate 11% of its population. We show that the population distribution was not stationary nor followed a random walk. Instead, the results suggest that a temporary shock set off a self-reinforcing process, where municipalities receiving the largest flows of displaced persons grew faster also in later periods. These findings are robust to using two distinct sources of exogenous variation and to controlling for previous growth trend, industry structure and geographical indicators.

JEL Classification: R12, J11, F12, N94

Keywords: Economic geography, regional growth, instrumental variables

2.1. Introduction

The relative importance of first and second nature on determining the spatial distribution of economic activity is a subject of intense debate among economists and geographers. Traditional economic models are based on the assumption that first nature – factor endowments, natural resources, landscape and the like – determine the location of production in space. Accordingly, there would be only one possible steady-state of regional structure.

While ideas challenging this prediction date back to at least Marshall (1890), formal economic models focusing on second nature – the distance between economic agents – did not occur until relatively recently. However, since Krugman’s (1991) influential work, a large theoretical literature on

¹This essay is joint work with Aki Kangasharju.

new economic geography (NEG) has provided formal models implying that several alternative steady-states of regional structure might exist (see Fujita et al., 1999 and Baldwin et al., 2003 for reviews). This class of models predicts that initial conditions, history of shocks and expectations determine which of these steady-states takes place.

For policy, the importance of second nature makes a fundamental difference. In the case of unique steady-state, regional policies are merely income transfers from one area to another. In the case of multiple steady-states, a sufficiently large policy intervention could, in principle, permanently change the regional structure. Thus it would be important to understand the real world relevance of the second nature.

Unfortunately, empirically testing the models has proved a daunting task (see Overman et al., 2003 and Head and Mayer, 2004 for reviews). The challenge is that the theoretical predictions rise from circular causation: the intuition is that it makes sense for everyone to locate where everyone else are. Testing for the existence of such circular causation is difficult. To overcome this problem, a branch of recent literature has focused on testing one implication of the models – that a temporary shock could have permanent effects – by exploiting variation arising from quasi-experimental settings.

Interestingly, this approach has produced results pointing towards the dominance of first nature. In particular, Davis and Weinstein (2002) show that even the massive shock created by the Allied bombings in Japan did not have long-term effects on the country's city size distribution. Brakman et al. (2004) and Miguel and Roland (2006) reach similar conclusions in their analysis of the effects of the Allied bombings in Germany and the U.S. bombings in Vietnam, respectively. These findings are reinforced by Davis and Weinstein (2008), who show that in the aftermath of the World War II, the specific industries in Japan tended to relocate back to their original sites.

This paper focuses on a different kind of natural experiment. That is, we study the impact of large-scale population displacement on rural municipality size distribution over the next half a century. These displacements took place after the World War II as Finland ceded a tenth of its territory to the Soviet Union and evacuated the entire population living in these areas. In total 430,000 persons – approximately 11 percent of the population – were resettled to the remaining parts of the country. As a consequence, the

population share of the displaced was up to 36 percent in some municipalities, while no-one was settled to others.

We use an instrumental variables approach exploiting features of the allocation policy to identify the causal effect of these migrations on the long-term spatial distribution of population. The instruments are constructed using the fact that latitude, the amount of government owned land, distribution of privately owned land and the share of Swedish-speaking population in the municipality had a strong effect on the number of displaced persons each municipality received. All these factors are plausibly exogenous to spurious shocks that took place after the World War II.

Our results are in stark contrast with the previous quasi-experimental evidence. We find that the resettlement had a permanent impact. Furthermore, the estimates suggest that it set off a self-reinforcing process, where municipalities experiencing large exogenous migration inflows during the war and its immediate aftermath also experienced a more positive population growth trend during the post-war period. The findings are robust to using different instruments and to controlling for previous growth trend, industry structure and geographical indicators. To the best of our knowledge, these are the first quasi-experimental results consistent with theories predicting "spatial catastrophes".²

The paper is organized as follows. Section 2.2 discusses the resettlement in detail. Section 2.3 presents our data and Section 2.4 introduces the empirical approach. Section 2.5 presents the results. Section 2.6 concludes.

2.2. Post-WWII Resettlement in Finland

During the World War II, Finland fought twice against the Soviet Union. As part of the peace treaties, Finland ceded three distinct land areas, totaling over tenth of its territory, and evacuated the entire population living in these areas. The evacuation created approximately 430,000 displaced persons corresponding to 11 percent of the total population. The most populous part of ceded land was located in the South-Eastern Finland, in a region of

²Previously Bosker et al. (2006, 2007) have presented some evidence on permanent change in the distribution of population across West German cities after the World War II. Furthermore, Redding et al. (2007) find that the division of Germany moved the German air hub from Berlin to Frankfurt, but it did not return to Berlin after the re-unification. However, while these papers show that temporary shocks may have permanent effects, they do not provide evidence of self-reinforcing processes.

Karelia. Two other parts were located in the extremely sparsely populated Lapland in Northern Finland.

The plan of resettling the evacuated population was designed in three pieces of legislation: Rapid Resettlement Act, Land Acquisition Act and Settlement Plan. According to these plans, those who had derived their principal income from agriculture in the ceded areas were entitled to receive cultivable land in the remaining parts of country. As more than half of the labor force was working in agriculture, this decision had a major effect on the allocation of displaced persons. The plans provided land to displaced farmers, landowners, war invalids, war widows and war orphans. Also, some non-displaced groups, such as landless ex-servicemen with families, were entitled to receive land, typically from their municipality of residence.

The aim was to resettle the displaced farmers to areas resembling their former location in terms of climate and soil. As a consequence, the location of the provided land depended on the pre-war municipality and the displaced farmers were not able to choose their destination. In contrast, non-agrarian displaced persons received compensation from their lost property in the form of government bonds, and were free to settle wherever they could find accommodation and employment.

The eligible agrarian population received three different kinds of holdings. Properties of at most 15 hectares of cultivable land were primarily intended for the displaced farmers. Front-line soldiers, war invalids and war widows typically received smaller homesteads and building lots. There was also an option of a "cold farm". These were located in Eastern and Northern Finland and had no cultivated land or buildings. Thus the receiver had to clear the land and found the farm by himself. Not surprisingly, displaced farmers were reluctant to opt for cold farms in spite of special awards and payment arrangements offered by the state.

In total, 245,724 hectares of already cultivated land was used for resettlement and 149,675 hectares was cleared for cultivation (Laitinen, 1995). The land was first taken from the state, municipalities, business corporations, church, other public bodies, land speculators and landowners not practicing farming. However, "secondary sources" – private landowners who lived in their farms – ended up providing roughly half of the cultivated fields, a quarter of the land that could have been cultivated and a fifth of forest land. The land was purchased either on voluntary basis or through expropriation

TABLE 2.2.1. Scale for Expropriation of Privately Owned Land

Size (hectares)	Expropriation Percentage	
	At the Lower Limit (2)	Part Exceeding the Lower Limit (3)
25–35	10	45
35–50	20	53
50–100	30	60
100–200	45	75
200–400	60	80
400–800	70	90
over 800	80	90

Note: The scale for land expropriation for private land owners. Set by Resolution of the Council of State in June 1945 and amended in July 1946. The size of the farm was determined on a basis of the total area of cultivated land, cultivable meadow and open pasture land. Farmers with two or more dependent children received some exemptions. Source: Pihkala (1952, Table II).

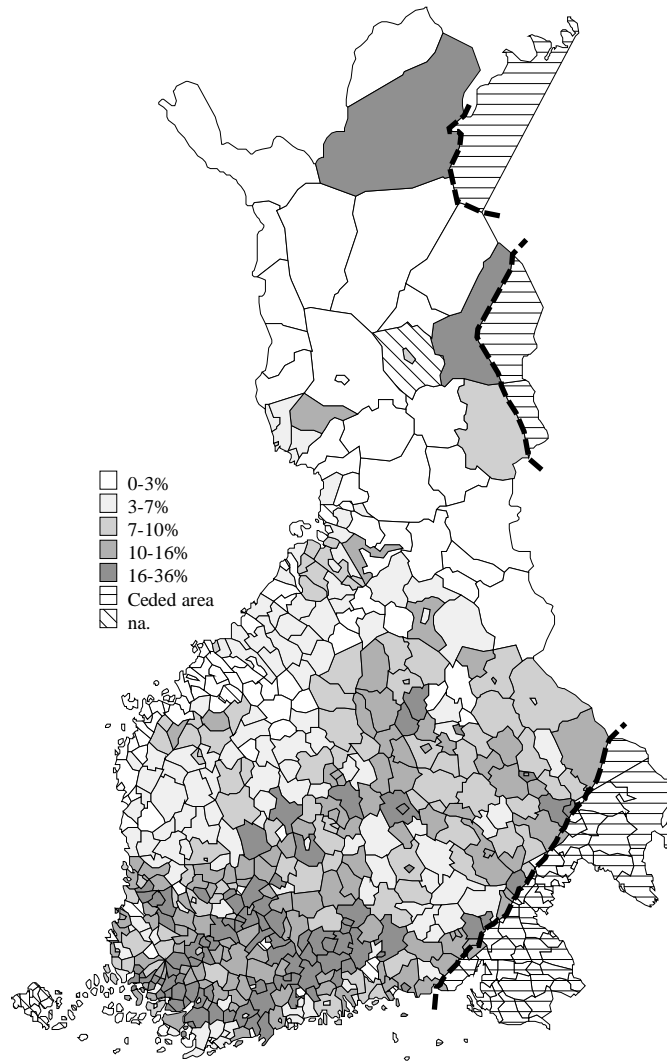
using a progressive scale presented in Table 2.2.1. Landowners were paid a “justifiable current local price” for the expropriated land.³

As a result, the amount of land available for displaced farmers within the borders of a given municipality – and hence the amount of displaced farmers allocated to the municipality – was primarily determined by the pre-war farm size distribution and the amount of land owned by the public sector. Two other factors created variation in the inflow of displaced persons. First, no-one was settled to Northern Finland, where the conditions for agriculture are the least favorable. Second, the Land Acquisition Act included a clause demanding that the resettlement should not alter the linguistic composition of municipalities. Since the vast majority of the displaced farmers spoke Finnish as their mother tongue, very few received land from the Swedish-speaking municipalities on the Western and Southern coasts.

While adjacent neighbors of ceded areas in Lapland received the most of those few displaced in Lapland, the Karelian displaced were more evenly distributed across Middle and Southern parts of remaining Finland. Figure 2.2.1 present the share of the displaced on municipalities’ population in 1948,

³In practice, the price was fixed by the Ministry of Finance. These compensations were paid by government bonds with 15 years to maturity and yielding 4% nominal interest. Eventually inflation wiped out four fifths of their value. However, the bonds could be used for paying the Second Property Expatriation Tax, which was collected from all capital owners in order to cover the costs of the resettlement. According to Pihkala (1952, 153), landowners probably did not suffer more than other owners of property.

FIGURE 2.2.1. Share of Displaced Persons on Municipalities' Population in 1948



which ranges between zero and 36 percent. In particular, the proportion of displaced persons in many Swedish speaking municipalities on the Western Coast is markedly low. Importantly, the settlement plan appears to have a large effect on the location of the displaced. According to Waris et al. (1952), 53 percent of all displaced persons lived in their designated placement areas in June 1949. Given that roughly 35 percent were entitled to receive land, this suggests that many followed the placement plan regardless, perhaps due

to family ties and employment opportunities with their former landowner employers.

2.3. Data

We use data at the level of 425 municipalities drawn from the National Archives, National Land Survey of Finland and from various Statistical Yearbooks, Population Censuses and Agricultural Censuses published by Statistics Finland.⁴ The estimation sample is restricted to 215 rural municipalities that were subject to the 1945 Resettlement Plan. This sample selection rule is motivated in two ways. First, since the exogenous variation essential for our identification strategy discussed below is created by the Resettlement Plan, it seems reasonable to focus on the area affected by the plan. However, this restriction comes with the caveat of excluding some Swedish-speaking municipalities that could have been comparable to municipalities in the settlement area. Nevertheless, part of the relevant Swedish-speaking municipalities is still included due to the fact that they were subject to the plan and received a small number of Swedish speaking displaced persons. Cities and market towns are excluded from the data, as the non-agrarian population was free to choose their location. This exclusion also allows us to focus on a relatively homogeneous group of agriculturally oriented rural municipalities. In order to test for the sensitivity of our results, Section 5 presents the results using an alternative sample selection criterion.

Figure 2.3.1 provides the first look at the data. It plots the population growth rates of the municipalities between 1948 and 2000 on the growth rates between 1939 and 1948. The figure provides three lessons. First, some municipalities experienced very large changes in their populations. Second, there is a positive association between war time growth rates and later growth rates.⁵ Third, while 95 percent of the municipalities grew during the war and its immediate aftermath, 79 percent of the sample experienced decline

⁴The precise data sources are the following: Total population, share of Swedish speaking population, taxable income per capita: *Statistical Yearbooks* (Statistics Finland, several years); Pre-war industry structure: Statistics Finland (1979): *Väestön elinkeino 1880-1975*; Amount and distribution of agricultural land: *Agriculture Census 1930 and 1941* (Statistics Finland); Number of displaced persons in municipalities: SM/Siirtoväenasiainosasto/ Kansiot H1, H1a, H2, H3, H4, H5, H6, H7, H8, H9 (National Archives). Longitude and latitude: National Land Survey of Finland. We deal with the fact that several municipalities have merged since the war by aggregating the data to correspond to the municipality borders of year 2000 throughout the analysis period.

⁵Regressing the growth rate between 1948 and 2000 on the growth rate between 1939 and 1948 yields a coefficient of 0.92 (standard error 0.20). The coefficient is 0.86 (standard

FIGURE 2.3.1. Growth Rates between 1939–1948 and 1948–2000

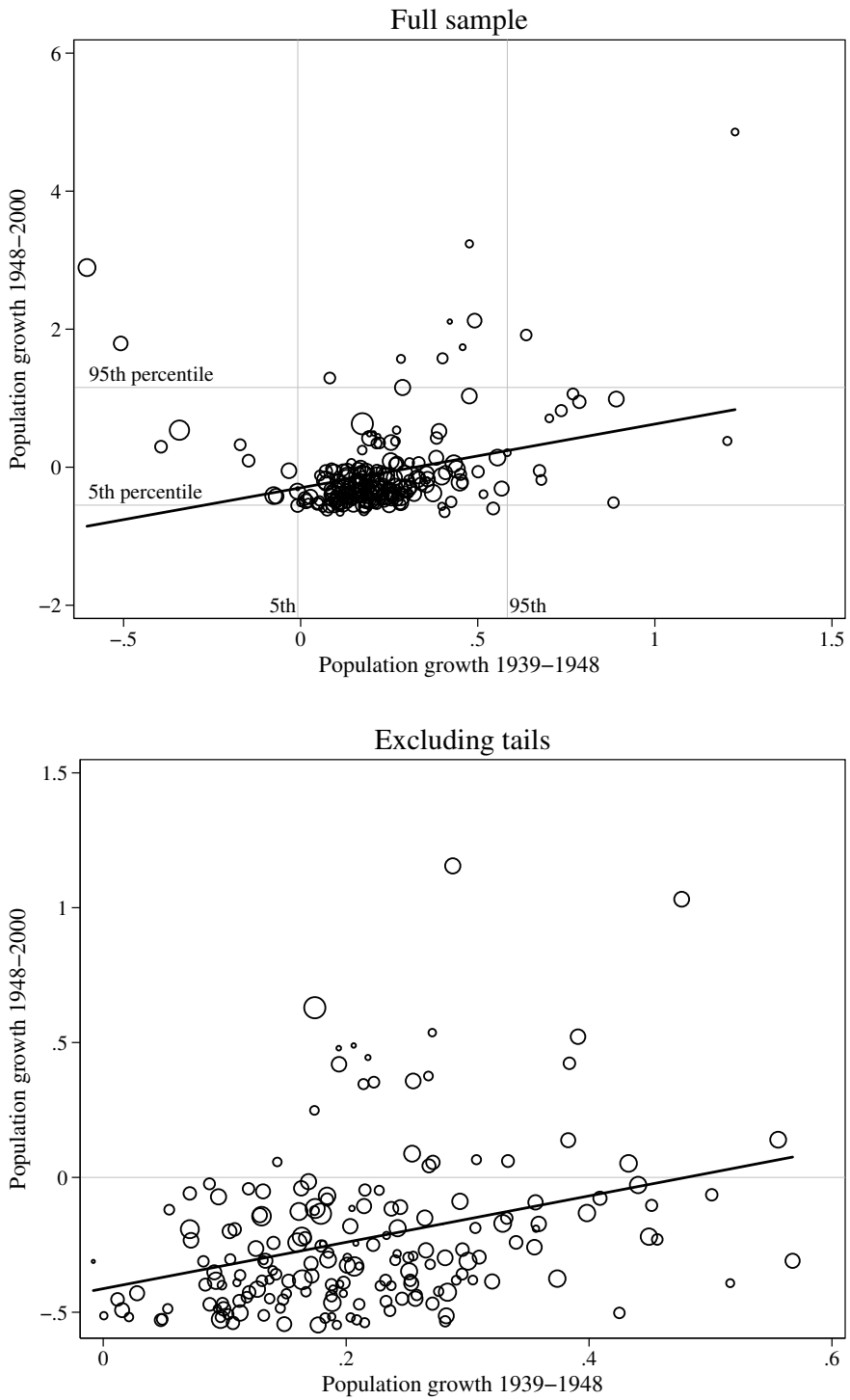
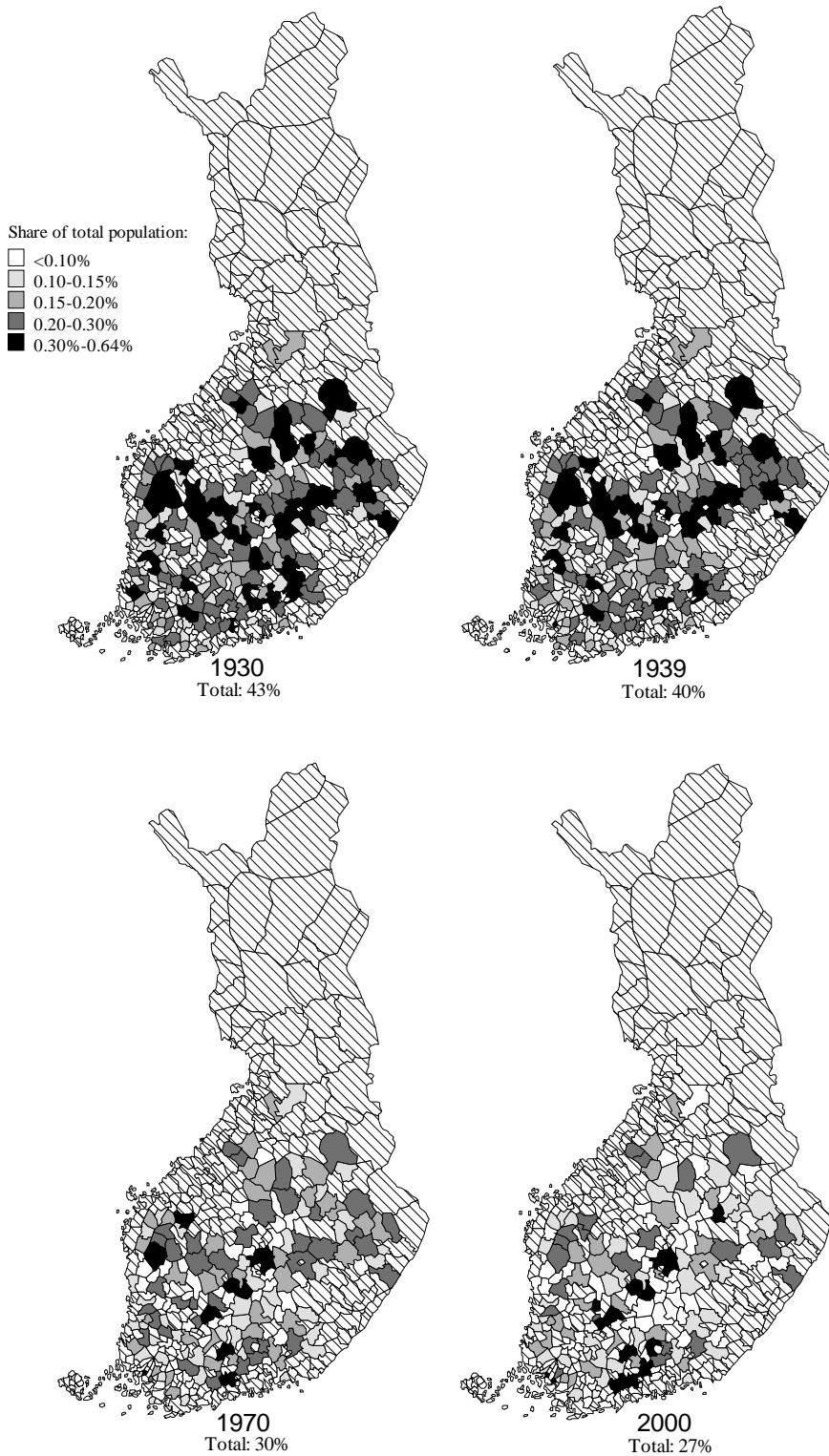


FIGURE 2.3.2. Municipalities Share of the Total Population
1930–2000



in their population between 1948 and 2000. This decline was driven by emigration and, more importantly, urbanization. In particular, there were large migration flows from rural to urban areas during the late 1960s, early 1970s and late 1990s. This process was already taking place before the war, as illustrated by Figure 2.3.2. In total, the share of population living in the sample municipalities declined from 40 percent just before the war to 27 percent in 2000.

2.4. Empirical Model

Above we saw that wartime population growth was positively correlated with later population growth. However, this association could be a result of confounding factors affecting both the wartime and the post-war population growth. We next set up an empirical framework – closely following the approach taken by Davis and Weinstein (2002) – which allows us to assess the causal effect of temporary shocks.

Let s_{it} be the natural logarithm of region i 's share of total population at time t . Suppose that the relative size of the region is a function of its initial size Ω_i and region- and time-specific shocks ε_{it} . The size of the region i at time t is

$$(2.1) \quad s_{it} = \Omega_i + \varepsilon_{it}$$

We model the persistence of shocks on population shares as

$$(2.2) \quad \varepsilon_{it+1} = \rho\varepsilon_{it} + \nu_{it+1}$$

where ν_{it+1} is the independently and identically distributed shock and parameter ρ determines the persistence of shocks. We test three hypothesis. First, if $0 < \rho < 1$, region's share of population is stationary and shocks will dissipate over time. Second, if $\rho = 1$, regional population growth follows a random walk and all shocks are permanent. Finally, if $\rho > 1$, a temporary positive shock leads to faster growth in later periods.

The last possibility is consistent with theoretical models where circular causation may lead to "spatial catastrophes". Note, however, that we are unaware of any theoretical model implying the form of equation (2) in the case of $\rho > 1$. That is, the NEG models, such as Krugman (1991), typically predict that a change of a steady-state would be a highly non-linear process.

error 0.19), when we exclude observations below the 5th or above the 95th percentile in either the wartime or post-war growth distributions.

Once a threshold in the parameter space has been crossed, the economy moves to a new steady-state. The shock should have no effect after this shift has occurred. Thus estimates implying that ρ is larger than unity would, in our view, suggest that the empirical model is likely to be misspecified and estimates of ρ would be informative mostly in the sense of rejecting both the stationarity and random walk hypothesis.

In order to estimate the persistence of temporary shocks, we take first differences of (1) and use (2) to get

$$(2.3) \quad s_{it+1} - s_{it} = (\rho - 1) \nu_{it} + [\nu_{it+1} + \rho(\rho - 1) \varepsilon_{it-1}]$$

where the growth in region's share of total population depends on the past and current shocks. We use growth in region's i share of total population during the war as a proxy for ν_{it} . The estimation equation is thus

$$(2.4) \quad s_{i1948+t} - s_{i1948} = \alpha [s_{i1948} - s_{i1939}] + \beta_0 + X_i \delta + error_i$$

where $\alpha = (\rho - 1)$, β_0 is a constant and X_i is a vector of control variables one might want to include. Clearly, war time population growth was affected by many other factors than the settlement policy. This can be considered as a measurement error problem, which – given a valid instrument – can be solved with an instrumental variables approach.⁶

The essence of our identification strategy is that flows of the displaced are an observable shock. That is, we will exploit the Settlement Plan determining the size of this shock as an instrument for war time population growth. However, there are two potential concerns. First, it is conceivable that past shocks to population growth may have affected the plan. For example, the government officials could have preferred evenly distributed population and hence more displaced persons might have been targeted to areas with lower pre-war growth rates. On the other hand, the displaced may have preferred to move to fast growing areas and thus might have lobbied for such allocation. The second potential problem is the possible correlation between the Settlement Plan and future shocks. For instance, if the Settlement Plan was affected by expectations on future population growth or by factors affecting this growth (e.g. industry structure), the estimates would be biased.

To address these concerns, we use only two components of the Settlement Plan. The first instrument is the share of Swedish-speaking population in

⁶Note that there is also an endogeneity problem since ε_{it-1} is a part of the error term in equation (3). Hence $s_{it} - s_{it-1}$ is correlated with the error term by construction. Instrumental variables approach solves this problem as well.

the municipality in 1930. The second instrument is the per capita amount of agricultural land available to be distributed to the displaced farmers, calculated as

$$(2.5) \quad I_{i39} = \frac{G_{i39} + \sum_{s=1}^n (\tau_l^s h_l^s + \tau_m^s h_m^s) N_{i39}^s}{P_{i39}}$$

where G_{i39} is the pre-war amount of government owned agricultural land in municipality i , τ_l^s is the expropriation rate at the lower limit of the size class s (column 2 in Table 2.2.1), τ_m^s is the expropriation rate for the part exceeding the lower limit in this bracket (column 3 in Table 2.2.1), h_l^s is the bracket's lower limit in hectares, h_m^s is the midpoint of the exceeding part, N_{i39}^s is the pre-war number of farms in the municipality belonging to the bracket, and P_{i39} is the pre-war population of the municipality.

The identifying variation we use is thus solely determined by the amount of land owned by the public sector, the distribution of privately owned land and the lingual composition of the municipalities, all measured before the war. These are outcomes of long historical processes and hence unlikely to be correlated with spurious shocks affecting post-WWII population growth. To further safeguard against such spurious correlations, we also report results from specifications controlling for observable pre-war characteristics of the municipalities and from specification where we use only one of the instruments and add the other as a control.

2.5. Results

We begin by addressing the plausibility of our instruments by regressing pre-war population changes on the share of Swedish-speaking population in 1930 and an “quasi-instrument”. The latter is constructed as in equation (2.5), but using the 1930 distribution of privately owned land. The motivation for this falsification test is that the validity of our approach relies on the assumption that linguistic composition and the distribution of land affects population growth *only* through the Settlement Plan. A finding that these factors were associated with pre-war population growth would raise concerns on the validity of this identifying assumption. However, results presented in Table 2.5.1 illustrate that such association did not exist. Thus the available data does not suggest that our instruments would be invalid.

We next examine the strength of the instruments to predict wartime population growth. Table 2.5.2 reports the first-stage results in the three

TABLE 2.5.1. Quasi-Instruments and Pre-War Change in log Population Shares

	(1)	(2)	(3)
“Land instrument” using 1930 distribution	0.00 (0.03)	0.01 (0.04)	-0.11 (0.05)
Share of Swedish-speaking population (1930)	-0.01 (0.04)	0.04 (0.05)	0.02 (0.04)
<i>Controlling for:</i>			
Municipality Characteristic	no	yes	yes
Geography	no	no	yes
Robust F-statistics	0.0	0.4	2.2
Partial R ²	0.00	0.00	0.01

Note: The impact of the instruments on the change in log population share between 1930 and 1939, and robust standard errors (in parenthesis). Municipality characteristics: share of labor force in agriculture 1930, share of population in manufacturing 1930. Geography: indicator for being a neighbor of a city or market town, longitude, latitude, longitude*latitude, distance from Helsinki, population density in 1930.

specifications used in the analysis below. As expected, larger stock of available agricultural land in 1939 is positively associated with the change in log population shares between 1948 and 1939. Similarly, municipalities with larger share of Swedish-speaking population in 1930 received fewer displaced persons and thus grew less. Together the instruments explain roughly 15 percent of the variance in the change of population shares during this period. Furthermore, the robust F-statistics for the excluded instruments vary between 17 and 26 across the specifications, implying that our inference does not suffer from problems related to weak instruments.

Table 2.5.3 presents the main results. Consider first panel A, where the first column reports the estimate from regressing the change in log population shares between 1948 and 1950 on the change in log population shares between 1939 and 1948, using available agricultural land and the pre-war share of Swedish-speaking population as instruments. This yields a point estimate of -0.01 for α . Recall that $\alpha = (\rho - 1)$. Thus a coefficient of zero corresponds to random walk, indicating that the population shock due to the resettlement did not affect growth rates during the first two years after its implementation was completed.

In the subsequent columns we gradually extend the period examined. For instance, the last column reports estimates from regressing the change in log

TABLE 2.5.2. First-Stage

	(1)	(2)	(3)
Available cultivated land per capita	0.28 (0.06)	0.21 (0.09)	0.38 (0.08)
Share of Swedish-speaking population (1930)	-0.33 (0.09)	-0.40 (0.06)	-0.36 (0.07)
<i>Controlling for:</i>			
Pre-War Municipality Characteristic	no	yes	yes
Geography	no	no	yes
Robust F-statistics	17.1	24.6	27.8
Partial R ²	0.13	0.15	0.19

Note: The impact of the instruments on the change in log population share between 1939 and 1948, and robust standard errors (in parenthesis). Pre-war municipality characteristics: change in log population share between 1930 and 1938, mean taxable income per capita between 1936 and 1939, share of labor force in agriculture 1939, share of population in manufacturing 1939. Geography: indicator for being a neighbor of a city or market town, longitude, latitude, longitude*latitude, distance from Helsinki, population density in 1938.

population shares between 1948 and 2000 on the change in log population shares between 1939 and 1948 using the same instruments as above. The point estimate now suggests that $\rho = 2.33$, implying that the impact of a war-time shock would more than double during the next half of century. Furthermore, we reject the null hypothesis of the regional population growth being stationary or following random walk.

There are at least two possible explanations for these results. First, they are consistent with the idea that a temporary shock on region's population share set off a self-reinforcing process. Interestingly, the estimates grow in almost a linear fashion as we extend the study period. This suggests that if we are truly witnessing a "spatial catastrophe", the shift from one steady-state to another was a rather slow process.

The alternative explanation is that municipalities with the most favorable fundamentals with respect to post-WWII population growth would have had more available land or larger Swedish-speaking populations. If this was the case, our instruments would be correlated with these (omitted) fundamentals and thus the estimates would be biased. We investigate this possibility in several ways. First, we begin by studying whether controlling for pre-war municipality characteristics affects the results. Panel B reports the estimates

TABLE 2.5.3. Impact of the Resettlement

	Dependent variable:					
	Change in log pop. share between 1948 and 1950					
	1960	1970	1980	1990	2000	
<i>A: No controls</i>						
	-0.01	0.05	0.29	0.57	0.93	1.33
	(0.05)	(0.18)	(0.26)	(0.38)	(0.49)	(0.56)
<i>B: Controlling for Pre-War Municipality Characteristics</i>						
	-0.03	0.25	0.45	0.67	0.90	1.17
	(0.06)	(0.20)	(0.27)	(0.39)	(0.48)	(0.55)
<i>C: ...and Geography</i>						
	-0.02	0.19	0.35	0.51	0.60	0.79
	(0.05)	(0.16)	(0.23)	(0.34)	(0.41)	(0.47)

Note: 2SLS estimates, robust standard errors (in parenthesis). Instruments: Agricultural land available for distribution per capita, share of Swedish-speaking population in 1930. Pre-war municipality characteristics: change in log population share between 1930 and 1938, mean taxable income per capita between 1936 and 1939, share of labor force in agriculture 1939, share of population in manufacturing 1939. Geography: indicator for being a neighbor of a city or market town, longitude, latitude, longitude*latitude, distance from Helsinki, population density in 1938.

in a specification controlling for previous growth trend (change in log population share between 1930 and 1938), taxable income per capita and labor force shares of agriculture and manufacturing, all measured before the war. The results are virtually unaffected. Panel C reports the estimates when we control for the same municipality characteristics as above and further add an indicator for the municipality neighboring a city or market town, longitude and latitude and their interaction, distance from Helsinki and population density in 1938. While the point estimates are now somewhat smaller, we still strongly reject the null of the population growth being stationary. However, we are no longer able to reject the null of random walk.

Table 2.5.4 provides further robustness checks. Panel A presents the results when we use only one of the instruments and add the other as a control. Both instruments yield virtually identical results. In Panel B, we extend the sample by including 13 rural municipalities that were not listed in the Settlement Plan, presumably due to having large Swedish speaking populations. The point estimates now become somewhat sensitive to the

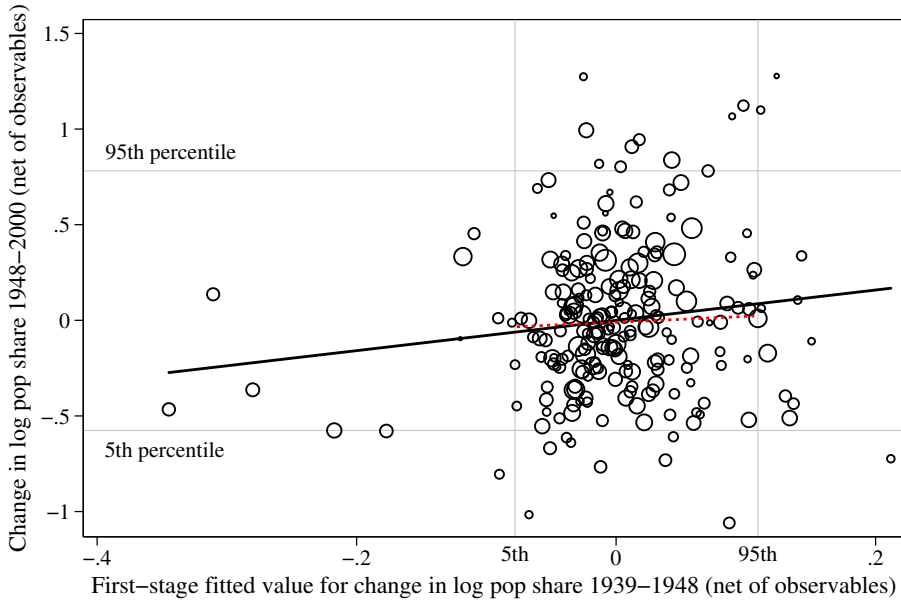
TABLE 2.5.4. Robustness Checks

	Dependent variable: Change in log pop. share between 1948 and					
	1950	1960	1970	1980	1990	2000
<i>A: Base data</i>						
Instrument 1	-0.07 (0.08)	-0.14 (0.20)	0.09 (0.31)	0.34 (0.47)	0.35 (0.56)	0.63 (0.64)
Instrument 2	0.04 (0.07)	0.57 (0.25)	0.65 (0.37)	0.70 (0.51)	0.87 (0.64)	0.98 (0.72)
Both	-0.02 (0.05)	0.19 (0.16)	0.35 (0.23)	0.51 (0.34)	0.60 (0.41)	0.79 (0.47)
<i>B: Extended data</i>						
Instrument 1	-0.08 (0.09)	-0.09 (0.20)	0.22 (0.33)	0.49 (0.50)	0.53 (0.60)	0.85 (0.69)
Instrument 2	0.03 (0.06)	0.61 (0.19)	0.91 (0.27)	1.01 (0.36)	1.09 (0.43)	1.08 (0.47)
Both	-0.01 (0.04)	0.34 (0.12)	0.65 (0.18)	0.81 (0.26)	0.88 (0.31)	0.99 (0.35)

Note: 2SLS estimates, robust standard errors (in parenthesis). Instrument 1: Agricultural land available for distribution per capita; Instrument 2: share of Swedish-speaking population in 1930. Controlling for change in log population share between 1930 and 1938, mean taxable income per capita between 1936 and 1939, share of labor force in agriculture 1939, share of population in manufacturing 1939, indicator for being a neighbor of a city or market town, longitude, latitude, longitude*latitude, and distance from Helsinki. Extended data includes 17 Swedish-speaking rural municipalities that are likely to have been excluded from the Resettlement Plan due to large Swedish-speaking populations.

choice of the instrument. However, over-identification test based on Hansen J-statistics rejects the null of the coefficient being identical only for the regressions where the dependent variable is the change in log population share between 1948 and 1960. Note, however, that even if different instruments would yield different results, this would not necessarily imply that one or both of them would not be valid. That is, theoretical models imply that the impact of a temporary shock should vary across receiving locations. If this is the case, different instruments identify different weighted averages of local average treatment effects (Angrist and Imbens, 1995). Thus difference between the estimates in the lower panel of Table 2.5.4 are consistent with the effect being heterogeneous across municipalities.

FIGURE 2.5.1. Resettlement shock and change in log population shares 1948–2000



Note: Second-stage regression of change in log population shares between 2000 and 1948 on the first-stage fitted value of change in log population share between 1948 and 1939, netting out the trend between 1939 and 1930, mean taxable income per capita between 1936 and 1939, share of labor force in agriculture in 1939, share of population in manufacturing in 1939, indicator for being a neighbor of a city or market town, longitude, latitude, longitude*latitude and distance from Helsinki. The size of the dots corresponds to the level of the municipality's population in 1939. Solid line is the linear fit using all observations. Dotted line is the linear fit omitting observations below the 5th or above the 95th percentile in either the first-stage fitted value or the post-war population distributions.

These results provide strong evidence on that a temporary shock led to a long-term change in the regional structure. We next ask, what drives these results. The first possibility is that some municipalities would have experienced extreme population growth or decline. However, this does not seem to be the case. Figure 2.5.1 plots the change in log population shares between 1948 and 2000 on the first-stage fitted value of the change in log population share between 1939 and 1948, once the impact of pre-war municipality characteristics and geographical indicators are netted out. In other words, this is the second-stage leading to the estimate presented in the last column of panel C in Table 2.5.3. The figure does not suggest that outliers would

TABLE 2.5.5. Pre-war Municipality Characteristics and the Impact of the Shock

Dependent variable:						
	Change in log pop. share between 1948 and					
	1950	1960	1970	1980	1990	2000
<i>A: Pre-War Population</i>						
Baseline	0.00 (0.06)	0.20 (0.21)	0.22 (0.35)	0.32 (0.56)	0.19 (0.65)	0.35 (0.77)
Interaction	-0.01 (0.13)	-0.08 (0.34)	0.16 (0.50)	0.28 (0.77)	0.65 (0.92)	0.70 (1.07)
<i>B: Pre-War Taxable Income per capita</i>						
Baseline	0.01 (0.09)	-0.11 (0.47)	0.30 (0.74)	0.90 (1.06)	-0.28 (0.81)	-0.20 (0.90)
Interaction	-0.07 (0.11)	0.35 (0.52)	0.09 (0.80)	-0.44 (1.13)	0.94 (0.91)	1.06 (1.02)
<i>C: Pre-War Labor Force in Agriculture</i>						
Baseline	0.06 (0.06)	0.40 (0.23)	0.60 (0.32)	0.80 (0.49)	1.01 (0.61)	1.26 (0.70)
Interaction	-0.17 (0.11)	-0.54 (0.30)	-0.78 (0.38)	-0.97 (0.61)	-1.37 (0.72)	-1.61 (0.82)

Note: 2SLS estimates, robust standard errors (in parenthesis). Instruments: Agricultural land available for distribution per capita, share of Swedish-speaking population in 1930, both interacted with with an indicator for the municipality having above median level of pre-war population (panel A) or above median taxable income per capita (panel B) or above median labor force share in agriculture (panel C). Controlling for pre-war trend (change in log population share between 1930 and 1938), mean taxable income per capita between 1936 and 1939, share of labor force in agriculture 1939, share of population in manufacturing 1939, an indicator for being a neighbor of a city or market town, longitude, latitude, longitude*latitude, distance from Helsinki, population density in 1938, and the individual effect of the interacted terms.

be driving the results. In fact, when we exclude observations that are in the tails of either the first-stage fitted value distribution or the post-WWII growth distribution, we get similar results (dotted line) as when using the full sample (solid line).

We next ask whether municipalities with different pre-war initial conditions reacted differently to the shocks. Table 2.5.5 reports the results from regressions where we interact the change in log population share between 1939 and 1948 with indicators of the pre-war municipality characteristics.

In panel A, the interaction term is an indicator variable taking value one if the municipality had a pre-war population above the median and zero otherwise.⁷ None of the interaction terms is significantly different from zero. Similarly, results in panel B do not provide evidence that municipalities with above median taxable income per capita would have reacted differently to the shock than poorer municipalities.

However, estimates presented in panel C of Table 2.5.5 suggest that the impact of the flows of displaced persons was substantially larger for those rural municipalities, which were less agrarian. Taken at face value, the point estimates imply that the size distribution of municipalities with above median share of agricultural labor force before the war followed a random walk. In contrast, the point estimates for municipalities where a larger share of the labor force worked in non-agrarian occupations suggest that the temporary shock set of a self-reinforcing process.⁸ We note that one should not read too much into the point estimates as the standard errors are rather large. Yet, the interaction terms from 1960 onwards are statistically significant.

2.6. Conclusions

In this paper, we have provided quasi-experimental evidence on the importance of second nature in determining the distribution of economic activity in space. More precisely, we have shown that resettling 11% of the Finnish population during and after the World War II had a permanent impact on spatial population distribution. Furthermore, the results point towards a possibility that these shocks set of a self-reinforcing process where rural municipalities receiving large exogenous flows of displaced persons grew faster or declined slower also in the next half a century. These findings are in line with the predictions of a large class of theoretical models. However, most previous quasi-experimental empirical work has reached the opposite conclusions.

⁷We chose to interact with dummies rather than with levels due to the level interactions leading to some extreme values. For instance, the largest value from interacting the change in log population share between 1939 and 1948 with the share of labor force working in agriculture in 1930 was almost 12 times larger than the third largest value. As a consequence, the estimates become too imprecise to be informative. However, point estimates provide similar patterns as those in Table 2.5.5. Furthermore, if we exclude observations with extreme values of the interaction terms, the results are similar also in terms of statistical significance.

⁸The labor force share of agriculture in 1930 varies between 23 and 96 percent with a median of 82 percent. The means below and above the median is 70 percent and 87 percent, respectively.

There are several reasons why we might expect the shock examined here to be more persistent than those studied previously. Most importantly, previous literature has focused on the impact of war-time bombings. After such shocks, people may return home due to location specific preferences, partial survival of the infrastructure and the remaining legal incentives such as property rights. In the case we study, the land was ceded to the Soviet Union and return was impossible. Furthermore, Finland is a sparsely populated country with many alternative locations for economic activity.

Perhaps the most striking feature of our results is that, unlike previous studies, we focus on rural areas. Second nature should be far less important for the municipalities in our data than for cities. Interestingly, the results suggest that this reasoning is valid also within rural areas. That is, the estimated effects were substantially higher among those rural municipalities, which were less agrarian prior to the resettlement. It seems reasonable to speculate that the resettlement was more important to these locations, perhaps because large flows of displaced persons may have pushed them to modernize their production structure more rapidly and thus helping them to adapt to the rapid structural change that Finland went through in the latter part of the 20th century.

Finally, several questions are left for future research. In particular, while we are able to argue that the change in steady-state occurred due to path dependence rather than a change in expectations, we cannot distinguish whether internal economies of scale, human capital externalities or some other possible mechanism drive the results. These questions are best addressed with a different empirical approach and are thus left for another paper. Furthermore, while our findings are broadly in line with the NEG models, we are not aware of a theoretical model that would predict exactly the shape of our results. Thus we hope to inspire also further theoretical work.

References

- ANGRIST, J. AND G. IMBENS (1995): "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity." *Journal of the American Statistical Association*, 90.
- BALDWIN, R., R. FORSLID, P. MARTIN, G. OTTAVIANO, AND F. ROBERT-NICOUD (2003): *Economic Geography and Public Policy*, Princeton University Press.
- BOSKER, M., S. BRAKMAN, H. GARRETSEN, AND M. SCHRAMM (2006): "A Century of Shocks: The Evolution of the German City Size Distribution 1925-1999," CESifo Working Paper 1728.
- (2007): "Looking for multiple equilibria when geography matters: German city growth and the WWII shock," *Journal of Urban Economics*, 61, 152–169.
- BRAKMAN, S., H. GARRETSEN, AND M. SCHRAMM (2004): "Bombing of German Cities during WWII and its Impact on City Growth," *Journal of Economic Geography*, 4, 201–218.
- DAVIS, D. AND D. WEINSTEIN (2002): "Bones, Bombs, and Break Points: The Geography of Economic Activity," *American Economic Review*, 92, 1269–1289.
- (2008): "Search for Multiple Equilibria in Urban Industrial Structure," *Journal of Regional Science*, 48, 29–65.
- FUJITA, M., P. KRUGMAN, AND A. VENABLES (1999): *The Spatial Economy: Cities, Regions, and International Trade*, MIT Press.
- HEAD, K. AND T. MAYER (2004): "The empirics of agglomeration and trade," in *Handbook of Regional and Urban Economics*, ed. by J. V. Henderson and J. F. Thisse, Elsevier, vol. 4, chap. 59, 2609–2669.
- KRUGMAN, P. (1991): "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99, 483.
- LAITINEN, E. (1995): "Vuoden 1945 maanhankintalain synty," in *Rintamalla raiviolle. Sodanjälkeinen asutustoiminta 50 vuotta*, ed. by E. Laitinen, Jyväskylä: Atena Kustannus, 52–138.
- MARSHALL, A. (1890): *Principles of Economics*, London: Macmillan.
- MIGUEL, E. AND G. ROLAND (2006): "The Long Run Impact of Bombing Vietnam," NBER Working Paper No. W11954.
- OVERMAN, H., S. REDDING, AND A. VENABLES (2003): "The Economic Geography of Trade, Production, and Income: A Survey of Empirics," in

Handbook of International Trade, ed. by J. Harrigan and E. Kwan Choi, Oxford: Blackwell.

PIHKALA, K. (1952): "The Land Settlement Program of Finland," *Land Economics*, 28, 147–159.

REDDING, S. J., D. M. STURM, AND N. WOLF (2007): "History and Industry Location: Evidence from German Airports," CEPR Discussion Papers 6345.

WARIS, H., V. JYRKILÄ, K. RAITASUO, AND J. SHIPI (1952): *Siirtoväen sopeutuminen*, Otava.

CHAPTER 3

Long-Term Effects of Forced Migration

Abstract.¹ We study the long-term effects of human displacement using individual-level panel data on forced migrants and comparable non-migrants. After World War II, Finland ceded a tenth of its territory to the Soviet Union and resettled the entire population living in these areas in the remaining parts of the country. We find that displacement increased geographical and occupational mobility. Furthermore, displacement increased the long-term income of men, but had no effect on that of women. We attribute a large part of the effect to faster transition from traditional (rural) to modern (urban) occupations among the displaced.

JEL classification: J60, O15, R23

Keywords: Migration, displaced persons, regional labor markets

3.1. Introduction

Armed conflicts, natural disasters and infrastructure projects force people to migrate. While international refugee flows have dominated the attention of the Western world, most forced migrants are displaced within their home countries. According to the UNHCR's (2008) estimate, more than three quarters of the 67 million people forced to move at the end of 2007 were internally displaced. Sadly, forced migration is not likely to cease. If anything, new causes, such as global warming, may increase the number of displaced persons. For instance, a one meter rise in sea levels would permanently inundate the land currently hosting 11 percent of the Bangladeshi population (Agrawala et al., 2003). Even if such scenarios would be far fetched, ordinary peace-time public policies sometimes lead to large-scale forced migrations. According to the World Commission on Dams (2000), for example, dam construction alone has displaced between 40 and 80 million people in the past half a century.

¹This essay is joint work with Roope Uusitalo and Markus Jäntti.

These displacements are likely to have important consequences for those who are forced to migrate as well as the receiving and the sending areas. They also create a major policy challenge. Understanding the consequences of human displacement and the mechanisms leading to these consequences would be essential for developing effective programs to assist people uprooted from their homes. Yet, the topic has been subject to relatively little economic research.²

This paper studies the long-term effects of being internally displaced in Finland after the World War II. The war led Finland to cede a tenth of its territory to the Soviet Union. The entire population from 60 rural municipalities and three cities was evacuated within a few weeks. Altogether 430,000 individuals, 11 percent of the 1940 population, were resettled to the remaining parts of the country. We focus on the long-term impact of migration on those who were forced to migrate. To perform the analysis, we have access to unusual individual-level longitudinal data on the displaced and non-displaced populations from strictly comparable sources. Our data record the situation before the war as well as post-war outcomes up to fifty years after the displacement.

We find that forced migration increased long-term income. Before the war, those living in the areas that were to be ceded did not differ from the rest of the population. By contrast, they were substantially more mobile in the post-war period and earned significantly more in 1971. These findings illustrate that successful resettlement policies are possible. Moreover, they suggest that increased mobility had a positive effect on the long-term earnings of the displaced.

While displacement is an important topic on its own right, we argue that these findings also contribute to a more general understanding of migration. Economists typically considered migration as an equilibrating mechanism promoting efficient resource allocation. In this framework, migration is a human capital investment that involves both costs and benefits for the migrants (Sjaastad, 1962). In the absence of costs, labor flows would equalize the value of the marginal product of workers across labor markets and the output of the economy would be maximized. However, large regional wage

²In his survey, the only papers Lucas (1997) finds on displaced persons are those by Schultz (1971) and Gottschang (1987). More recent research include Czaika and Kis-Katos (2007), Ibáñez and Vélez (2008) and Kondylis (2007, 2008).

differentials persist even after conditioning for observable characteristics of the residents.

The situation we analyze provides an interesting natural experiment that allows us to assess whether differences in unobservable characteristics or high migration costs give rise to regional disparities. That is, in equilibrium, the returns to migration are equal to its costs for a marginal migrant. Therefore, a consistent estimate of returns to migration would be informative about the magnitude of these costs. The challenge to estimation is that if returns to migration vary across individuals, those who decide to migrate are a self-selected group of the population. As a consequence, correlations between migrant status and outcomes of interest in a typical observational dataset are likely to overstate the returns to migration (see Greenwood, 1997, for a survey). In our case, the selection problem does not occur, since the entire population in a certain area was forced to migrate. Therefore, we estimate returns to (forced) migration for a randomly chosen individual in the sample.

Our results suggest that large economic gains from migrating would have been available also for those who were not displaced. However, many chose not to exploit these opportunities. This is consistent with migration carrying a high cost. In essence, being displaced reduced these costs. This finding is highly relevant for regional policy. While we certainly do not advocate policies forcing people to migrate, the results suggest that less brutal policies, such as providing assistance for voluntary movers, may be more efficient than policies supporting those who remain in the less-advantaged regions. Furthermore, higher mobility would also be likely to help those who stay put by reducing regional and occupational wage differences (Borjas, 2001).

The rest of this paper is organized as follows. Section 3.2 provides the details of the post-WWII situation in Finland and the implemented settlement policy. Section 3.3 describes the data. Section 3.4 discusses our empirical approach. Section 3.5 present the results and section 3.6 provides robustness checks. We discuss the implications of the results in section 3.7. Section 3.8 concludes.

3.2. Historical Background

In August 1939, just before the outbreak of the World War II, Germany and the Soviet Union signed a non-aggression pact that included a secret clause in which the two signatories divided up Eastern Europe between themselves. Finland was consigned to the Soviet sphere of influence. Following

unsuccessful negotiations about Soviet Union's territorial demands, the Red Army attacked Finland in November 1939, three months after the German invasion into Poland. In the peace treaty ending the battles between Finland and Soviet Union in March 1940, Finland ceded roughly a tenth of its territory to the Soviet Union. The entire population of these areas had been evacuated during the war. An Emergency Settlement Act (*Pika-asutuslaki*) was enacted in July 1940 to settle the displaced into the rest of the country.

The execution of the Act was suspended in June 1941, when Finland joined Germany's attack to the Soviet Union. By the end of August, the Finnish troops had reoccupied the ceded areas and on December 6th, the Finnish Parliament declared them re-united with the rest of Finland. Approximately two thirds of the displaced returned to their pre-war homes.

However, in the summer of 1944, the Red Army pushed the Finnish troops back to roughly the same line of defense they had held at the end of the first war. The entire population of the ceded areas was again evacuated. The armistice signed in September 1944, and later ratified in the Paris peace treaty in 1947, restored the 1940 borders with some additional areas ceded to the Soviet Union. Finland also agreed to pay USD300 million in war reparations and to expel German troops from its territory, which led to Finland's third armed conflict during World War II.

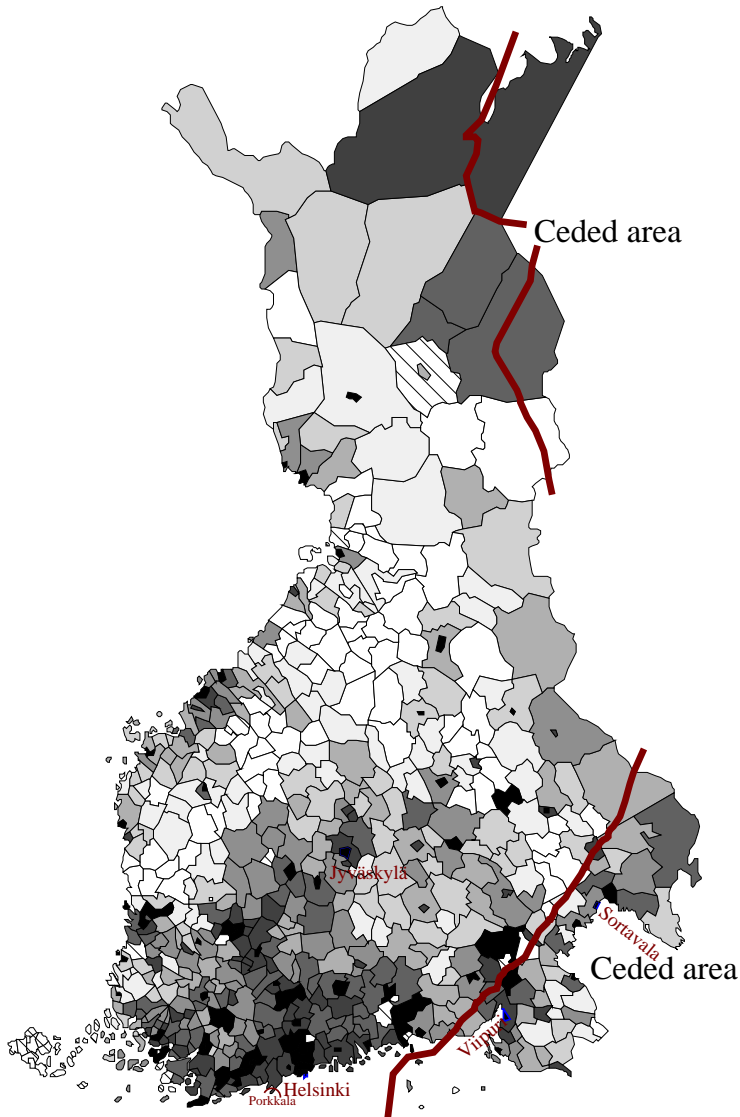
Figure 3.2.1 presents the ceded areas on a 1938 map of Finland.³ The map also displays average income, showing that the ceded area include both rich and poor areas. As we will show below, the average income in the ceded area was close to the national average.

The war left Finland with approximately 92,000 dead and 228,000 injured out of total population of four million. Much of the country's production capacity was destroyed in the war and further cuts in capacity were caused by war reparations. For example, a quarter of the Finnish commercial fleet was handed over to the Soviet Union. Altogether war reparations took about 15 percent of the government budget between 1945 and 1949 (Tarkka, 1988).

Even without such a dire situation, settling the 430,000 displaced persons would have been a major burden. Finland was still a predominantly agrarian society, in many ways resembling current middle-income developing

³In addition to the area ceded in 1940, the Petsamo area, in the North of Finland, was ceded to the Soviet Union in 1944. Furthermore, the Porkkala Peninsula near Helsinki was leased for a Soviet naval base for fifty years. Following an improvement in international relations and changes in military technology that made land-based artillery less important for protecting Leningrad from the sea, Porkkala was returned to Finland in 1956.

FIGURE 3.2.1. Ceded Areas and Taxable Income per capita, 1938



Note: Darker shades indicate higher taxable income per capita.

countries, with roughly half of the working age population employed in agriculture. Similarly, almost one half of those forced to migrate were farmers. The only feasible option at the time was to resettle a large fraction of the displaced to areas where they could derive their main income from farming.

In May 1945, the Parliament approved the Land Acquisition Act (*Maanhankintalaki*), which guided the settlement policy. The displaced who had

owned or rented land in the ceded areas and had received their principal income from agriculture were entitled to receive land from remaining parts of the country. Others received compensation for their lost property in the form of government bonds. Land was primarily taken from the state, the local governments (municipalities) and the church, but the required amount far exceeded the capacity of the public sector. Thus, roughly two thirds of the cultivated fields, one half of land that could be cleared for cultivation and a third of forest land was seized from private owners. The committee drafting the law proposed an explicit progressive expropriation schedule for seizing private land.⁴

The implementation of the Land Acquisition Act was entrusted to the Department of Land Settlement in the Ministry of Agriculture. Altogether 147 local land redemption boards were responsible for the expropriation measures and the same number of local settlement boards had a duty to locate applicants for land. Those coming from each Karelian village were settled into a designated target municipality. The number of displaced allocated in each municipality was mainly affected by the availability of suitable land, which again mainly depended on the pre-war farm size distribution and on the quantity of state-owned land in the municipality. The most important factor in allocation across receiving municipalities was the location of their municipality in the ceded area. Those from the western parts of the Karelian peninsula were settled along the southern coast, those from the eastern part of the Karelian peninsula north of the first group and those from Northern Karelia even further north. None were placed in Northern Finland, where conditions for agriculture are unfavorable. Those from the municipalities surrounding Viipuri, the largest city in Karelia, were settled close to the capital, Helsinki, and those from Sortavala, the second largest Karelian city, close to Jyväskylä, a city located in the middle part of the country. Persons from the same village were settled in the same municipality and neighboring villages were settled close to each other.

The destination of the non-agrarian displaced was mainly determined by the availability of housing and the distance from the ceded areas. Cities in Eastern Finland received flows of displaced persons that constituted almost

⁴The proposed schedule required private land owners to cede up to 80% of their land holdings depending on the size of their farms. No land was expropriated from farms smaller than 25 hectares. The landowners were compensated with government bonds yielding 4% nominal interest. Inflation eventually wiped out about four fifths of their value. Pihkala (1952) discusses the acquisition policy and its consequences in detail.

ten percent of the population, while cities further west and cities with the most severe housing problems received much less.⁵ While the non-agrarian population was not explicitly allocated, the settlement plan influenced also their migration due to family ties and employment opportunities with their former landowner employers. In June 1949, 53 percent of the displaced lived in their designated placement areas (Waris et al., 1952).

3.3. Data

Our primary source of data is the 1950 population census, the first full census implemented in Finland. Data were collected by personal interviews and the information for each dwelling unit was stored on a single form. These forms were sorted by municipality, within municipalities in alphabetical order and then filed in boxes. In 1997, Statistics Finland drew a sample from the full 1950 census by picking every tenth box. Nearly all of the information on the census forms was then keyed into a database. The resulting sample contained about 114,000 dwelling units with 411,629 persons from 392 of Finland's 547 municipalities. Based on the first and the last names, along with date and place of birth, Statistics Finland matched these data to the Population Register in order to find the social security number – which had been introduced in 1964 – for each person. Thus, in order to end up in the final sample, the person had to be alive and live in Finland in mid-1960s. Social security numbers allowed merging the 1950 census file to a longitudinal census file containing information from population censuses performed every five years starting in 1970. Hence, the 1950 census sample can be followed through each subsequent census up to the year 2000.⁶ Statistics Finland (1996) provides a detailed discussion of the data.

The 1950 census contains information on various household characteristics and person-level information such as place of birth and residence, education, occupation and sector of employment. Importantly for our purposes, it also contains retrospective information concerning municipality of residence

⁵The share of the displaced in 1948 living in urban areas varied from 9.2% in Mikkeli, 8% in Jyväskylä and 7.8% in Lahti (all located in Central or Eastern Finland) to 2.6% in Pori (on the west coast) and 0.4% in Pietarsaari (a Swedish-speaking town on the western coast). Housing shortages in the capital city, Helsinki, led to direct regulation. In 1945, those who wished to move to Helsinki had to apply for a specific permission from the local housing board.

⁶Ideally we would also have data from the 1960 census. However the original punch cards were destroyed in the 1970s and the magnetic tapes (where the data had been stored) were damaged in storage.

in 1939. Other information from 1939 includes occupational status and industry codes referring to September 1st, 1939 – two months before the war began. The same information is available for 1950. This creates an unusual situation where we have longitudinal microdata on the displaced and the non-displaced from pre- and post-displacement periods. The same survey instruments were administered to both groups and hence all information is fully comparable.

Linking the 1950 census to the longitudinal census data for 1970–2000 further increases the amount of available information. Most importantly, the 1970 census include tax record data from 1971. This provides an opportunity to evaluate the long-term effects of migration 27 years after the war ended. Even longer-term effects on income can be evaluated by using data on taxable income around 1980–1990 when most of the (surviving) displaced are already retired from the labor force. Since pension income is affected by the accrued pension rights from each employment spell during the career, the taxable income of pensioners is a reasonable proxy for lifetime income.

The most important shortcoming of the data is that the 1950 census did not collect any direct information on income. However, the data contain several variables that are informative about the economic status of the individuals. In order to summarize this information efficiently, we construct two measures of imputed income. Our main approach is to use the information available in the microdata on the 1971 earnings. That is, we break down the data into 590 industry-socioeconomic status-living in an urban area categories using the 1939 and 1950 data and assign each observation the average 1971 log annual income of the respective cell. To complement this measure, we also use detailed annual tables listing taxable earnings based on reports from local tax boards (Statistics Finland, 1942, 1953). These data allow us to classify the observations into 38 industry-occupation-socioeconomic status groups for 1950 and 12 industry-living in urban area groups for 1939 and assign each individual the mean income of their reference group. The Appendix discusses our procedures in detail.

We have access to a smaller random sample of the data originally stored by Statistics Finland. In order to focus on those who were of working age throughout the period from 1939 to 1971, we further limit the analysis to individuals born between 1907 and 1925. Moreover, we exclude observations where municipality of residence in 1939 is unknown (312 persons), who lived in the 12 partly ceded municipalities (642 persons) or on the Åland Islands

(178 persons) as well as those who lived in the 13 municipalities for which pre-war data on taxable income per capita is missing (498 persons). This leaves us with information on 22,771 individuals of whom 2,558 were displaced.

3.3.1. Descriptive Statistics. Table 3.3.1 reports sample means for the displaced and the rest of the population in 1939, 1950 and 1970/71. The pre-war means are relatively similar. The main difference is the lower share of blue-collar workers among the future displaced. This reflects partly the lower share working in manufacturing and partly the slightly higher share of assisting family members and those out of the labor force. Furthermore, the share working in the formal labor market, i.e. working as an entrepreneur or hired labor, and the share living in urban areas are lower among those who were later displaced. However, these differences are rather small in magnitude. Importantly, as we will see in more detail in Section 3.5, there are no substantial income differences by either definition of imputed income or by average taxable income per capita in the municipality of residence. The only large difference is that very few of those living in the ceded areas spoke Swedish as their mother tongue.

The next two columns report means in 1950. The most important changes are decreases in the share of the population employed in agriculture, by 17 and 9 percentage points among the displaced and the rest of the population, respectively. The shares of population working in manufacturing are now equal for the displaced and the others. The share working in service or other sectors increased among the displaced, but remained constant among the rest of the population. The fast flow away from agriculture among the displaced is also evident in that they are now more likely to work as blue-collar workers than the non-displaced. As a result, their imputed incomes are roughly six percent higher than those of the non-displaced. There are no large differences in educational attainment between the displaced and the non-displaced. Furthermore, the displaced were more likely to live in cities or market towns, though the difference remains small in magnitude. Not surprisingly – given that the displaced had lost their homes only six years earlier – there was a clear difference in the fraction living in owner-occupied housing.

The remaining columns report means in 1970/71. The share of the sample working in agriculture continues to decrease, but the gap between the displaced and the rest of the population remains at ten percentage points.

TABLE 3.3.1. Sample means

	1939		1950		1970	
	Disp.	Other	Disp.	Other	Disp.	Other
Age	22.9	22.8	33.9	33.8	53.8	53.7
Female	0.53	0.53	0.53	0.53	0.55	0.54
Lives in urban area	0.24	0.27	0.33	0.30	0.48	0.41
Native tongue Swedish	0.01	0.09	0.01	0.09	0.00	0.09
Public sector	na.	na.	0.20	0.14	0.15	0.13
Single	na.	na.	0.13	0.12	0.20	0.18
Owner occupied housing	na.	na.	0.43	0.54	0.71	0.76
Migrated	na.	na.	1.00	0.29	0.35	0.24
Taxable income per capita in the mun. of residence	46.8	50.3	688	635	16.0	14.7
Imp. log income (def. 1)	1.91	1.94	2.35	2.28	2.32	2.26
Imp. log income (def. 2)	2.93	2.92	5.40	5.34	na.	na.
Annual taxable income	na.	na.	na.	na.	14.5	13.2
<i>Socioeconomic status</i>						
entrepreneur	0.09	0.08	0.14	0.18	0.17	0.22
white collar worker	0.10	0.10	0.21	0.19	0.20	0.18
blue collar worker	0.27	0.34	0.47	0.44	0.37	0.33
assisting family member	0.16	0.14	0.16	0.17	0.04	0.06
unemployed/out of LF	0.38	0.35	0.02	0.02	0.22	0.21
Employed	0.47	0.52	0.82	0.81	0.74	0.73
<i>Sector (conditional on not unknown/missing)</i>						
Agriculture	0.45	0.47	0.28	0.38	0.22	0.32
Manufacturing etc.	0.19	0.21	0.26	0.26	0.26	0.23
Construction	0.07	0.06	0.11	0.09	0.12	0.10
Service etc.	0.30	0.27	0.34	0.27	0.40	0.36
Sector unknown	0.40	0.37	0.05	0.05	0.23	0.22
<i>Education</i>						
Primary	na.	na.	0.78	0.75	na.	na.
Secondary (9 years)	na.	na.	0.07	0.06	na.	na.
Secondary (12 years)	na.	na.	0.03	0.04	na.	na.
None / unknown	na.	na.	0.12	0.15	na.	na.
Observations	2,558	20,213	2,558	20,213	2,326	18,454

Note: Taxable income per capita in the municipality of residence are in nominal terms and not comparable across time periods. Imputed earnings (definition 1) assigns the 1971 earnings distribution to all years on industry-socioeconomic status-living in an urban area level. Imputed earnings (definition 2) is based on 1939 and 1950 actual earnings distributions in nominal terms and not comparable across time periods. See the Appendix for details.

The displaced also remain more likely to work in the service and construction sectors, and now also in manufacturing. Interestingly, the difference in the proportion living in urban areas has increased markedly in comparison to 1950. As a consequence, the displaced now live in municipalities with higher mean taxable income. They are still less likely to live in owner-occupied housing. Most importantly, the displaced have higher annual earnings than the non-displaced, suggesting that forced migration or perhaps higher post-war mobility may have had a positive long-term effect on income.

3.4. Empirical Approach

We will evaluate the impact of being displaced by comparing the outcomes of those living in the ceded area to comparable persons who were not forced to move due to the war. Our base estimation equation is

$$(3.1) \quad y_{ijt} = \alpha D_i + \mathbf{X}_{0i}\beta + \varepsilon_{ijt}$$

where y_{ijt} is the outcome of interest for individual i living in location j at time t , D_i is a dummy indicating displacement status, \mathbf{X}_{0i} is a vector of observable characteristics measured before the war and ε_{ijt} summarizes the unobservable factors. In order to interpret the results and to state our identifying assumption clearly, we divide the unobservables into two parts

$$(3.2) \quad \varepsilon_{ijt} = u_{jt} + \nu_{it}$$

where u_{jt} captures the unobserved “quality” of the labor market j at time t and ν_{it} are individual specific unobserved characteristics.

Our identifying assumption is that, once we condition for the observed characteristics, displacement status is uncorrelated with unobserved individual characteristics:

$$(3.3) \quad \text{Cov}(D_i, \nu_{it} | \mathbf{X}_{0i}) = 0.$$

Since the location of the new border was determined as an outcome of the battles, this assumption seems plausible. Those who happened to live on the ceded area had no option but to move. Nevertheless, one could argue that those living in the Western part of the country prior to the war could have differed from those living in the Eastern part. Section 3.5 provides two types of evidence suggesting that this was not the case. First, the available data suggest there were few important pre-war differences between the future displaced and non-displaced persons. Second, the estimates are not sensitive to

controlling for pre-war observables characteristics or for excluding the Western part of the country from the analysis. Thus, we argue that assumption (3.3) holds and our estimates can be considered causal in a sense discussed next.

3.4.1. Interpretation of the Base Estimates. The key to interpreting our main results is to note that the correlation between displacement status (D_i) and post-war labor market quality (u_{jt}) is positive. This occurs for at least three reasons. First, resettlement moved individuals to new labor markets and occupations. While the displaced had limited possibilities to choose their initial destination, the authorities assigned more displaced persons to more prosperous municipalities. Second, the displacement appears to have hastened the transition from traditional to modern occupations. Third, the likelihood of moving again between geographic areas after the resettlement appears to have been higher among the displaced than among the rest of the population. These later moves are likely to be correlated with job opportunities. If the displaced had lower costs of moving, perhaps because they had less reason to stay in the placement area, post-war sorting across municipalities may differ between the displaced and the rest of the population.

These considerations can be summarized formally by noting that a least-squares estimator of α has the probability limit

$$(3.4) \quad \text{plim } \hat{\alpha} = \pi_D + \theta$$

where π_D is the partial correlation between displacement status and labor market quality – where labor market refers to the interaction between a spatial location and an occupation – and θ captures direct effects of being displaced.⁷ That is, the impact of forced migration works through two mechanisms. The first is moving to better labor markets. The second are “other” effects due to e.g. loss of location specific human capital, trauma or loss of property. The same interpretation carries over to quantile regressions.

Given assumption (3.3), $\hat{\alpha}$ has a causal interpretation. However, it is important to carefully consider the nature of this causality. Clearly, resettlement of such a large scale was likely to lead to general equilibrium effects.

⁷To derive this probability limit, suppose that the data generating process is $y_{ijt} = \theta D_i + \mathbf{X}_{0i}\beta + u_{jt} + \nu_{it}$. Equation (3.4) then follows from the familiar omitted variables bias, i.e. π_D is the probability limit of the OLS estimator of the displacement status in a regression of u_{jt} on D_i and \mathbf{X}_{0i} .

Thus, the estimates are not informative about the counter-factual world where Finland had not ceded any areas and no one was evacuated. In other words, the evacuation “experiment” probably affected both the “controls” and the “treated”. In this case the causal effect of being displaced is the difference between outcomes that the displaced experienced and outcomes in a counter-factual state where the post-war border would have been drawn so that their homes remained on the Finnish side, but many others were still forced to move.

3.4.2. Controlling for Current Labor Market Quality. Below we will also report estimates from specifications adding post-war labor market fixed-effects to equation (3.1). The motivation is to estimate θ by conditioning on the current labor market. This would allow us to divide the impact of being displaced into a part due to increase in labor market quality and into a part due to other reasons. Interestingly, however, adding labor market fixed effects to equation (3.1) may not produce unbiased estimate of θ . This is due to the fact that even if the displacement status D_i and individual specific factors ν_{it} are uncorrelated, they are uncorrelated *conditional on* u_{jt} only if $Cov(D_i, u_{jt}|\mathbf{X}) = 0$ or $Cov(u_{jt}, \nu_{it}|\mathbf{X}) = 0$. We have already argued that the first equality does not hold. The second equality would imply that individuals do not sort into localities based on their unobservable characteristics. This is also unlikely to be true. If those with above average unobserved characteristics are more likely to be located in above average labor markets, our estimate of θ will be biased downwards. Even in this case, estimates controlling for labor market fixed effects contain useful information. They are the average differences between a displaced and a non-displaced person, who were similar in their pre-war characteristics and who lived in similar labor markets in 1970.

3.5. Results

We now report the results from regression models discussed above. In the following two Sections, we focus on summarizing the data and postpone a more speculative discussion on the implications of the results to section 3.7. We begin by asking whether we can find any evidence that the displaced differed from the rest of the population prior to the war. We then present our main results on the causal effect of forced migration among the displaced on several long-term outcomes. This is followed by studying the impact of

conditioning on post-war labor market and heterogeneity of the main effects as well as going through several robustness checks.

3.5.1. Pre-War Differences. In Section 3.3 we saw that the pre-war mean characteristics of the displaced were rather similar to those of the rest of the population. We now perform simple tests of the significance of the pre-war differences by regressing the available pre-war information on an indicator variable taking value one if the person was to become displaced after the war and zero otherwise. Table 3.5.1 reports the results separately for men (panel A) and women (panel B), from a bivariate regression (column 2) and from a specification controlling for age, longitude and latitude of the 1939 residence municipality and dummies for living in urban area and speaking Swedish as one's mother tongue in 1939 (column 3). The outcomes considered are a dummy for being employed as an entrepreneur or hired labor, two definitions of imputed income and taxable income per capita in the 1939 municipality of residence. Both imputed income measures vary only between industry, socioeconomic status and urban versus rural residence. Thus, the estimates are informative only on whether the future displaced were employed in occupations carrying above average earnings.

The results suggest that pre-war differences between future displaced and the rest of the population were small or nonexistent. Out of the 16 estimates, only two are statistically significant at the 10% level or lower. These significant estimates would imply that men who were later displaced were working in pre-war occupations with a slightly higher average wages and that women living in the ceded areas were less likely to be employed in the formal labor market than comparable women in other parts of the country. Note, however, that given a large number of regressions, one should expect to get statistically significant estimates occasionally even if the outcomes were randomly generated. Thus we conclude that, on balance, we find little evidence suggesting that the economic performance of the displaced differed from the rest of the population prior to the war. Furthermore, as our base specification controls for these pre-war observable characteristics, assumption (3.3) is very likely to hold.

3.5.2. Main Results. We next turn to post-war outcomes. Tables 3.5.2, 3.5.3, 3.5.4 and 3.5.5 present our main results. Each table has a similar structure, where column (2) reports the results from regressing several post-war outcomes on a binary indicator of the person being displaced. Each

TABLE 3.5.1. Pre-War Differences

Dependent Variable	Mean	Displacement status		Obs.
		w/o Cov. (2)	w Cov. (3)	
<i>A: Men</i>				
Employed in the Formal Labor Market	0.71	-0.033 (0.034)	-0.038 (0.023)	10,673
Imputed log Earnings (def. 1)	2.48	0.001 (0.098)	0.008 (0.031)	10,270
Imputed log Earnings (def. 2)	2.87	0.015 (0.032)	0.027* (0.017)	8,590
log Taxable income per capita in the municipality of residence	3.54	0.089 (0.222)	-0.076 (0.122)	10,673
<i>B: Women</i>				
Employed in the Formal Labor Market	0.33	-0.063 (0.052)	-0.042** (0.018)	12,098
Imputed log Earnings (def. 1)	1.47	-0.053 (0.102)	-0.031 (0.024)	11,914
Imputed log Earnings (def. 2)	2.98	0.010 (0.059)	0.003 (0.018)	5,603
log Taxable income per capita in the municipality of residence	3.63	0.022 (0.239)	-0.114 (0.128)	12,098

Note: OLS estimates of future displacement status and pre-war outcomes. Each coefficient comes from a separate regression. Standard errors (in parentheses) robust to clustering at municipality of residence in 1939 level. Specification for column (3) controls for age, latitude and longitude of the 1939 residence municipality, and indicator variables for speaking Swedish as ones mother tongue and living in urban area in 1939. ***, **, * indicate statistical significance at 1%, 5%, 10% level, respectively.

coefficient comes from a separate regression. Under the assumption that being displaced was random, these estimates correspond to the parameter $\hat{\alpha}$ in equation (3.4). That is, they are informative about the causal effect of being forced to migrate inside Finland after the Second World War. Column (3) reports similar estimates from a specification controlling for age, imputed earnings in 1939, longitude and latitude of the 1939 municipality of residence, and indicator variables for being Swedish-speaking and living in urban area in 1939. The main impact of conditioning on these pre-war observable characteristics is the improvement in precision, while point estimates are virtually

TABLE 3.5.2. Short-Term Outcomes (Men)

Dependent Variable	Mean	The Effect of Displacement		Conditioning on Post-War Labor Market and Covs			Obs.
		w/o Covs (2)	w/ Covs (3)	1950a (4)	1970a (5)	1970b (6)	
Emp. in Form	0.90	-0.02 (0.02)	-0.02 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	10,673
Imp. log Income (1)	2.7	0.06 (0.08)	0.07*** (0.02)	0.02 (0.02)	0.06** (0.03)	0.04* (0.02)	10,509
Imputed log Income (2)	5.3	0.09** (0.04)	0.11*** (0.02)	0.07*** (0.02)	0.10*** (0.03)	0.09*** (0.03)	9,543
Lives in urban area	0.29	0.04 (0.10)	0.06 (0.04)	.	0.02 (0.04)	0.02 (0.04)	10,673
Changed Sect. (1-dig)	0.42	0.10*** (0.01)	0.10*** (0.02)	0.04 (0.02)	0.09*** (0.03)	0.09*** (0.03)	10,673
Changed Sec. (2-dig)	0.48	0.09*** (0.02)	0.11*** (0.02)	0.06*** (0.02)	0.11*** (0.03)	0.11*** (0.03)	10,673

Note: Impact of being displaced. OLS estimates on 1950 outcomes, standard errors (in parentheses) robust to clustering at 1939 residence municipality level. Each coefficient comes from a separate regression. Covariates: age, imputed earnings in 1939, latitude, longitude and log taxable income per capita of the 1939 residence municipality, and indicator variables for speaking Swedish as ones mother tongue and living in city or market town in 1939. ***, **, * indicate statistical significance at 1%, 5%, 10% level, respectively.

unaffected. The remaining columns present results conditioning on post-war labor markets. We return to the interpretation of these estimates in the next subsection.

Consider first the impact of forced migration on short-term outcomes, measured in 1950. The point estimates in bottom rows of Table 3.5.2 suggest that displaced men were roughly ten percentage points more likely to change their sector of employment than non-displaced men. The estimates for imputed earnings show that these flows took place towards occupations carrying roughly 10 per cent higher wages (recall that imputed earnings are constructed as the mean earnings in cells defined by industry-socioeconomic status-living in urban area). As we already saw above, this predominantly reflects more frequent transitions from agriculture to manufacturing and services among the displaced, while transitions between categories of socioeconomic status are similar among both groups. Interestingly, however, we find no effect on the propensity to work in the formal labor market.

TABLE 3.5.3. Short-Term Outcomes (Women)

Dependent Variable	Mean	The Effect of Displacement		Conditioning on Post-War Labor Market and Covs			Obs.
		w/o Covs (2)	w/ Covs (3)	1950a (4)	1970a (5)	1970b (6)	
Emp. in Form	0.73	0.04 (0.05)	0.04* (0.02)	-0.00 (0.02)	0.02 (0.02)	0.02 (0.02)	12,098
Imp. log Income (1)	1.9	0.08 (0.10)	0.11*** (0.03)	0.04 (0.03)	0.06** (0.03)	0.06** (0.02)	11,982
Imputed log Income (2)	5.4	0.03 (0.03)	0.07*** (0.02)	0.04* (0.02)	0.04* (0.02)	0.04* (0.02)	8,765
Lives in urban area	0.32	0.01 (0.10)	0.06 (0.04)	.	0.01 (0.03)	0.01 (0.03)	12,098
Changed Sect. (1-dig)	0.69	0.05*** (0.01)	0.01 (0.02)	0.02 (0.02)	0.02 (0.02)	0.02 (0.02)	12,098
Changed Sec. (2-dig)	0.73	0.04** (0.02)	0.02 (0.02)	0.03 (0.02)	0.03 (0.02)	0.03 (0.02)	12,098

Note: Impact of being displaced on 1950 outcomes. OLS estimates, standard errors (in parentheses) robust to clustering at 1939 residence municipality level. Each coefficient comes from a separate regression. Covariates: age, imputed earnings in 1939, latitude, longitude and log taxable income per capita of the 1939 residence municipality, and indicator variables for speaking Swedish as ones mother tongue and living in city or market town in 1939. ***, **, * indicate statistical significance at 1%, 5%, 10% level, respectively.

The impact on urbanization is less clear. While the point estimates suggest that displacement increased the likelihood of moving to urban areas, the results are not statistically significant.⁸ Furthermore, the estimates presented in Table 3.5.3 suggest that displaced women were as likely to change their sector of employment as other women. However, this result is due to three out of four of sector changes between 1939 and 1950 among women being due to leaving the category “unknown”. A closer look reveals that in comparison to other women, the displaced are less likely to switch from “unknown” sector of employment to agriculture and more likely to services. These transitions are also captured by the impact of the displacement on imputed earnings, which is similar to the estimates for men.

⁸Note, however, that our inference is rather conservative as the standard errors are clustered at 1939 residence municipality level. We chose this approach in order to take into account that persons coming from same areas might have been exposed to common shocks, in particular due to the settlement plan affecting individuals based on their pre-war residence municipality.

Consider next the long-term effects on the 1970 situation presented in Tables 3.5.4 and 3.5.5. We now find a positive impact on the propensity to be employed in the formal labor market. More importantly, the data also allow us to study the impact of displacement on actual annual taxable income in 1971. The estimates for men suggest that forced migration substantially increased long-term income relative to the non-displaced. The point estimate for income levels (including zeros) indicates a difference of 2,600 Finnish marks, which corresponds to roughly an 11% difference in annual income. The estimate for log income implies that the impact would have been even larger. As we will show below, the difference in the estimated mean effect when measuring income in levels or logs is mostly due to the effect being larger in the lower end of the income distribution. Using log income as an outcome measure places more weight on those in the lower end of the income distribution, where the effect is large, and therefore yields higher estimates. The impact on pension income suggests a roughly 11 percent increase in lifetime income. For women, the results are mixed. While the impact on pension income is similar to that for men, the 1971 income of displaced women did not differ from that of other women.

On balance, these results suggest that being forced to migrate after World War II in Finland had a sizable positive causal effect on lifetime income. We next turn to possible sources of this effect. One candidate explanation is the transformation from traditional to modern occupations. As we already saw, being displaced increased the pace of this process. However, while the displacement had a permanent impact on the occupational distribution, the largest changes occurred soon after the war. For the period between 1950 and 1970, we find no difference in the likelihood of changing sector of employment. With regard to mobility on 1-digit occupational category level, we find a small but statistically significant positive effect for women. Nevertheless, the difference in imputed earnings remains virtually unchanged in comparison to short term effects.

However, estimates presented in the bottom rows of Tables 3.5.4 and 3.5.5 show that the displaced remained more mobile also in the period between 1950 and 1970. Instead of changing occupation, this mobility now takes the form of a higher propensity to change municipality of residence. Furthermore, among those who migrated, the displaced tended to move longer distances than the others. Finally, while the point estimate for living in an

TABLE 3.5.4. Long-Term Outcomes (Men)

Dependent Variable	Mean	The Effect of Displacement		Conditioning on Post-War Labor Market and Covs			Obs.
		w/o Covs (2)	w/ Covs (3)	1950 (4)	1970a (5)	1970b (6)	
Emp. in Formal Labor Market	0.82	-0.00 (0.01)	0.03** (0.02)	0.02 (0.02)	0.01 (0.02)	.	9,496
Imputed log Income (def. 1)	2.7	0.06 (0.05)	0.10*** (0.02)	0.04 (0.02)	0.03 (0.02)	.	9,494
Taxable Income ('000 marks)	20.8	2.26 (2.17)	2.64*** (0.99)	2.03** (0.96)	2.30** (0.96)	1.46** (0.72)	9,496
log Income	2.8	0.17** (0.08)	0.24*** (0.04)	0.12*** (0.04)	0.11** (0.04)	0.08** (0.03)	8,572
log Pensions (various years)	2.4	0.13* (0.07)	0.13*** (0.05)	0.12** (0.05)	0.11** (0.05)	0.07 (0.04)	7,383
Changed Sec. (1-dig)	0.46	0.04** (0.02)	-0.00 (0.03)	0.01 (0.03)	-0.01 (0.03)	-0.01 (0.02)	9,496
Changed Sec. (2-dig)	0.56	0.04** (0.02)	-0.01 (0.03)	0.02 (0.03)	0.01 (0.03)	0.02 (0.02)	9,496
Changed Occ. (1-dig)	0.44	0.03* (0.02)	-0.02 (0.02)	-0.00 (0.02)	-0.04 (0.03)	-0.04* (0.02)	9,496
Lives in urban area	0.40	0.08 (0.07)	0.08* (0.05)	0.04* (0.02)	.	.	9,496
Migrates 1950–1970	0.36	0.11*** (0.02)	0.04 (0.03)	0.02 (0.03)	0.02 (0.04)	0.01 (0.04)	10,510
Distance (cond. on migrating)	130.0	10.97 (11.66)	61.42* (36.34)	0.98 (11.88)	38.34 (27.75)	35.86 (28.26)	2,535

Note: Impact of being displaced. OLS estimates, standard errors (in parentheses) robust to clustering at 1939 residence municipality level. Income measured in 1971 except pensions, which are measured as the first observed pension income in 1975, 1980, 1985 or 1990. Other outcomes measured in 1970. Each coefficient comes from a separate regression. Covariates: age, imputed earnings in 1939, latitude, longitude and log taxable income per capita of the 1939 residence municipality, and indicator variables for speaking Swedish as ones mother tongue and living in city or market town in 1939. Estimates reported in column (4) are also conditional on 1950 residence municipality, estimates presented in column (5) are conditional on 1970 residence municipality, estimates presented in column (6) are conditional on 1970 residence municipality, industry and socioeconomic status fixed-effects. ***, **, * indicate statistical significance at 1%, 5%, 10% level, respectively.

TABLE 3.5.5. Long-Term Outcomes (Women)

Dependent Variable	Mean	The Effect of Displacement		Conditioning on Post-War Labor Market and Covs			Obs.
		w/o Covs (2)	w/ Covs (3)	1950a (4)	1970a (5)	1970b (6)	
Emp. in Formal Labor Market	0.66	0.01 (0.02)	0.06*** (0.02)	0.03* (0.02)	0.03* (0.02)	.	11,284
Imputed log Income (def. 1)	1.9	0.06 (0.06)	0.11*** (0.03)	0.06*** (0.02)	0.04** (0.02)	.	11,280
Taxable Income ('000 marks)	7.2	0.48 (1.02)	-0.02 (0.48)	0.09 (0.57)	-0.43 (0.52)	-0.80* (0.45)	11,284
log Income	2.3	-0.01 (0.08)	-0.01 (0.05)	0.08 (0.05)	-0.01 (0.05)	-0.05 (0.04)	5,664
log Pensions (various years)	1.8	0.09 (0.07)	0.11** (0.05)	0.09** (0.04)	0.08** (0.04)	0.06 (0.04)	8,554
Changed Sec. (1-dig)	0.54	0.05*** (0.01)	0.02 (0.02)	0.02 (0.02)	-0.01 (0.02)	0.01 (0.02)	11,284
Changed Sec. (2-dig)	0.64	0.05** (0.02)	0.02 (0.02)	0.01 (0.02)	-0.01 (0.02)	0.00 (0.02)	11,284
Changed Occ. (1-dig)	0.51	0.03* (0.01)	0.03* (0.02)	0.05** (0.02)	0.02 (0.02)	0.01 (0.02)	11,284
Lives in urban area	0.43	0.07 (0.07)	0.07 (0.05)	0.05** (0.02)	.	.	11,284
Migrates 1950–1970	0.32	0.10*** (0.02)	0.05* (0.03)	0.04 (0.03)	0.02 (0.04)	0.02 (0.04)	11,916
Distance (cond. on migrating)	129.3	11.83 (8.85)	56.91* (30.45)	2.78 (11.23)	29.61 (19.96)	28.79 (19.95)	2,989

Note: Impact of being displaced. OLS estimates, standard errors (in parentheses) robust to clustering at 1939 residence municipality level. Income measured in 1971 except pensions, which are measured as the first observed pension income in 1975, 1980, 1985 or 1990. Other outcomes measured in 1970. Each coefficient comes from a separate regression. Covariates: age, imputed earnings in 1939, latitude, longitude and log taxable income per capita of the 1939 residence municipality, and indicator variables for speaking Swedish as ones mother tongue and living in city or market town in 1939. Estimates reported in column (4) are also conditional on 1950 residence municipality, estimates presented in column (5) are conditional on 1970 residence municipality, estimates presented in column (6) are conditional on 1970 residence municipality, industry and socioeconomic status fixed-effects. ***, **, * indicate statistical significance at 1%, 5%, 10% level, respectively.

urban area changes only a little, it now becomes statistically significant for men.

3.5.3. Conditioning on Post-War Outcomes. We next study how the coefficients for the displacement status change after conditioning on post-war labor market. Columns (4) and (5) in Tables 3.5.2, 3.5.3, 3.5.4 and 3.5.5 present the estimates when we control for the full vector of indicator variables for the residence municipality in 1950 and 1970, respectively. In columns (6), we control for the 1970 residence municipality, sector of employment (54 categories) and socioeconomic status (5 categories). All specifications also control for the same pre-war observable characteristics as the results presented in column (3). As we discussed in Section 3.4, the coefficients in columns (4) to (6) do not have a causal interpretation, as they condition on post-war outcomes, which themselves are affected by the displacement. Rather, we consider them as descriptive statistics that are informative about the differences between a displaced and non-displaced person, who were similar before the war and were located in the same labor market after the war.

The main insight provided by these conditional differences is that displaced men had, on average, higher annual earnings than non-displaced men working in the same place, in the same industry and having the same socioeconomic status in 1970/71. However, this difference is considerably smaller than the estimate when we condition only on pre-war observable characteristics. Furthermore, while the point estimates for pension income remain positive, they are no longer statistically significant once we condition on all these post-war characteristics. By contrast, displaced women had similar or slightly smaller earnings as non-displaced women regardless of the control variables added to the specification.

We return to the possible explanations for these rather surprising results in Section 3.7. Before turning to more speculative discussion, however, we collect the remaining facts and perform some initial robustness checks (to be followed by additional ones in Section 3.6). With regard to the remaining estimates, most are almost unaffected for conditioning on the post-war outcomes considered. Perhaps the most interesting result is that displaced men were *less* likely to change their occupation between 1950 and 1970 than non-displaced men who were working in the same municipality and industry and having the same socioeconomic status in 1970. While the conditional results for industry/occupation mobility between 1950 and 1970 are weak in terms

of statistical significance, these findings are consistent with the results above, suggesting that the impact of displacement on the pace of moving from traditional to modern occupations was predominantly a short-term phenomenon, although it led to persisting differences. Another interesting finding is that once we condition for the post-war residence municipality, displaced men are not more likely to be employed in the formal labor market than non-displaced men.

3.5.4. Heterogeneity of the Effect. We now turn to the question of whether the impact of forced migration varies across different groups of the displaced. We first look at the heterogeneity of the impact across age groups. Such heterogeneity could occur, for instance, because the younger cohorts have greater incentives to respond to the displacement by acquiring more education or learning a new profession. To some extent, the results support this hypothesis. Columns (1) to (4) in Table 3.5.6 present the estimates from specifications including an interaction term between age and displacement status as well as controls for pre-war observable characteristics. The results suggest that younger displaced were more likely to shift to higher earning industries. That is, the estimates indicate that forced migration increased the 1970 imputed income by 0.15 log-points (standard error 0.03) among those displaced at age 14 in comparison to .03 log-points (standard error 0.04) among those displaced at age 32. Results for women are similar. Furthermore, the impact on being employed in the formal labor market is estimated to be 7.4 percentage points (standard error 2.3) among the youngest men, in comparison to -1.1 percentage points (standard error 2.9) among the oldest. However, the interaction terms are not statistically significant for actual income.

Another likely dimension along which the impact of displacement could vary is the pre-war sector of employment. The results above suggest that one important mechanism through which income effect occurred was a shift from traditional (rural) sectors to modern (urban) industries. The remaining columns in Table 3.5.6 support this hypothesis. We find that the interaction terms between displacement status and a dummy for working in agriculture before the war are significant for both imputed and actual income. They are also large in magnitude. For example, the point estimates for men suggest that the impact of forced migration on actual log income was 0.11 log-points larger for those working in agriculture in 1939 than for those working in

TABLE 3.5.6. Impact of Displacement, Interacted with Age and Working in Agriculture in 1939

Dependent Variable	Age (deviation from mean, years)				Works in Agriculture in 1939			
	Men		Women		Men		Women	
	Disp. (1)	Interaction (2)	Disp. (3)	Interaction (4)	Disp. (5)	Interaction (6)	Disp. (7)	Interaction (8)
Employed (1950)	-0.017 (0.014)	0.001 (0.003)	0.046* (0.024)	-0.004 (0.003)	0.002 (0.015)	-0.023 (0.024)	0.049** (0.023)	-0.000 (0.034)
Employed (1970)	0.033** (0.016)	-0.005** (0.002)	0.058*** (0.016)	-0.003 (0.002)	0.032* (0.018)	0.011 (0.024)	0.052*** (0.017)	0.039 (0.030)
Imputed log Income (1950)	0.068*** (0.022)	-0.001 (0.003)	0.111*** (0.034)	-0.007* (0.004)	0.037* (0.020)	0.072** (0.031)	0.110*** (0.033)	0.040 (0.042)
Imputed log Income (1970)	0.093*** (0.022)	-0.007** (0.003)	0.109*** (0.026)	-0.006** (0.003)	0.053** (0.024)	0.097*** (0.033)	0.095*** (0.028)	0.093** (0.041)
Income (1970)	2.602*** (0.980)	-0.111 (0.110)	-0.007 (0.482)	-0.043 (0.041)	1.922 (1.171)	2.002 (1.353)	-0.181 (0.524)	1.397** (0.628)
log Income (1970)	0.239*** (0.044)	0.002 (0.005)	-0.010 (0.046)	-0.005 (0.007)	0.186*** (0.048)	0.114** (0.055)	-0.023 (0.046)	0.277*** (0.102)
log Pensions (various years)	0.137*** (0.048)	0.005 (0.008)	0.115*** (0.047)	0.005 (0.006)	0.090* (0.052)	0.069 (0.058)	0.112** (0.045)	0.018 (0.091)

Note: Impact of being displaced and interactions with age (columns 1 to 4) and a dummy for working in agriculture in 1939. OLS estimates, standard errors (in parentheses) robust to clustering at 1939 residence municipality level. Controlling for age, imputed earnings in 1939, latitude, longitude and log taxable income per capita of the 1939 residence municipality, and indicator variables for speaking Swedish as ones mother tongue and living in city or market town in 1939. ***, **, * indicate statistical significance at 1%, 5%, 10% level, respectively.

TABLE 3.5.7. Quantile Regressions

	Quantile				
	.1	.25	.5	.75	.9
<i>A: Men</i>					
Taxable Income, 1971 (’000 marks)	0.77 (0.63)	3.15*** (0.83)	3.90*** (0.90)	4.08*** (1.05)	0.65 (2.18)
Quantiles among non-displaced	1	6	17	27	41
log Income, 1971	0.44*** (0.12)	0.29*** (0.07)	0.23*** (0.05)	0.13*** (0.04)	0.04 (0.06)
log Pensions, several years	0.27*** (0.10)	0.08 (0.06)	0.11** (0.05)	0.13** (0.07)	0.08 (0.07)
<i>B: Women</i>					
Taxable Income, 1971 (’000 marks)	.	.	0.41** (0.21)	0.81 (0.71)	1.01 (0.83)
Quantiles among non-displaced	0	0	0	13	21
log Income, 1971	0.12 (0.13)	0.05 (0.08)	-0.07 (0.06)	-0.05 (0.05)	-0.07 (0.06)
log Pensions, several years	0.27*** (0.10)	0.17** (0.07)	0.07 (0.05)	0.04 (0.05)	-0.05 (0.06)

Note: Impact of being displaced. Quantile regressions, bootstrapped standard errors (in parentheses) based on 1,000 replications and robust to clustering at 1939 residence municipality level. Pensions measured as the first observed pension income in 1975, 1980, 1985 or 1990. Controlling for age, imputed earnings in 1939, latitude, longitude, log taxable income per capita of the 1939 residence municipality and indicator variables for speaking Swedish as ones mother tongue and living in city or market town in 1939. ***, **, * indicate statistical significance at 1%, 5%, 10% level, respectively.

other industries. For women, we find a positive effect on actual income *only* for those who worked in agriculture before the war. However, we find no statistically significant evidence on that the impact on pension income would have differed between farmers and the others.

These results imply that the displacement had very different effects for persons differing in at least two observable pre-war characteristics. Another way to study effect heterogeneity is to look at the impact on different parts of the income distribution. To do this, Table 3.5.7 reports estimates from quantile regressions of income on displacement status. There is some variation in the results, depending on the dependent variable and whether we

look at men or women. On balance, however, the results suggest that the impact is largest in the lower end of the income distribution and gradually fades away as we move towards higher quantiles. This is in particular the case for men and log income, where the lowest decile of income among the displaced is 0.44 log points higher than for the non-displaced, compared to a difference in medians of 0.23 log points and an insignificant difference of 0.04 points for the ninth decile. For women, the log income regressions yield no statistically significant differences, although the point estimates are decreasing across the distribution. For pension income, the lowest decile and quartile suggest higher incomes among the displaced. Furthermore, looking at the coefficient for level of taxable income, the median regression now suggests that forced migration had a positive albeit small effect also on female income.

3.6. Robustness Checks

3.6.1. Attrition. Our first robustness check considers the possibility that those observed in 1970 may not be a representative sample of those who were displaced during the war. It is conceivable, for instance, that proximity to the battles or stress caused by forced migration might have led to higher mortality rates among the displaced than the rest of population. On the other hand, the increased mobility caused by the displacement could have been reflected in increased propensity to emigrate, in particular given that a large share of the cohort we study left Finland during the 1950s and 1960s.⁹ Furthermore, one might argue that those with the weakest earnings potential may have been the most likely to die or emigrate. Note that if attrition rates were higher among the displaced and if those leaving the sample were negatively selected, our baseline estimates would be biased upwards.

We address this question in two ways. First, we study attrition rates by 1939 residence municipality. As discussed in more detail in the Appendix, we do this by comparing the number of individuals in the microdata to pre-war data on the population of municipalities. Table 3.6.1 reports the results from regressing number of individuals observed in 1970 in our microdata by their 1939 residence municipality on total pre-war population of this municipality, a dummy indicating that the municipality was ceded after the

⁹According to Statistics Finland, net outmigration between 1945 and 1970 was some 270,000 individuals. By far the most important destination was the neighboring Sweden.

TABLE 3.6.1. Attrition

	$y = \log \# \text{observations, 1970}$			
	(1)	(2)	(3)	(4)
Ceded municipality	0.026 (0.114)	-0.002 (0.109)	-0.011 (0.109)	0.039 (0.139)
log Population, 1939	1.090 (0.045)	1.032 (0.044)	1.027 (0.044)	1.033 (0.046)

Note: Differences in attrition rates. The coefficients correspond to β (first row) and θ (second row) of equation (3.3), see the Appendix. Specifications: (1) no additional covariates, (2) controlling for log mean age, (3) controlling for age and share of Swedish-speaking population, (4) controlling for age, Swedish speaking population, longitude and latitude. All control variables are measured in the level of 1939 residence municipality.

war and some pre-war municipality characteristics. The results suggest that there was no systematic differences in attrition rates between the displaced and non-displaced. We acknowledge, however, that the estimates are rather imprecise.

Our second approach is to ask how the results would change *if* important non-random attrition was present. Specifically, we consider the hypothetical case where (a) attrition among the displaced was five percentage points higher than among the non-displaced, and (b) this attrition was extremely negatively selected in a sense that those with the lowest earnings capacity are not present in 1970. To be clear, we do not claim that such attrition took place. Rather, the idea is to consider a scenario that would be particularly worrying for our key conclusions.

Under the assumption that this type of attrition occurred, consistent estimates can be obtained by running the regressions on a sample omitting the lowest five percent of the non-displaced.¹⁰ The second row of Table 3.6.2 presents these estimates for key outcome variables. As expected, the figures are smaller than the baseline estimates reproduced at the first row. Yet, we still find that displacement increased long-term income among men.

¹⁰See Angrist et al. (2006) for formal discussion. The key insight is that under the assumption of monotone treatment response – that becoming displaced never decreased the likelihood to die or to emigrate before 1970 – dropping the lower tail of the control distribution and running OLS for the remaining sample provides an estimate of lower bound of the treatment effect. Similarly, the baseline estimates provide an upper bound.

TABLE 3.6.2. Robustness Checks

Specification	Imputed Income		Income		log Income	
	Men (1)	Women (2)	Men (3)	Women (4)	Men (5)	Women (6)
Baseline	0.10*** (0.02)	0.11*** (0.03)	2.64*** (0.99)	-0.02 (0.48)	0.24*** (0.04)	-0.01 (0.05)
Assuming selected attrition	0.06*** (0.02)	0.06*** (0.02)	1.75* (1.00)	-0.34 (0.50)	0.20*** (0.04)	-0.13*** (0.05)
Inflating with local prices	0.13*** (0.03)	0.14*** (0.03)	3.08*** (0.92)	0.22 (0.46)	0.26*** (0.04)	0.03 (0.04)
Conditional on 1950 Education	0.10*** (0.02)	0.11*** (0.03)	3.56*** (0.73)	0.16 (0.47)	0.25*** (0.04)	0.00 (0.05)
Conditional on 1950 Wealth	0.09*** (0.02)	0.09*** (0.02)	2.51*** (0.92)	-0.04 (0.44)	0.22*** (0.04)	-0.01 (0.04)
Excluding Agri- culture 1970	0.10*** (0.03)	0.07*** (0.02)	1.99* (1.17)	-0.67 (0.58)	0.22*** (0.05)	-0.04 (0.05)
Alternative area (187km)	0.07** (0.03)	0.07* (0.04)	3.55*** (1.14)	-0.03 (0.49)	0.21*** (0.06)	-0.03 (0.07)
Alternative area (50km)	0.09** (0.04)	0.06* (0.04)	3.68** (1.44)	-0.03 (0.54)	0.20*** (0.07)	0.09 (0.08)

Note: Impact of being displaced. OLS estimates, standard errors (in parentheses) robust to clustering at 1939 residence municipality level. Each coefficient comes from a different regressions. Each specification controls for age, imputed earnings in 1939, latitude, longitude and log taxable income per capita of the 1939 residence municipality, and indicator variables for speaking Swedish as ones mother tongue and living in city or market town in 1939. Additional controls: four categories of 1950 education (fourth specification); 17 dummies for the amount of agricultural land owned, rooms per resident and dummies for living in owner occupied premises and having a servant (fifth specification); ***, **, * indicate statistical significance at 1%, 5%, 10% level, respectively.

In short, we find no evidence of non-random attrition. Furthermore, we show that even if large differences in attrition rates were present and attrition would be extremely negatively selected, the results would remain qualitatively intact, at least for men. Thus we conclude that our main findings are not driven by non-random attrition.

3.6.2. Regional Price Variation. We next turn to the question whether the increase in nominal income among the displaced translates into an increase in consumption possibilities. This question is motivated by the notion that the displaced were more likely to move to urban areas. Thus, at least

part of the gains in nominal income could have been lost due to higher prices, in particular for housing and part of the surplus created by migration would be transferred to landlords in the areas attracting many displaced.

However, this does not seem to be an important part of the story. The third row of Table 3.6.2 reports estimates for the impact on real income, which is obtained by dividing the monetary outcomes measured in 1971 by the local price index.¹¹ Interestingly, we find that the impact on real income was even slightly larger than the impact on nominal income. This finding is a bit surprising, but partly reflects the fact that the most expensive rural municipalities are in the north, not in Southern Finland where most displaced were settled.

3.6.3. Conditioning on Other Post-War Characteristics. We next study whether our main results are affected by conditioning on further observable characteristics. One unfortunate feature of the data is that they do not contain information about pre-war educational attainment. However, we have a good measure for level of education in 1950. Given that forced migration could affect the availability of schooling or the decision to acquire education, we chose not to control for 1950 education in our baseline specification. However, as the results reported in Table 3.6.2 show, conditioning on education has no effect on the estimates – cf. the first- and fourth-row estimates. These results strongly imply that differences in education are not driving our results.

Another possibility is studied in the third row of Table 3.6.2. The purpose of these estimates is to address a hypothesis that wealth effects give raise to the higher 1971 income among the displaced. That is, since the settlement policy provided only partial compensation for wealth lost, higher earnings among the displaced could result from the displaced working harder in order to recover from this shock rather than their higher productivity. Studying the relevance of this possibility is difficult because of the limited amount of information on wealth available in the data. Yet, some observed variables are reasonable proxies for wealth. Most importantly, we observe the amount of agricultural land owned by the household in 1950. Other available wealth

¹¹These price data were collected in 1971 cost of living study that collected prices of 135 items from each municipality. A local price index was calculated based on consumption shares estimated in the 1966 Household Budget Survey. According to this study variation in local price level was mainly due to differences in (quality adjusted) rental prices. (Statistics Finland, 1972)

proxies are the number of rooms per persons in the residence and dummies for living in an owner-occupied dwelling and employing a servant, all measured in 1950. Conditioning on these variables has little impact on the estimates, suggesting that differences in post-war wealth, at least to the extent that we observe them, are not driving the results.

In addition to these alternative specifications, we have experimented with controlling for alternative industry/occupation definitions in the specifications leading to the estimates reported in columns (6) in Tables 3.5.2, 3.5.3, 3.5.4 and 3.5.5. The estimates reported in the tables use the 1939 industry coding, which is constructed from the richer categorization available only for 1970. However, it makes no difference whether we control for the baseline 54 industry categories or for the 274 categories available in the 1970 coding. Similarly conditioning on the 312 occupational categories available for 1970 (or for both the occupation and industry) has little impact on the estimates.

3.6.4. Measurement Issues. One possible objection to the results concerning the income effect is that taxable income may be a poor measure of true consumption possibilities or productivity. In particular, one could hypothesize that an important part of consumption among farmers would consist of consuming part of their own harvest and thus would not be recorded as part of taxable income. Since the displaced were more likely to move from agriculture to other sectors, such under-measurement of agricultural income would over-estimate the impact of the displacement. However, as reported in the fourth row of Table 3.6.2, running the regressions for a sample excluding those working in agriculture in 1970 yields estimates that are virtually identical to those for the whole sample. Thus, systematic measurement error, at least in this dimension, does not account for our results.

3.6.5. Alternative Control Areas. So far we have studied the impact of the displacement comparing the displaced to the entire population. In principle, one could argue that people living far away from the ceded area would systematically differ from those living close to the ceded area and thus the former should not be included in the analysis. In the two last rows of Table 3.6.2 we checked whether the estimates are sensitive to the choice of the comparison region by including in the sample only those who lived in 1939 close to the post-war border. As the ceded areas extended 187 kilometers east from the post-war border we first limit the comparison region to municipalities that were no further than 187 kilometers west from the

border. In the bottom row we limit both treatment and comparison groups to those living within 50 kilometers from the new border. The point estimates change very little, but the restrictions do affect the estimated standard errors.

In short, the data strongly suggests that forced migration increased the earnings of displaced men. For women, the results are somewhat mixed. However, the estimates are remarkably stable across various specifications and alternative sample selection criteria. We next turn to discuss more informally the implications of these results.

3.7. Discussion

Our results show that forced migration increased the long-term economic outcomes of the displaced persons. Thus, at a minimum, the Finnish experience in settling 430,000 displaced persons illustrates that successful resettlement policy is possible. However, the finding that the displaced did significantly better than the rest of the population calls for explanations beyond resettlement policies. We next suggest some possibilities that are consistent with our empirical findings. These explanations are inevitably more speculative than the evidence presented in the previous sections.

The first part of our explanation is based on the idiosyncratic factors related to a specific historical period. The post-war years in Finland were an era of rapid industrialization. While half of the persons in our sample were employed in agriculture in 1939, this share had declined to a third among the non-displaced by 1970. The change was considerably faster among the displaced. Even though the displaced farmers were settled in farming communities and given an opportunity to continue farming, their farms were small and often provided low standards of living. In addition, the displaced had fewer non-pecuniary reasons to stay in their assigned settlement areas. The resulting faster transformation from agrarian to modern occupations partially explains higher long-term income among the displaced.

However, we find that the displaced also earned more than those living in similar regions and employed in similar occupations after the war. This surprising result is substantially harder to explain. In our view, the most plausible explanation again has to do with lower mobility costs and the resulting higher mobility across regions and sectors. As shown above, the displaced were less likely to live in owner-occupied housing and were more likely to move not only from ceded areas to the rest of Finland, but also from the initial placement areas to other regions. A substantial fraction of these

moves is likely to be a response to better employment opportunities. Higher mobility may have led to better matching between workers and jobs and hence to higher earnings among the displaced even within a labor market. As the displaced moved from the agrarian sector to industrial occupations before others, they also had a first-mover advantage and more time to find the best matches. Furthermore, earnings growth or return to tenure may have been higher in the modern sector. Since the displaced moved to modern occupations earlier, they simply may have enjoyed faster earnings growth for more years than similar non-displaced persons. Unfortunately, we do not have data on earnings in 1950, or indeed any of the intervening years, and thus we are unable to examine this hypothesis in detail.

The fact that women appear to have experienced much smaller gains from displacement warrants some discussion. If the economic advantages from early moves are driven by better labor matches, married women may not have been able to exploit those as effectively if the main bread-winner was the male. Thus, if women were more likely the secondary bread-winner in the family, moves were more likely driven by the better labor market matches of their husbands.

The final question concerns the general lessons this particular displacement provides for public policy. In our view, the results inform at least two policy debates. First, our findings provide some guidance on the question on how displaced persons should be helped. The resettlement policy studied here was quite generous, providing both land and monetary compensation for lost property. The displaced farmers had the option of not taking up the offered land and were free to sell it – and many did. The displaced urban population received monetary compensation and was free to choose where to live. Whether the later economic success of the displaced should be viewed as a result of increased mobility at a time of rapid structural change or as the result of a successful resettlement policy is not evident. Arguably, it is a bit of both. Yet, in the absence of a settlement plan and compensation for lost property, the displaced might have stayed long in the evacuation areas and formed an impoverished underclass. The settlement plan did not lock them into the traditional sector, but provided the means to start over.

Our results are also relevant for regional policy. We find that moving early into the modern sector – which often entailed geographic mobility – was financially rewarding for the displaced. The hypothesized explanation, their lower pecuniary and non-pecuniary moving costs, suggests that policies

aimed at lowering the costs of migration may be an effective way to reduce regional disparities. Thus, policies encouraging mobility, such as subsidizing voluntary moves or discontinuing tax subsidization of owner-occupied housing, might be more efficient than place-based policies that attempt to help people who stay in the economically declining regions.

3.8. Conclusions

Post World War II evacuation of the Eastern parts of Finland created an situation where 430,000 persons had to relocate to the remaining parts of the country. We have exploited this historical episode to study the impact of displacement on those who were forced to migrate. Our findings indicate that being displaced had significant positive effects on long-term economic outcomes.

Since we focus on migrants who had no choice but to move, the setup allows us to avoid the fundamental selectivity problems present in much of the previous research on the economic effects of migration. Thus, we provide a consistent estimate of returns to (forced) migration on an average person in the data. The results suggest that higher mobility enhanced the efficiency of resource allocation in post-war Finland.

Much migration consists of either voluntary moves or population displacements in very disintegrated countries. Yet, there are many important situations resembling the one studied here. Examples of such displacements include large-scale public infrastructure projects and conflict-induced displacements in many parts of the world, as well as displacements caused by natural disasters permanently turning some areas uninhabitable. These situations call for active settlement policies, possibly financed by rich countries. The Finnish experience of resettling the displaced and their subsequent economic success may hold useful lessons for dealing with these forced migrations.

References

- AGRAWALA, S., T. OTA, A. A. UDDIN, J. SMITH, AND M. V. AALST (2003): *Development and Climate Change in Bangladesh: Focus on Coastal Flooding and the Sundarbans*, OECD, Environment Directorate and Development Co-operation Directorate.
- ANGRIST, J., E. BETTINGER, AND M. KREMER (2006): "Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia," *American Economic Review*, 96, 847–862.
- BORJAS, G. (2001): "Does Immigration Grease the Wheels of the Labor Market?" *Brookings Papers on Economic Activity* 2001, 69–119.
- CZAIKA, M. AND K. KIS-KATOS (2007): "Civil Conflict and Displacement. Village-Level Determinants of Forced Migration in Aceh," HiCN Working Paper 32.
- GOTTSCHANG, T. (1987): "Economic Change, Disasters, and Migration: The Historical Case of Manchuria," *Economic Development and Cultural Change*, 35, 461.
- GREENWOOD, M. (1997): "Internal Migration in Developed Countries," in *Handbook of Population and Family Economics*, ed. by M. Rosenzweig and O. Stark, Amsterdam: Elsevier, vol. 1, 647–720.
- IBÁÑEZ, A. M. AND C. E. VÉLEZ (2008): "Civil Conflict and Forced Migration: The Micro Determinants and Welfare Losses of Displacement in Colombia," *World Development*, 36, 659–676.
- KONDYLIS, F. (2007): "Conflict Displacement and Labour Market Outcomes: Evidence from post-war Bosnia and Herzegovina," CEP Discussion Paper 778.
- (2008): "Agricultural Households and Conflict Displacement: Evidence from a Policy Intervention in Rwanda," *Forthcoming in Economic Development and Cultural Change*.
- LUCAS, R. (1997): "Internal Migration in Developing Countries," in *Handbook of Population and Family Economics*, ed. by M. Rosenzweig and O. Stark, Amsterdam: Elsevier, vol. 1, 721–798.
- PIHKALA, K. U. (1952): "The Land Settlement Program of Finland," *Land Economics*, 28, 147–159.
- SCHULTZ, T. (1971): "Rural-Urban Migration in Colombia," *Review of Economics and Statistics*, 53, 157–163.

- SJAASTAD, L. (1962): "The Costs and Returns of Human Migration," *The Journal of Political Economy*, 70, 80–93.
- STATISTICS FINLAND (1942): *IV Varallisuustilastoa: Tulo ja omaisuusverotilastoa vuodelta 1938*, Helsinki: Valtioneuvoston kirjapaino.
- (1953): *IV Varallisuustilastoa: Tulo ja omaisuusverotilastoa vuodelta 1950*, Helsinki: Valtioneuvoston kirjapaino.
- (1972): *Vuoden 1971 kuntien kalleustutkimus*, Studies 19, Statistics Finland.
- (1996): *The Finnish Longitudinal Census Data File*, Helsinki: Statistics Finland.
- TARKKA, J. (1988): "Paasikiven aika," in *Suomen Historia 8*, ed. by J. Tarkka, Espoo: Weilin ja Göös, 1–2.
- UNHCR (2008): *2007 Global Trends: Refugees, Asylum-seekers, Returnees, Internally Displaced and Stateless Persons*, UNHCR.
- WARIS, H., V. JYRKILÄ, K. RAITASUO, AND J. SHIPI (1952): *Siirtoväen sopeutuminen*, Otava.
- WORLD COMMISSION ON DAMS (2000): *Dams and Development: A New Framework for Decision-Making*, Earthscan Publications.

Appendix

Constructed Variables.

Imputed income. We construct imputed income variable for each individual using the coefficient estimates from a regression of log annual taxable income in 1971 on a full set of age, socioeconomic status and industry dummies for a sample aged 18–59 in 1971. In essence, this procedure assigns each industry–age–socioeconomic status combination in all years a value representing the expected income of persons of the same socioeconomic status working in that particular industry in 1970/71. The regressions are run separately for men and women.

While providing a useful summary measure, this procedure has several shortcomings. Most importantly, imputed income is not informative on whether there was intra-industry earnings differences between ceded areas and the rest of the country. Unfortunately, there is little we can do about this problem. Second, the wage structure may have changed substantially between 1939 and 1971. This can be addressed by using tables listing taxable earnings in 38 industry-occupation-socioeconomic status groups for

1950 (Statistics Finland, 1953) and 12 industry-living in urban area groups for 1939 (Statistics Finland, 1942).

Lives in urban area. Statistics Finland categorizes municipalities into cities, market towns and rural municipalities. Our definition of urban area is based on the pre-war category of cities augmented with two municipalities (Espoo and Vantaa) bordering Helsinki (the capital). The municipalities classified as urban are: Helsinki, Espoo, Vantaa, Viipuri (ceded), Tampere, Turku, Vaasa, Lahti, Oulu, Kuopio, Kotka, Kemi, Pori, Lappeenranta, Mikkeli, Rauma, Hämeenlinna, Jyväskylä, Kokkola, Savonlinna, Hanko, Porvoo, Kajaani, Pietarsaari, Joensuu, Hamina, Sortavala (ceded), Käkisalmi (ceded), Loviisa, Tammisaari, Iisalmi, Raahe, Uusikaupunki, Heinola, Kristiinankaupunki, Tornio, Kaskinen, Uusikaarlepyy and Naantali.

Taxable income per capita in the residence municipality. For 1939 and 1950, this variable is constructed by dividing the sum total of taxable income (*veroäyri*) by the number of residents in the municipality. For 1970, we calculate average income from the microdata using 1971 data on taxable income.

3.8.1. Attrition. The aim is to learn how the available pre-war population data relate to the number of individuals observed in our sample. This relationship is by definition

$$(3.1) \quad N_{j,70} = o_j S_j P_{j,39} \xi_j$$

where $N_{j,70}$ is the number of individuals observed in microdata in 1970 who lived in municipality j in 1939, o_j is the sampling rate for this municipality¹², S_j is the survival rate, $P_{j,39}$ is population in 1939 and ξ_j captures measurement error in $P_{j,39}$, which is due to our having restricted interest to particular age groups but the municipality data are for the entire population. We model the survival rate as

$$(3.2) \quad \ln S_j = \ln \bar{S}_j + \beta D_j + X_j \gamma + v_j$$

where $\ln \bar{S}_j$ is the logarithm of the mean survival rate among municipalities that remained part of Finland, D_j is a dummy indicating that the municipality was ceded after the war, X_j is vector of pre-war characteristics such as the age structure and v_j is an error term. Taking logs of (3.1) and substituting

¹²As the sampling was made by taking every tenth file, this will vary across municipalities

with (3.2), we get an estimation equation

$$(3.3) \quad \ln N_{j,70} = \alpha + \beta D_j + \theta \ln P_{j,39} + X_j \gamma + \epsilon_{j,70}$$

where $\alpha = \ln \bar{o} + \ln \bar{S}_j$, and $\epsilon_{j,70} = (\ln \bar{o} - \ln o_j) + v_j + \ln \xi_j$. Under the assumption that ϵ is uncorrelated with the covariates, parameter β is informative on whether attrition among the displaced differs from the rest of the population. Furthermore, the fact that (3.1) is an accounting relation also suggests a natural specification check: we should find that $\theta = 1$. If this does not hold, either $\ln S_j$ is wrongly specified or the error term is correlated with covariates.

Assimilation to a Welfare State: Labor Market Performance and Use of Social Benefits by Immigrants to Finland

Abstract. I study the assimilation of immigrants to the Finnish labor market and welfare system. The initial immigrant-native earnings gaps are large. While long-term immigrants experience rapid earnings growth, only men from OECD countries converge to natives' earnings. Earnings grow predominantly due to improving employment rates rather than wage growth. Earnings profiles for temporary immigrants are flat. Furthermore, direct study of the use of social benefits suggests that immigrants learn to use the welfare system gradually. In particular, non-OECD households substantially increase their use of social assistance during their first five years in the country despite simultaneously doubling their earnings.

JEL Classification: J61, J31, F22

Keywords: Immigrants, assimilation, welfare state, social benefits

4.1. Introduction

Immigration is among the most controversial policy topics in many countries. One of the key themes in the debate concerns immigrants' performance in the labor market and the consequent impact on public finances. The discussion is fueled by the fact that, at least initially, immigrants tend to have lower earnings and to receive more public assistance than natives.

A central, and to some extent unresolved, question is whether immigrants recover from their initial disadvantage. In a seminal paper, Chiswick (1978) argued that while immigrants to the United States earned significantly less than comparable natives upon arrival, they overtook natives in 10 to 15 years. Later studies have shown that while the earnings of immigrants grow faster than those of natives, Chiswick's early results were overly optimistic

due to biases created by changes in the cohort “quality” (Borjas, 1985) and non-random return migration (Hu, 2000; Lubotsky, 2007). Furthermore, studies focusing on other countries suggest that both the initial gaps and the assimilation profiles differ vastly across countries and time periods (see Borjas, 1994; Boeri et al., 2002; Pekkala, 2005, for surveys).

This paper contributes to the literature in three ways. First, it is the first study on assimilation of immigrants to Finland. While Finland may not be of great general interest per se, focusing on a country with short immigration history and generous welfare state provides useful insights. In particular, it adds to the emerging literature on the impact of labor market institutions on assimilation. In a recent paper, Antecol et al. (2006) suggest that compressed wage distribution and generous welfare benefits may force assimilation to occur through improving employment and to reduce the potential for wage growth. My results are consistent with this argument. This suggests that differences in labor market institutions – along with differences in immigration policy – may explain why immigrants’ labor market performance varies between countries. It also has an important implication on the distribution of the costs of assimilation. That is, if immigrants earn lower wages but have high employment rates, the costs fall primarily on immigrants themselves. When assimilation occurs through improving employment and unemployed immigrants are eligible for social benefits, the costs are split between immigrants and host country’s tax payers.

Second, I assess the extent of which lower earnings truly translate into higher social benefits. The answer is surprisingly complex. I find that while immigrants from OECD countries earn substantially less than comparable natives, they receive similar amounts of social benefits. While temporary immigrants have lower earnings than long-term immigrants, they receive less social benefits. Furthermore, among non-OECD households the propensity to receive means-tested social assistance increases during the first five years to Finland, in spite of rapidly increasing earnings. Eligibility criteria or other characteristics of the welfare system do not explain these findings. Rather, the results point towards the possibility that immigrants’ take-up rate is initially very low. However, over time they may assimilate also in the sense of learning to use the welfare system.

The final contribution is methodological. Most importantly, I show that the standard practice of using log earnings or log wages as a measure of immigrants labor market performance – and thus excluding those with zero

earnings – may yield severely distorted estimates. Since immigrants with lower earnings capacity enter the employed sample gradually, the standard approach may underestimate both the initial earnings gaps and the rate of earnings growth. Another methodological contribution is to study assimilation profiles of long- and short-term immigrants separately. This may be important, because incentives to invest in host-country specific human capital depend on the length of the period the immigrant plans to stay in the host-country (Dustmann, 1993). Thus earnings profiles for temporary migrants are likely to be flatter than those of permanent migrants. This is exactly what I find. Hence this paper highlights the possibility that non-random return migration biases conventional assimilation measures both due to temporary migrants differing from permanent migrants in their time-invariant unobservable characteristics and due to differing assimilation profiles.

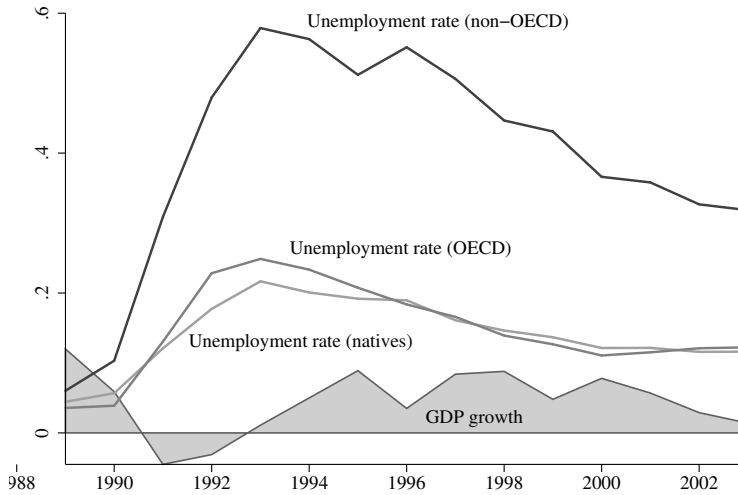
The paper proceeds as follows. The next Section discusses the institutional setting. Section 4.3 describes the data. Section 4.4 presents the empirical framework and Section 4.5 reports the results. Section 4.6 concludes.

4.2. Background

4.2.1. Immigration to Finland. The main reasons to immigrate to Finland have traditionally been either having a Finnish spouse or facing strong push factors in the home country. For the most part of its history, Finland has been characterized by emigration. The most intense period of emigration occurred in 1969–1970 when over 100,000 Finns (roughly 2% of the 1968 population) left the country, mainly to Sweden. In total, roughly 806,000 persons have emigrated during the period from 1945 to 2003, while net migration for the same period is some 224,000 negative.¹ As a result, early immigration to Finland consisted primarily of return migrants and their families. First refugees arrived in 1973, but the numbers remained small throughout the 1980s. The circumstances changed rapidly in the early 1990s as Finland simultaneously received a large flow of immigrants from the former Soviet Union and an increased number of refugees from the former Yugoslavia, Iran, Iraq and Somalia. The largest new immigrant group consisted of ethnic-Finns from the former Soviet Union, who were granted the

¹These figures are from Statistics Finland website (PX-Web Statfin, Table: “Väestönmuu-
tokset ja väkiluku 1749–2007”), visited in August 15th, 2008.

FIGURE 4.2.1. Unemployment Rate and GDP growth, 1989–2003



right to return in 1990.² Also relatively large group of Estonians – whose lingual proximity distinguishes them from other immigrants – migrated to Finland. As a result, the population share of immigrants raised from 0.4% to 2.4% between 1980 and 2003.

The period of rapidly increasing immigration coincided with a period of severe recession. Figure 4.2.1 plots the annual GDP growth and unemployment rates for natives and immigrants. By the first quarter of 1993, the GDP had shrunk by 14.4% from its peak of the first quarter of 1990. The situation was particularly harsh for the non-OECD immigrants, whose unemployment rate rose to 55 percent in 1996.

High unemployment rates of natives were also likely to contribute to a strict policy towards economic migrants. Throughout the analysis period of this study, current residents were given prioritized access to vacant posts. Before granting a work permit, the labor administration evaluated whether a worker is available “in a reasonable time”. In practice, the requirement seems to have been that no suitable EU/EEA citizen or an immigrant already in the country applied for the job in a few weeks.

²The evaluation of their Finnish ancestry was based on Soviet documents and it is not clear how much the ethnic-Finns differ from other Russian immigrants. For example, Ministry of Labour (1998) reports that in the in particular the younger cohorts of Soviet-born ethnic-Finns living in Finland have weak Finnish language skills.

4.2.2. Labor Market Institutions. The Finnish wage bargaining system is based on centralized negotiations between the labor unions and the employer organizations. Typically, union and employer federations negotiate a framework agreement, which is followed by industry level negotiations between individual unions and employer organizations. These collective labor contracts set minimum wages at job-complexity–education level. The agreements bind also non-union members, if more than half of the employees in the industry are members. Due to the high unionization rate, 95% of the employees are covered by these contracts.

Naturally, the collective agreements may be violated. The National Bureau of Investigation (NBI) estimates that annually about 20,000 employees are involved in the informal labor market, out of which 3,000–4,000 are immigrants (Adam and Laitinen, 2006). Most of these migrants are in the country legally, but violate employment or tax legislation. According to the NBI, there are only “a few hundred” migrants working in Finland without a work permission. However, it appears that the vast majority of the violations are related to foreign companies sending short-term posted workers to Finland.

4.2.3. Social Benefits. The Finnish welfare system provides a wide range of public services and high level of income security to all permanent residents (see the Appendix for details). For most benefits, eligibility does not depend on nationality or the residence permit status, but on living in Finland on “permanent basis”, defined as planning to stay for at least a year.³ In practice, one has to file applications to a magistrate and the Social Insurance Institution of Finland and the plausibility of the applicant’s plans are evaluated on an individual basis. Given that the analysis below is based on register data and observations from the first months since arrival are excluded, it seems safe to assume that almost everyone in the sample are fully eligible.

³The exceptions are student allowance (immigrants are eligible only if they have initially migrated to Finland for other reasons than to study) and pensions not tied to past employment (available only for immigrants who have stayed in Finland for more than five years). However, these transfers make up a negligible share of total transfers also for natives in the (25–60 year old) sample used in the analysis below.

4.3. Data

The analysis is based on individual-level panel data. Statistics Finland has put together these data by linking several administrative registers, including population register, tax register and register on social assistance maintained by the National Research and Development Centre for Welfare and Health (STAKES). The base sample contains annual observations of a 15% (2%) random sample of working age immigrants (natives) living in Finland in 1989 and a similar sample of new immigrants arriving to Finland (natives turning 15 years old) between 1990 and 2004. Each person is followed until the end of year 2004, emigration or death. Furthermore, the data include detailed information on the characteristics of a possible spouse and an indicator on whether the person is still living in Finland at the end of year 2005. Immigrants are defined as individuals born abroad, who do not speak Finnish as their native tongue and who enter the sample as non-citizens. Those still residing in Finland at the end of year 2005 are classified as “long-term” immigrants and the others as “temporary” immigrants.

The estimation sample is constructed as following. First, since complete benefit information is available only for 1993–2003, only data from these years are used. In order to focus on working age population, the analysis is further restricted to those born between 1944 and 1968. Everyone in this cohort had turned 25 year old by 1993 and were under 60 years of age in 2003. Finally, everyone arriving to Finland after year 2000, those who were less than 16 years of age at the time of arrival and those in the top percentile of the earnings and personal tax distributions are excluded.⁴ The final estimation sample consists of 55,927 observations for 6,949 immigrants and 407,929 observations for 38,419 natives.

Table 4.3.1 reports the sample means and standard deviations. It shows that immigrants are far more likely to live in urban, relatively low unemployment areas than natives. In most other respects, long-term immigrants from OECD countries closely resemble natives. In contrast, non-OECD immigrants differ substantially from the rest of the population. In particular, they are far less likely to be employed than natives. As a consequence, their

⁴The data include some extremely high levels of earnings and benefits. Most of these observations are likely to be either typing errors or atypically high earnings from bonus programs based on stock options (which the Finnish tax code regards as earnings). The main impact of exclusion the top 1% of the earnings and benefits distribution is a decrease in the standard errors, while impact on point estimates is small and qualitatively unimportant.

TABLE 4.3.1. Descriptive Statistics

	Long-Term Immigrants				Temporary Immigrants	
	Natives	OECD	non-OECD	OECD	non-OECD	
Age	42.2 (7.75)	40.5 (7.24)	39.0 (6.92)	37.7 (7.17)	36.8 (7.27)	
Age at arrival	(6.86)	(6.61)	(6.87)	(6.73)	(6.96)	
Women	0.50 (0.50)	0.36 (0.48)	0.53 (0.50)	0.37 (0.48)	0.40 (0.49)	
Single	0.29 (0.46)	0.21 (0.41)	0.28 (0.45)	0.29 (0.45)	0.35 (0.48)	
Has an immigrant spouse	0.01 (0.11)	0.09 (0.28)	0.44 (0.50)	0.20 (0.40)	0.35 (0.48)	
... a native spouse	0.69 (0.46)	0.70 (0.46)	0.28 (0.45)	0.50 (0.50)	0.29 (0.46)	
Single parent	0.05 (0.22)	0.03 (0.18)	0.09 (0.28)	0.02 (0.14)	0.05 (0.21)	
Number of children	0.92 (1.16)	1.12 (1.20)	1.12 (1.27)	0.75 (1.04)	0.80 (1.11)	
Lives in urban municipality	0.61 (0.49)	0.76 (0.43)	0.88 (0.33)	0.84 (0.36)	0.86 (0.35)	
... semi-urban municipality	0.18 (0.38)	0.09 (0.29)	0.07 (0.25)	0.08 (0.27)	0.06 (0.24)	
... rural municipality	0.21 (0.41)	0.15 (0.36)	0.06 (0.23)	0.08 (0.27)	0.08 (0.27)	
Local unemployment rate	16.2 (5.56)	12.9 (5.74)	13.5 (5.03)	15.3 (5.13)	15.0 (4.88)	
Owner occupied housing	0.72 (0.45)	0.56 (0.50)	0.25 (0.43)	0.44 (0.50)	0.24 (0.43)	
Months employed	8.7 (4.97)	7.5 (5.30)	4.4 (5.31)	4.3 (5.08)	3.0 (4.65)	
Zero Earnings	0.14 (0.34)	0.21 (0.41)	0.44 (0.50)	0.40 (0.49)	0.52 (0.50)	
Annual Earnings	19,176 (14,329)	17,105 (16,310)	8,097 (11,459)	10,631 (14,514)	6,172 (10,661)	
Annual Benefits	5,130 (5,814)	5,086 (5,776)	9,077 (7,473)	4,328 (5,552)	6,711 (7,046)	
Receives social assistance	0.09 (0.29)	0.10 (0.29)	0.33 (0.47)	0.12 (0.32)	0.24 (0.43)	
...housing allowance	0.07 (0.26)	0.06 (0.24)	0.32 (0.47)	0.06 (0.24)	0.22 (0.42)	
...unemployment benefits	0.23 (0.42)	0.22 (0.42)	0.54 (0.50)	0.23 (0.42)	0.37 (0.48)	
...other benefits	0.22 (0.41)	0.20 (0.40)	0.20 (0.40)	0.16 (0.37)	0.19 (0.39)	
Individuals	38,419	989	4,665	563	732	
Observations	407,929	9,063	41,836	1,981	3,047	

Note: Sample means and standard deviations (in parenthesis).

mean annual earnings are less than half of native earnings, while they receive almost twice as much benefits. Furthermore, comparisons between long-term and temporary immigrants reveal that the latter have lower earnings, but nevertheless receive less social benefits. More than a third of the OECD-born immigrants in the sample remigrate prior to the end of year 2005, while the share of temporary immigrants is 14% among non-OECD immigrants.

4.4. Empirical Framework

The primary interest of this paper is in the evolution of the economic performance of immigrants and in how their performance compares to that of comparable natives. Formally, the quantities of interest can be written as

$$(4.1) \quad g(k, x) = \mathbb{E}[y|I = 1, YSM = k, X = x] - \mathbb{E}[y|I = 0, X = x]$$

which measures the difference in expectation between an immigrant ($I = 1$), who has stayed in the host country for k years and a native ($I = 0$) with identical observable characteristics x . Immigrants are said to be assimilating, if $g(k, x)$ approaches zero as k increases. Since it is also informative to learn how immigrants perform in absolute terms, the expectations will be reported separately.

Estimation of these expectations typically follows the framework based on Chiswick (1978) and Borjas (1985). The estimation equation for immigrants is

$$(4.2) \quad y_{jt} = YSM_{jt}\alpha + \sum_m \beta_m C_{jm} + A_{jt}\delta^I + \sum_s \gamma_s^I \Pi_{js} + X_{jt}\phi^I + \epsilon_{jt}$$

where y_{jt} is the outcome of interest for person j at time t , YSM_{jt} is a vector of polynomials on the number of years immigrant has resided in the host country, C_{jm} is a vector of indicator variables for the year of arrival m , A_{jt} is a vector of polynomials of age, Π_{js} is a vector of indicator variables denoting the year of observation, X_{jt} is a set of control variables and ϵ_{jt} summarizes the impact of unobservable characteristics. Similarly, the estimation equation for natives is

$$(4.3) \quad y_{jrt} = A_{jt}\delta^N + \sum_s \gamma_s^N \Pi_{js} + X_{jt}\phi^N + \epsilon_{jt}$$

The first challenge in consistently estimating (4.2) and (4.3) arises from the fact that year of arrival, time in the host country and calendar year

are perfectly collinear. Hence, some restriction must be imposed in order to separately identify α , β , γ^I and γ^N . I follow Barth et al. (2004) and model the time effects as

$$(4.4) \quad \begin{aligned} \gamma_{rt} = & \gamma_t + \gamma_r + \eta_0 \ln u_{rt} + \eta_I (I \times \ln u_{rt}) \\ & + (YSM \times \ln u_{rt}) \eta_{YSM} + (A \times \ln u_{rt}) \eta_A \end{aligned}$$

where γ_t captures the time-effects common to all regions, γ_r is a set of regional fixed-effects, u_{rt} is the local unemployment rate and other variables are as above.⁵

The second problem arises from nonrandom return migration. Hu (2000); Lubotsky (2007) and Edin et al. (2000) show that the least successful immigrants are the most likely to leave, at least from the U.S and Sweden. As a result, the observed immigrant population becomes increasingly favorably self-selected over time and consequently the estimates of the rate of assimilation are biased upwards.

My approach is to estimate the assimilation profiles separately for temporary and long-term immigrants. Since there is no attrition from the population of long-term immigrants, this yields unbiased assimilation profiles for this sub-sample. Furthermore, separate regressions allow me to assess whether assimilation profiles of temporary immigrants are flatter than those of long-term immigrants. This seems likely for two reasons. First, some immigrants may leave the country due to negative spurious shocks, such as becoming unemployed. In other words, they leave because they fail to assimilate. Second, some migrations were planned to be short-term. The assimilation profiles of these immigrants are likely to be flat, since their incentives to invest in host country specific human capital are low (Dustmann, 1993). Since equation (4.2) includes an implicit assumption of common α , mixing temporary and long-term migrants would decrease informativeness of the estimates for both groups.

A drawback of the approach is that I only observe whether the immigrant left Finland prior to the end of year 2005. Thus I am forced to set an ad hoc criterion on who is considered to be a long-term immigrant. Clearly, some

⁵A more common identifying assumption is that aggregate economic conditions have similar impact on immigrants and natives ($\gamma_t^I = \gamma_t^N$). However, Bratsberg et al. (2006) and Barth et al. (2004) present evidence that wages of immigrants are more sensitive to local unemployment rates than those of natives in the U.S. and Norwegian labor markets. Thus estimates based on the common time effects assumptions are biased in the presence of a trend in unemployment during the observation period.

of these immigrants will emigrate later and may thus make lower human capital investments. The problem is mitigated by including only immigrants who arrived prior to January 2000 in the estimation sample. Perhaps more importantly, the long-term sample is likely to over-present “lucky” immigrants, since those who left due to unsuccessful assimilation end up in the sub-sample of the temporary immigrants. Thus the assimilation profiles of long-term immigrants could, in principle, be upward sloping due to human capital investment or due to the stayers being favorably self-selected with respect to their assimilation profiles. Whatever the reason, the estimates are nevertheless informative about the population of immigrants who ended up being long-term immigrants. On the other hand, the estimates for temporary immigrants still suffer from the attrition bias. That is, those who leave after ten years are likely to differ in their unobserved characteristics from those who leave after a year. To decrease this source of bias, I control for the number of years the temporary immigrant eventually stays in Finland.

The second selection issue concerns selection into employment when estimating wage equations. To see why this might pose a problem, suppose that immigrants accept (or are allowed to accept) wage offers only if they exceed some threshold. If wage offers are an increasing function of time in the host country, those with the most favorable unobserved characteristics are most likely to be employed upon arrival, while others become employed later. Thus the unobserved component and time in the host country would be negatively correlated and the estimates for both the initial wage gap and wage growth would be biased downwards. One possibility for correcting such selection bias would be using a control function framework. However, in my view, the data does not include any plausible exclusion restrictions. Thus my approach is to acknowledge the problem when interpreting wage assimilation profiles and to draw conclusions mainly from annual data – including zeros – on employment, earnings and benefits.

To illustrate the results, I calculate the two expectations in (4.1) for each immigrant observation in the data. I remove the impact of business cycle by setting local unemployment rate at 13.4% (mean of the immigrant sample) and year dummies to their means. Other variables are left as they are. This yields two sets of expectations for each immigrant. The first are expectations of her outcomes, had the general labor market conditions remained constant over time. The second are expectations for natives with identical observable

characteristics. The assimilation profiles reported in the next section are averages of these expectations over years in Finland.⁶

4.5. Results

This section reports the results from estimating the model discussed above using several outcomes and running the regressions separately for men and women, for long-term and temporary migrants and for those born in the OECD and non-OECD countries. Given the specification used, natives and immigrants are comparable in the sense that they are of similar age, live in similar labor markets and have similar family structure.⁷ However, it is important to note that they may – and are likely to – differ in factors such as education and relevant work experience. Omitting education from the specification is motivated by the practical reason of not observing education obtained abroad. On the other hand, even if the data would include accurate information on educational attainment, it is not clear whether one should condition on it as immigrants often study in their host countries. Thus the specification used below allows for a broader notion of assimilation that includes post-migration investments on formal education. In order to provide another meaningful comparison group, I also report profiles separately for low-skilled natives, defined as those 28% of the native sample who have less than a secondary degree.

4.5.1. Labor Market Performance. I begin by studying how immigrants' labor market performance evolves as they spend more time in Finland. Figure 4.5.1 presents annual earnings of immigrants and comparable natives over the first two decades in Finland. Earnings are measured in thousands of euros and observations with zero earnings are included. The results

⁶ $\mathbb{E}[y|I = 1, YSM = k, X = x]$ is calculated using estimates from equation (4.2) and $\mathbb{E}[y|I = 0, X = x]$ by using estimates of equation (4.3). Native profiles are comparable to the profiles of long-term immigrants. Profiles for temporary immigrants come from a separate regression and allow temporary and long-term immigrants to differ in their observable characteristic

⁷The covariates are a cubic of time in the host country and age, year of entry to Finland, indicator for being single, number of children under 18-years old, three indicators for children younger than 3, 7 or 18 years old living in the household and type of municipality of residence indicators (urban, semi-urban, rural). The specification also includes interactions with all these variables and immigrant status. Furthermore, local unemployment rate, year dummies and 20 region dummies enter the specification as shown in equation (4.4) and the specification for temporary migrants include a vector of indicators for the number of years the immigrant eventually lives in Finland.

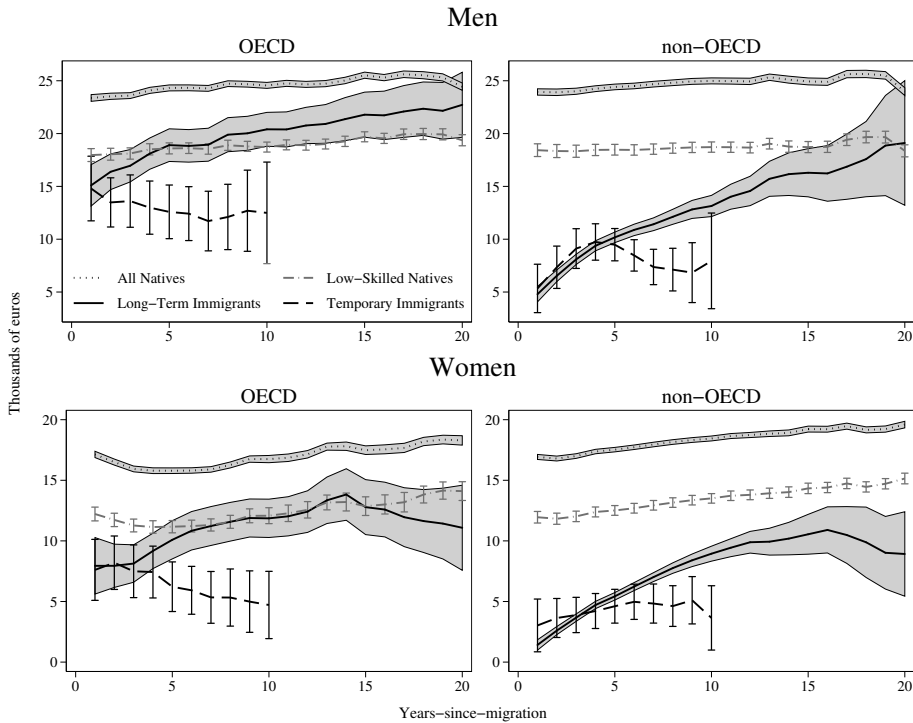
reveal a dramatic earnings gap between newly arrived immigrants and natives. According to the point estimates, non-OECD immigrants earn only 20% (men) and 8% (women) of the level of comparable natives during their first full year in Finland. The corresponding figures for OECD immigrants are 65% (men) and 46% (women).

While it is hard to find comparable estimates for other countries, these gaps appear to be among the largest in the literature.⁸ One can speculate on several possible reasons. For instance, compressed wage distribution and generous welfare state could attract negatively self-selected immigrants to Finland. Alternatively, large initial gaps are consistent with long-term immigrants initially investing heavily in the acquisition of Finland-specific human capital. The optimal investment may be larger than in many other countries, in particular because the Finnish language differs substantially from most European languages.

These explanations have different implications for the slope of the earnings profiles. If immigrants come to Finland simply to enjoy the welfare state, the gap between natives and immigrants should be permanent; if the human-capital investment explanation is valid, the gap should shrink as immigrants start to receive returns to their investments. Furthermore, given that temporary migrants have fewer incentives to invest in host-country specific human capital, their earnings profiles should be flatter than those of long-term immigrants.

⁸The key problems with comparability is that most studies report estimates for log earnings or log wages and thus omit individuals who are not employed. Essentially, they thus ask how large the earnings gap is among those immigrants and natives who have non-zero earnings, while this paper focuses on the earnings gaps between the entire native and immigrant populations. Further comparability problems arise due to some studies conditioning on education and most studying using repeated cross-sectional data. In order to get a more meaningful comparison, I have replicated the approach taken by Lubotsky (2007), who studies log earnings assimilation of immigrants to the U.S using longitudinal data. He reports (Lubotsky, 2007, Table 5) initial median earnings gaps ranging between .17 and .41 log points depending on the year of entry. Immigrants who have been the to U.S. for 11–15 years have .15 log points higher median earnings than otherwise similar immigrants who have been to the U.S. for 1–5 years. The corresponding estimates from Finnish data are an initial immigrant-native gap of .58 log points and .38 log points difference between those who have lived in Finland 11-15 years in comparison to recent arrivals. Thus the initial earnings gap is larger and the subsequent earnings growth faster in Finland than in the U.S. even when we study only the sub-population of immigrants who have non-zero earnings. Studies on immigrant assimilation to other Nordic countries include Barth et al. (2004) for Norway, Edin et al. (2000) for Sweden and Husted et al. (2001) for Denmark.

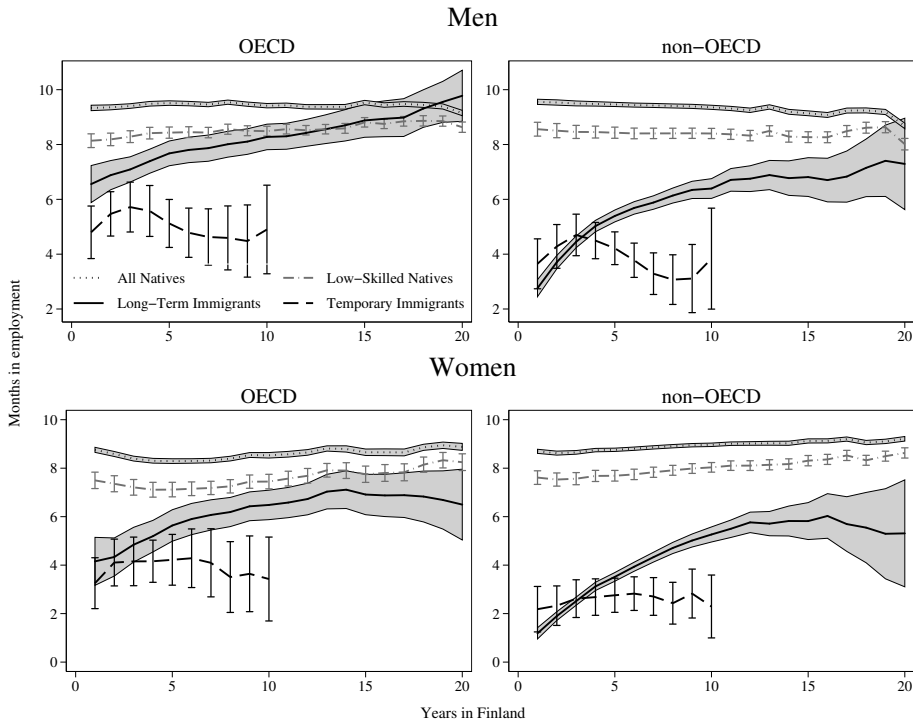
FIGURE 4.5.1. Earnings profiles



Note: Expected annual earnings and 95% confidence intervals over time in Finland for long-term immigrants (solid line), temporary immigrants (dashed line), comparable natives (dotted line) and comparable low-skilled natives (dotted gray line). Local unemployment rate fixed at 13.4%. Confidence intervals are robust to intra-individual autocorrelation.

Figure 4.5.1 is in line with the human-capital investment explanation. According to the point estimate, the earnings of long-term non-OECD immigrants grow a staggering 340% (men) and 870% (women) over the first 15 years in the country. Earnings growth for OECD immigrants is roughly 50% over the same period, while earnings of temporary immigrants remain almost constant. However, the earnings gap to natives is closed only by OECD men. For other groups, large immigrant-native earnings differences persist even after two decades in Finland. After roughly six years in the country, earnings profiles of OECD-born women become similar to those of low-skilled native women. For non-OECD men it takes two decades converge to the earnings level of low-skilled natives, while non-OECD women never reach even this comparison group.

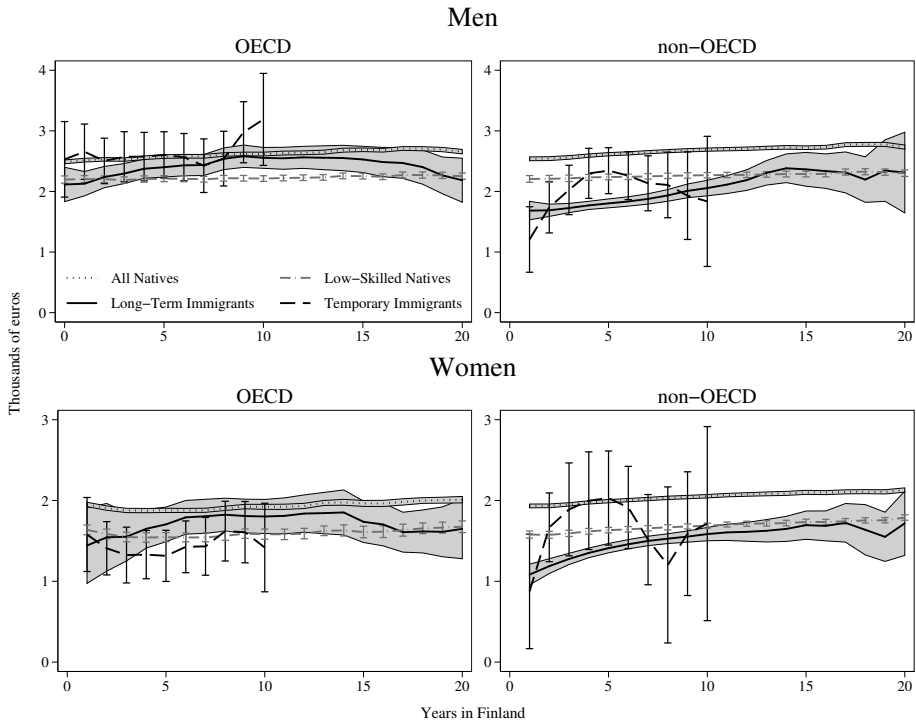
FIGURE 4.5.2. Employment profiles



Note: Expected months in employment (excluding subsidized work) and 95% confidence intervals over time in Finland for long-term immigrants (solid line), temporary immigrants (dashed line), comparable natives (dotted line) and comparable low-skilled natives (dotted gray line). Local unemployment rate fixed at 13.4%. Confidence intervals are robust to intra-individual autocorrelation.

I next turn to the sources of earnings growth. Note that earnings reflect both wages or entrepreneurial income when working and the hours worked. Figures 4.5.2 and 4.5.3 plot profiles for months employed during a year and monthly earnings. As discussed in the previous Section, non-random selection to employment is likely to bias the monthly earnings profiles downwards. Nevertheless, it seems fair to conclude that the increase in employment rather than wage growth is driving earnings assimilation. For OECD immigrants, monthly wages are similar to those of comparable natives and remain almost constant throughout the observation period. In contrast, expected months in employment increase by three months during the first 15 years in the country and, among men, reach the level of natives. Non-OECD immigrants

FIGURE 4.5.3. Monthly Earnings



Note: Expected monthly earnings and 95% confidence intervals over time in Finland for long-term immigrants (solid line), temporary immigrants (dashed line), comparable natives (dotted line) and comparable low-skilled natives (dotted gray line). Local unemployment rate fixed at 13.4%. Confidence intervals are robust to intra-individual autocorrelation.

assimilate in both the employment and wage dimensions. However, a back-of-the-envelope calculations, similar to those by Antecol et al. (2006), suggest that roughly 90% of the earnings growth among men and 85% among women can be attributed to increase in employment.

Comparing Figures 4.5.1 and 4.5.3 also reveals a simple, but important methodological point. Most studies on immigrant assimilation use either log earnings or log wages as the outcome variable and thus exclude those who are not employed from the estimation sample. At least in the Finnish case, the standard approach would lead to severe underestimation of both the initial earnings gap and the rate of assimilation. For instance, recall that the estimates above suggested that natives earn five times more than newly arrived non-OECD men. Using log annual earning or level of monthly

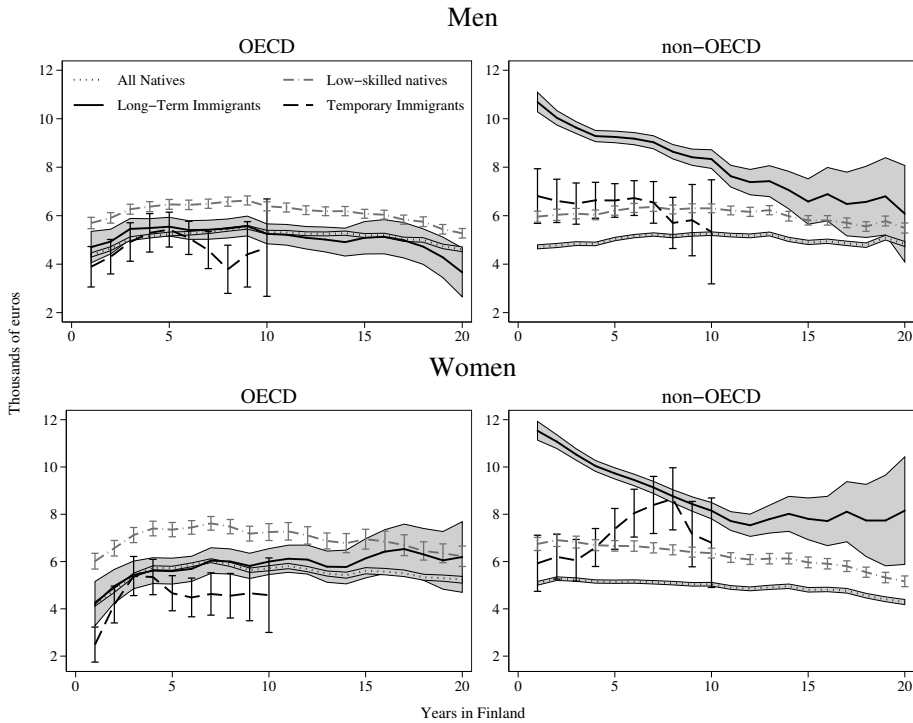
earnings would lead to a conclusion that the difference is only 3.5 fold or 1.5 fold, respectively. Furthermore, instead of concluding that the earnings of non-OECD men grow 3.4 fold over the first 15 years in Finland, estimates based on log annual earnings and level of monthly earnings would suggest that the growth is 2.1 fold and 1.5 fold, respectively. This is not surprising, of course, as the assimilation process is driven by increasing employment. Thus those with lowest earnings capacity are likely to have zero earnings upon arrival and enter to the lower part of the conditional-on-positive earnings distribution over time in the host country.

4.5.2. Social Benefits. Given these results, one would expect immigrants to collect considerably more social benefits than comparable natives upon arrival and this gap to decrease as immigrants' earnings increase over time. Figure 4.5.4 takes a direct look at the issue by presenting the benefit profiles. Since many benefits depend on total household income, the profiles are calculated at the household-level and "immigrants" are defined as a household where the adult male (top row) or female (bottom row) is an immigrant regardless of the immigrant status of the spouse.

Upon arrival, non-OECD households receive more than twice the benefits of comparable native households. While the gap decreases over time, the difference appears to be permanent. Surprisingly, however, OECD households receive the same amount of social benefits as native households throughout the observation period. This occurs in spite of their initially lower earnings and later earnings converge. Furthermore, among those born outside the OECD countries, temporary immigrants earn similar or lower earnings as long-term immigrants, but receive substantially less benefits.

Turning to the participation rates of four subcategories of benefits, presented in Figure 4.5.5, reveals a third surprising result: non-OECD households substantially increase their use of social assistance over the first five years in the country despite of their rapid earnings growth and decline in the use of other means-tested benefits. To the best of my knowledge, the institutional setting does not explain this finding. As discussed in Section 4.2.3, eligibility to the social benefits is only conditional on living in Finland on permanent basis and this requirement is fulfilled by virtually everyone in the sample. One possible explanation with regard to the rising social assistance rates could be that immigrants switch from receiving unemployment benefits to collecting social assistance. Such benefits substitution could be motivated,

FIGURE 4.5.4. Benefit profiles

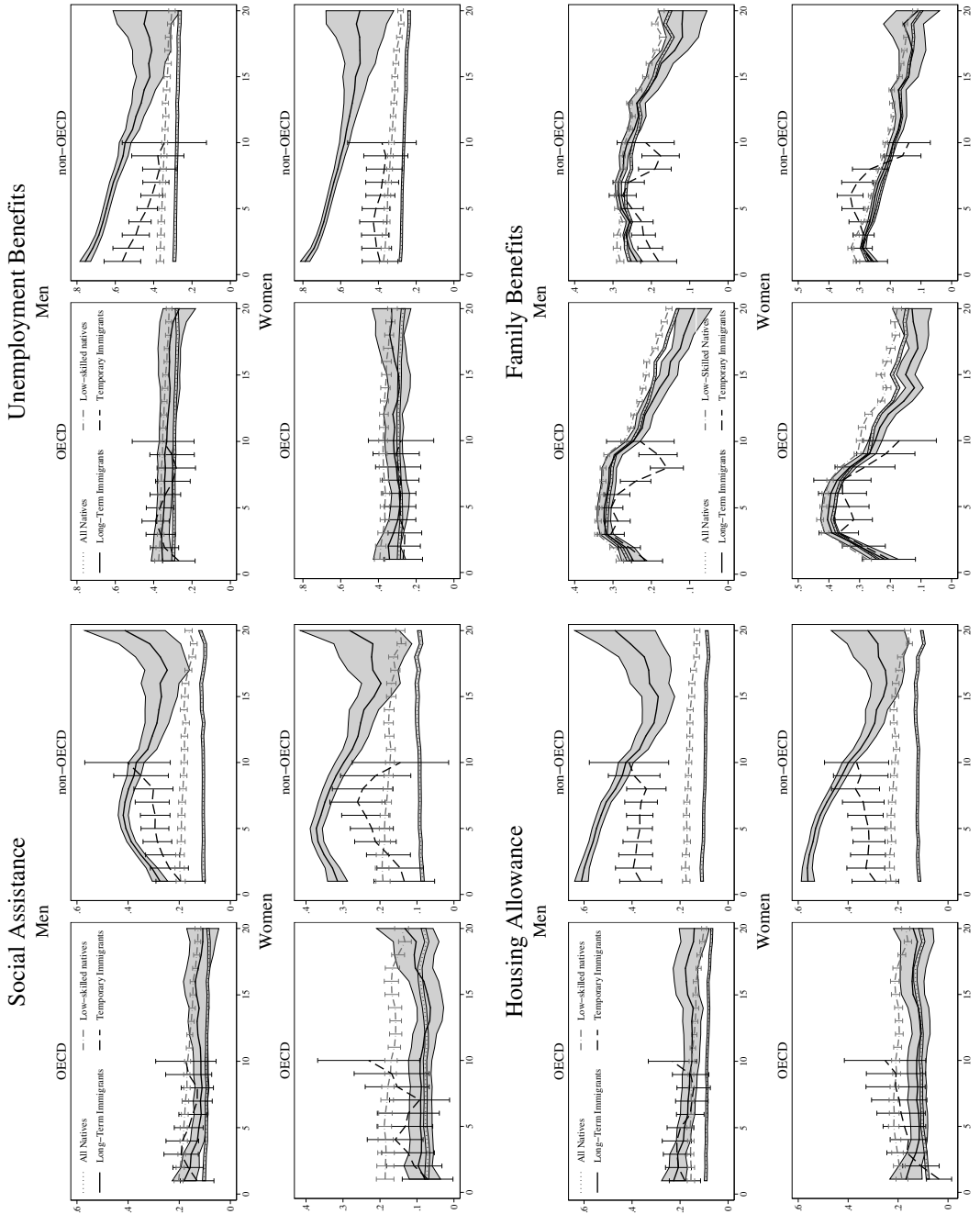


Note: Expected annual benefits and 95% confidence intervals over time in Finland for long-term immigrants (solid line), temporary immigrants (dashed line), comparable natives (dotted line) and comparable low-skilled natives (dotted gray line). Local unemployment rate fixed at 13.4%. Confidence intervals are robust to intra-individual autocorrelation.

for example, by unemployment benefits being conditional on participation to labor market training. Alternatively, administrative practice could be to pay essentially same benefits under different names at different stages of the assimilation process. Yet, social assistance seems to complement rather than substitute other benefits. That is, 78% of the immigrants who collect social assistance also receive unemployment benefits. Furthermore, only 9% of those who start to collect social assistance stop receiving unemployment benefits during the first two years on social assistance. Looking at the whole immigrant population, just 0.5% switches from receiving only unemployment benefits to collecting only social assistance during any three year period.

Thus the results seem to point towards the possibility that immigrants learn to use the welfare system as they spend more time in the host country (Borjas and Hilton, 1996). In other words, upon arrival they may simply

FIGURE 4.5.5. Participation Rates



be unaware of their eligibility to social assistance and hence their initial take-up rates may be low.⁹ Furthermore, the distinguishing feature of social assistance in comparison to other benefits is that the case-workers are able to apply some discretion when deciding on eligibility and the amount. Hence, spending time in the host country could increase both the information of the available benefits and the ability to talk to the case-workers.

4.6. Conclusions

Ever since Chiswick's (1978) study, a vast number of papers have assessed the labor market assimilation of immigrants. In an influential review article, Borjas (1994, p. 1671) summarized the motivation behind this literature as following: "These studies view the labor market performance of immigrants in the host country as a measure of the immigrant contribution to the economy's skill endowment and productivity. In addition, the trends in immigrant skills help determine the impact of immigration on the employment opportunities of native-born workers and on expenditures in social insurance programs".

The results presented above suggest that the contribution of immigrants to Finland is, at best, modest. The initial earnings gaps are large and while immigrants' earnings grow rapidly, only men from OECD-countries reach natives. Earnings of women from non-OECD countries do not converge even to the level of low-skilled natives.

In a sense, these findings are not surprising given that I study a country that has allowed primarily non-economic immigrants to enter. Furthermore, Finland has a short immigration history and a generous welfare system placing no restrictions on the eligibility of immigrants. Furthermore, many immigrants arrived during an unusually severe recession. Thus one may be willing to consider these results as something close to a worse case scenario. In particular, given that Finland appears to be moving towards an immigration policy designed to attract economic immigrants, there is no reason

⁹Previous research suggests that take-up rates for social assistance among natives are around 50–60% (Bargain et al., 2007). Unfortunately, the available data does not allow for direct study of take-up rates among immigrants. As an indirect way to assess take-up rates, I have experimented with regressing participation on social assistance on the same covariates as above plus a vector of dummies on household's income from other sources. Results from these regressions suggest that keeping household characteristics and income constant, the use of social assistance increases over the first six years to Finland and remains roughly constant thereafter.

to expect that future immigrants to Finland would fare as badly as those studied above.

The results also provide three more general lessons. First, they lend further support to the hypothesis that compressed wage distribution may force assimilation to take place in the employment dimension. As a consequence, countries where low wage jobs are not available should expect their tax-payers to bear a larger share of the costs associated with labor market assimilation.

Second, the estimates suggest that the relationship between immigrants' earnings and the cost they place on social insurance programs is not as straightforward as the quotation above seems to imply. Rather the findings point towards under-utilization of the welfare system among newly arrived immigrants. As a consequence, both earnings and the use of social benefits could, in principle, increase as immigrants become more accustomed to the host country.

Finally, I have illustrated some weaknesses in the standard methodology of estimating assimilation profiles. The results provide further evidence that neglecting non-random return migration leads to upward biased estimates of the rate of assimilation. Unfortunately, avoiding this bias requires longitudinal data and such data are rare. However, the second methodological point is simple and easily executed with typical datasets. That is, measuring immigrants' labor market performance with log earnings or hourly wages may lead to wrong conclusions. Neglecting non-random selection into employment is likely to bias estimates of immigrants' initial earnings upwards and estimates of their earnings growth downwards. Thus using the level of annual earnings and including those with zero earnings to the estimation sample should provide a more informative picture of the assimilation process than the current practice.

References

- ADAM, C. AND L. LAITINEN (2006): "Illegal Employment of Third Country Nationals in Finland," Working Paper. IOM Helsinki.
- ANTECOL, H., P. KUHN, AND S. TREJO (2006): "Assimilation via Prices or Quantities? Sources of Immigrant Earnings Growth in Australia, Canada, and the United States," *Journal of Human Resources*, 41, 821–840.
- BARGAIN, O., H. IMMERVOLL, AND H. VIITAMÄKI (2007): "How Tight Are Safety-Nets in Nordic Countries? Evidence from Finnish Register Data," IZA DP No. 3004.
- BARTH, E., B. BRATSBERG, AND O. RAAUM (2004): "Identifying Earnings Assimilation of Immigrants under Changing Macroeconomic Conditions," *Scandinavian Journal of Economics*, 106, 1–22.
- BOERI, T., G. HANSON, AND B. MCCORMICK (2002): *Immigration Policy and the Welfare System*, Oxford: Oxford University Press.
- BORJAS, G. (1985): "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants," *Journal of Labor Economics*, 3, 463–489.
- (1994): "The Economics of Immigration," *Journal of Economic Literature*, 32, 1667–1717.
- BORJAS, G. AND L. HILTON (1996): "Immigration and the Welfare State: Immigrant Participation in Means-Tested Entitlement Programs," *Quarterly Journal of Economics*, 111, 575–604.
- BRATSBERG, B., E. BARTH, AND O. RAAUM (2006): "Local Unemployment and the Relative Wages of Immigrants: Evidence from the Current Population Surveys," *The Review of Economics and Statistics*, 88, 243–263.
- CHISWICK, B. (1978): "The Effect of Americanization on the Earnings of Foreign-born Men," *Journal of Political Economy*, 86, 897–921.
- DUSTMANN, C. (1993): "Earnings adjustment of temporary migrants," *Journal of Population Economics*, 6, 153–168.
- EDIN, P., R. LALONDE, AND O. ÅSLUND (2000): "Emigration of Immigrants and Measures of Immigrant Assimilation: Evidence from Sweden," *Swedish Economic Policy Review*, 7, 163–204.
- HU, W. (2000): "Immigrant Earnings Assimilation: Estimates from Longitudinal Data," *American Economic Review*, 90, 368–372.
- HUSTED, L., H. NIELSEN, M. ROSHOLM, AND N. SMITH (2001): "Employment and Wage Assimilation of Male First-generation Immigrants in

- Denmark,” *International Journal of Manpower*, 22, 39–68.
- LUBOTSKY, D. (2007): “Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings,” *Journal of Political Economy*, 115, 820–867.
- MINISTRY OF LABOUR (1998): “Inkerinsuomalaisten maahanmuutto Suomeen 1990-luvulla,” Ministry of Labour.
- PEKKALA, S. (2005): “Economic Impacts of Immigration: A Survey,” Discussion Paper 362. Government Institute for Economic Research.

Appendix

The Social Benefits System. Unemployment benefits are paid through a three-tier system. First, all registered unemployed who have worked for at least 43 weeks during the past 28 months receive a basic benefit. On top of this, those who have contributed to a voluntary unemployment insurance fund for at least ten months prior to the claim, receive an earnings related benefit. These benefits are paid for maximum of 500 work days. Those unemployed who do not meet the employment condition or who have received the basic allowance for the maximum period are eligible for labor market subsidy (LMS). The subsidy is means-tested and has no maximum duration. The amount of LMS is equal to or less than the basic unemployment benefit. The income-test is applied to the household income and the benefit gradually decreases as household income increases. After a cut-off point, no subsidy is paid. However, if the person participates in a labor market policy measure, LMS is paid without means-testing.

Social assistance is a residual benefit which acts as a last resort of economic assistance. It is means-tested based on household's expenses, income (including other forms of transfers) and assets. The transfer consists of a fixed basic amount and an additional allowance. When these expenses exceed household's net income, social assistance makes up the difference. The basic amount is designed to cover the costs of food, clothes, hygiene, transport, newspaper, telephone, TV license, minor health care costs and 7% of housing costs. The additional allowance is designed to cover other cost such as rest of the "reasonable" housing costs, children's day care fees and large health care costs. Some municipalities also grant one-off supplements. These may be related to special situations (sickness, change in family circumstances and other "life events"); for supporting re-integration or rehabilitation measures; or as a preventive measure (e.g. support for job-search, to cope with sudden problems due to debts).

Housing allowance covers up to 80% of acceptable housing costs for low-income households. The definition of acceptable housing expenditure depends on the size of housing, geographical location, construction year and heating system. The allowance decreases as household income increases and stops after an upper threshold. Single students receive a separate housing allowance (see below).

Student allowances consists of a study grant, housing allowance and government guarantees for student loans. It is granted based on mechanical rules depending on the level of education, age, marital status, mode of accommodation and other income. Non-citizens arriving to Finland to study are not eligible, but those who have migrated to Finland on other purpose than studying are eligible. Student allowances are available for the maximum of 55 months.

Pensions are paid through two system: earnings-related pensions related to past employment and national pensions related to residence in Finland. All employees and self-employed are covered by the mandatory earnings-related pension insurance. The amount of national pension is determined by household's other pension income and becomes zero after a cut-off point. Both systems include a wide range of retirement benefits for individuals below the formal retirement age such as disability pension, early retirement pension and unemployment pension. In addition, the earnings-related pension system has provisions for partial disability and part-time pensions. Many of these pensions have eligibility rules that are hard to fulfill for immigrants. For example, the amount of national pensions is affected by the length of residence in Finland. However, a Special Assistance for Immigrants is available for immigrants, who have lived in Finland for at least five years. This, in effect, provides old-age and disability national pensions to those who are ineligible for a standard national pension.

Families with children receive a variety of cash-transfers. First, mothers are entitled to maternity allowance for 105 working days. This is followed by parental allowance for 158 working days to the parent who takes leave from work. The amounts are proportional to pre-parental-leave earnings. Second, home care allowance is available to families with children under the age of three, on the condition that the children do not use public day care. The allowance consists of care allowance and an income-tested home care supplement. Some municipalities also grant special municipal supplements. After the child turns three, private day care allowance is available for families with children under the age of seven, who participate in private day care. Again, the allowance consists of a basic part, of an income-tested supplement and sometimes of a special municipality supplement. On top of these, every family with children under the age of 17 receives family allowance. The amount of this allowance depends on the number of children and whether both parents are present in the household, but not on household income.

Moving Immigrants from Welfare to Work: The Impact of an Integration Program

Abstract.¹ We study the impact of an integration program for immigrants using a fuzzy regression-discontinuity approach. The program consists of an individualized sequences of active labor market policy measures combined with sanctions in the case of non-compliance. The program was introduced in 1999 in Finland, but only those who had entered the population register after May 1997 had an obligation to participate. Exploiting this date rule, we find that the program substantially increased immigrants' medium-term employment and earnings and decreased their dependency on social benefits.

JEL Classification: J61, J68, H43

Keywords: Immigrants, integration programs, welfare-to-work, regression-discontinuity

5.1. Introduction

The design of immigration policy is among the most important and controversial political issues in many countries. To a large extent, this follows from the increased number of immigrants residing in rich countries and the fact that the labor market performance of these immigrants is typically modest. In particular, immigrants tend to receive more social benefits than comparable natives (Hatton and Williamson, 2006). On the other hand, most OECD countries are facing aging populations and increasing immigration is sometimes argued to be a potential solution to the challenges caused by declining workforce.

The concerns with regard to immigrant integration and the costs immigrants impose on the welfare system have given rise to a variety of policy

¹This essay is joint work with Kari Hämäläinen

responses. Some countries have followed Canada and adopted a points system policy, which aims to prevent the entry of those who are the most likely to collect public benefits. Whether these policies are efficient remains controversial (Borjas, 1993; Antecol et al., 2003). The second approach, most notably taken by the United States in the 1996 welfare reform, is to limit immigrants' rights to claim social benefits (Borjas, 2002; Hanson, 2005).

The third alternative is to design policies that aim to move immigrants from welfare to work. The leading examples are the integration programs launched in many European countries. These programs are typically framed as a mutual contract between the state and the immigrant. The state commits to help immigrants by providing language training, civic courses, labor market orientation, vocational training and the like. The immigrants are expected to participate in these measures. Importantly, refusal to participate is typically sanctioned by substantial reductions of welfare benefits. Sometimes also gaining permanent residence permit is made conditional on participation.² However, little is known about the efficiency of these programs.

This paper studies the impact of "integration plans" implemented in Finland as part of the 1999 Act on the Integration of Immigrants and Reception of Asylum Seekers (henceforth the Integration Act). An integration plan is an individualized sequence of training and subsidized employment combined with reductions in welfare benefits in case of non-compliance. We evaluate the impact of these plans by exploiting a quasi-experimental setting, which was created by backdating the law by two years. That is, the Integration Act came into force in May 1st, 1999, but only those who had entered the population register after May 1st, 1997 had an obligation to participate. Given that the threshold date was set more than a year after the affected immigrants made their entry decisions, it seems plausible to assume that this date rule did not affect the time of entry. Accordingly, the introduction of the Integration Act creates exogenous variation in the likelihood of receiving an integration plan among immigrants arriving at different dates. We exploit this variation within a fuzzy regression discontinuity framework. The results suggest that the integration plans substantially improved immigrants' labor market performance and reduced their welfare dependency.

²Austria, Denmark, Finland, France, Germany, the Netherlands and Sweden have already introduced integration programs and new proposals along similar lines are under consideration in Hungary, Poland and Spain. See Carrera (2006) for a cross-country comparison of integration programs in Europe and Joppke (2007) for a discussion of recent trends in Western European immigration policies.

These findings contribute to two branches of literature. First, to the best of our knowledge, this is the first study to evaluate the impact of an integration program. Thus we provide evidence on the impact of a policy implemented in many countries and being contemplated in many others.³ Second, the Integration Act had many elements in common with welfare-to-work programs such as TANF in the US, the New Deal in the UK, SSP in Canada and welfare policy reforms in Denmark and the Netherlands (see Blundell, 2002; Moffitt, 2002, for reviews of the relevant research). One key element of these programs, also adopted in the 1999 Integration Act, is to make welfare benefits conditional on participation in activation measures.

The rest of this paper is organized as follows. The next two sections provide background information on immigration to Finland and details of the Integration Act. We discuss our empirical approach in Section 5.4 and present the data and main results in Section 5.5. Section 5.6 assesses the robustness of the findings. Section 5.7 concludes.

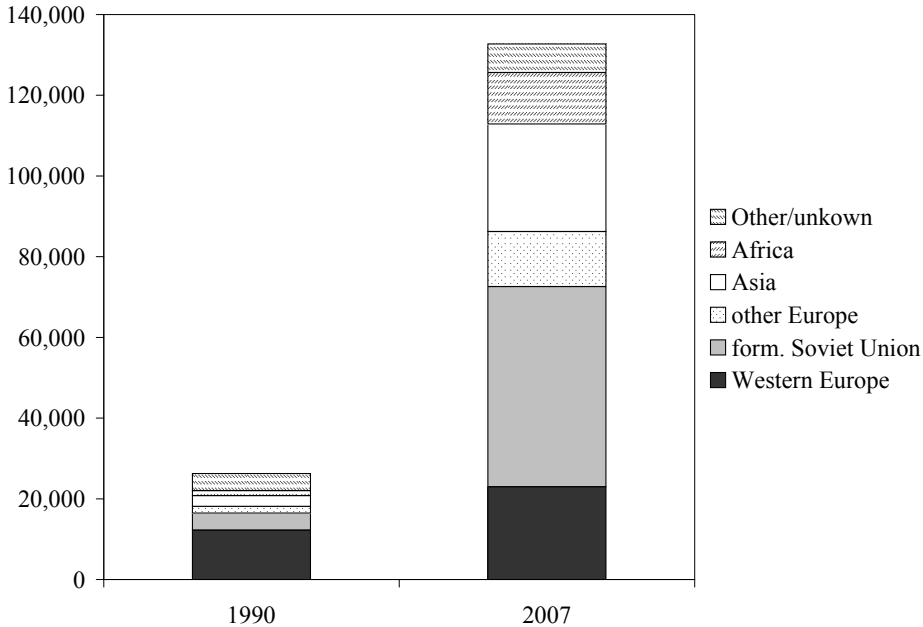
5.2. Immigration to Finland

For the most part of its history, Finland has been characterized by emigration. In particular, large emigration flows to Sweden followed the establishment of free migration between the Nordic countries in 1954. During the period between 1954 and 1970, the magnitude of net emigration was some 220,000 persons – or more than 5% of Finland’s 1954 population. Net-migration turned into positive in the 1980s, but a genuine immigration wave began only in the 1990s. After that the immigrant population has grown fivefold. Given the low initial level, however, immigrant population share is still relatively low at roughly 2.5 per cent in 2007.

As in other Western countries, increasing immigration was accompanied with a change in the composition of origin countries. Figure 5.2.1 presents the stock of foreign nationals living in Finland in 1990 and 2007. Before the immigration wave, almost half of immigrant population had arrived from Western countries. Today, the bulk of immigrants come from the former Soviet Union and Asia. The trend of declining share of Western Europeans coincides with the experience of most other OECD-countries. However, the

³While integration programs have not received much attention, there are few evaluations of the impact of training programs on immigrants (Åslund and Johansson, 2006; Cohen and Eckstein, 2008) and on the impacts of policy changes on labor market assimilation (e.g. Borjas, 1993, 2002; Edin et al., 2004; Constant and Zimmermann, 2005; Rosholm and Vejlin, 2007).

FIGURE 5.2.1. Foreign nationals in Finland



Source: Statistics Finland. Western Europe: Austria, Belgium, Denmark, France, Germany, Greece, Ireland, Island, Italy, Liechtenstein, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, the UK

share of immigrants arriving from the former Soviet Union is unusually high in Finland. An important share of these immigrants is ethnic-Finns (Ingrian-Finns), who were granted the right of return in 1991. Another large immigrant group is the Estonians owing to geographical proximity and the similarities between Estonian and Finnish language.

5.3. The Integration Act

The Integration Act came in force in May 1st, 1999 with the aim to promote integration, equality and freedom of choice by providing measures that help to achieve information and skills needed in the Finnish society. In practice, it introduced two reforms. First, it set new rules for the responsibilities between the central and local administrations. This part of the Integration Act is likely to affect all immigrants and therefore its impacts are hard to measure.

We focus on the second part of the reform, which introduced individualized integration plans for immigrants. The concrete content of an integration

plan depends on the personal factors of the immigrant. An integration plan may include language courses, other preparatory and/or vocational training, career counseling, rehabilitation, work practice, and so forth. Typically various measures are combined to “paths” in which one measure precedes another. The integration plan is aborted if the immigrant finds permanent, full-time employment or becomes a full-time student. Labor administration of the central government is responsible for preparing and implementing the integration plans for 18–64 year old immigrants. Municipalities take care of other age groups.

Eligibility for an integration plan depends on labor market status, the date of entering the population register and the type of residence permit. First, only those registered as unemployed job-seekers or living in a household that receives social assistance are eligible. Second, the right for a plan lasts for three years after the first registration to the population register. This requires a Finnish social security number, which effectively rules out immigrants with very short-term residence permits.

The right for an integration plan is combined with an obligation to participate. However, this obligation relates only to those who entered the population register after May 1st, 1997, while earlier cohorts only had a right to demand an integration plan. When the criteria above are fulfilled, a plan has to be drawn-up during the first five months of an unemployment or social assistance spell. Immigrants have to report how they have followed the plan and whether it needs to be updated. A refusal to participate in the preparation process or to follow the plan is sanctioned by a reduction or withdrawal of integration benefits. Given that labor market support and social assistance are paid as integration benefits during the first three years in Finland, these sanctions could reduce income considerably. In the absence of sanctions, the level of integration benefit corresponds to labor market support that is currently some 500 euros per month. The integration benefit is means-tested and depends on the income of the entire household. However, if the immigrant participates in a labor market policy measure, the integration benefit is paid without means-testing.

The final important feature of the Integration Act is that it did not allocate additional resources on immigrant integration, at least on the central

level. According to the Government Report on Implementation of the Integration Act (2003), the Act was expected to improve integration by reallocating existing resources, training staff, and increasing co-operation between local authorities.

5.4. Empirical Strategy

Our empirical strategy exploits the discontinuity caused by the rule that only immigrants entering the population register after May 1st, 1997 had an obligation to participate in drawing-up of an integration plan (henceforth the treatment). This rule creates a quasi-experimental setting which resembles the situation where the immigrants had been randomized to treatment and control groups.

More precisely, we are able to uncover the causal effect of the treatment under two identifying assumptions. First, those arriving just before and after May, 1997 need to be comparable. Formally, potential outcomes given the date of arrival are assumed to be continuous at the threshold. We argue that this is a fair assumption given that immigrants made their entry decisions two years before the Integration Act was introduced.⁴ Hence immigrants were not able to self-select into the treatment or control group by choosing their date of entry. Furthermore, there were no other policy reforms that would have affected potential outcomes at the threshold. The second identifying assumption is local monotonicity. That is, we need to assume that the probability of being treated never decreased if a person entered the population register after May, 1997 rather than before it. It seems very unlikely that this assumption would be violated.

Given the two assumptions, the causal effect of the treatment is given by the ‘local Wald’ estimator

$$(5.1) \quad \beta = \frac{y^+ - y^-}{p^+ - p^-}$$

where $y^+ = \lim_{z \downarrow z_0} \mathbb{E}[y_i | z_i = z]$ is the limit of the outcome when approaching the threshold z_0 from above and $y^- = \lim_{z \uparrow z_0} \mathbb{E}[y_i | z_i = z]$ is the limit from below (Hahn et al., 2001). In our application the forcing variable z is the month of arrival to Finland and the threshold z_0 is May, 1997. Similarly $p^+ = \lim_{z \downarrow z_0} \mathbb{E}[D_i | z_i = z]$ and $p^- = \lim_{z \uparrow z_0} \mathbb{E}[D_i | z_i = z]$ are the limits for

⁴The threshold date was published in May 8th, 1998 when the government introduced the bill to the parliament. Next day, the leading Finnish newspaper *Helsingin Sanomat* had a short article about the bill, but did not discuss this threshold date.

the probability of being treated when approaching the threshold from above and below.

There are two widely used approaches to estimate equation (5.1): local linear estimator discussed by Hahn et al. (2001) and the control function approach adopted by van der Klaauw (2002). We employ the latter for our baseline estimates due to the relatively small sample size. Local linear estimates are reported in Section 5.6.4 as a robustness check.

Our baseline estimation equation is

$$(5.2) \quad y_i = \alpha + \beta \mathbb{E}[D_i | z_i, X_i] + X_i \theta + k(z_i) + u_i$$

where X_i is a vector of observed background characteristics and the probability of being treated is modeled as

$$(5.3) \quad \mathbb{E}[D_i | z_i, X_i] = \gamma \mathbf{1}\{z_i > z_0\} + X_i \psi + g(z_i)$$

The key idea of the approach is that the underlying dependence between the date of arrival and the outcome is controlled by the smooth term $k(z_i)$. In our context this dependence follows from the assimilation process: immigrants' labor market performance tends to improve as they spend more time in the host country. Failing to take this into account would lead to biased estimates. Similarly, as we discuss in more detail below, the likelihood of being treated was larger among later cohorts than among those entering the population register just after the threshold. This process is controlled for by $g(z_i)$. If both control functions are smooth over the range of arrival dates, a discontinuous jump in $\mathbb{E}[D_i | z_i, X_i]$ allows for consistently estimating the causal effect of the treatment.

In practice, the sample size forces us to specify the control functions in a relatively parsimonious way and this adds a third critical assumption to our identification strategy: the estimates are consistent only under the assumption that the control functions are correctly specified. Since the true form of these functions is unknown, we experiment with several alternative parametrizations. We implement the estimation using a two-step procedure, where we first estimate the propensity to receive the treatment using a linear probability model and then plug this propensity score to the outcome equation. We also report results from an approach where we estimate the numerator and denominator of equation (5.1) in separate regressions and

report the ratio of these two reduced form estimates.⁵ Statistical inference is based on bootstrapping methods.

How should the resulting estimates be interpreted? The first part of the answer is to consider closely, what the treatment is. This is partly clear: the treatment consists of introducing integration plans and threat of sanctions at least two years after the immigrants entered Finland. Thus the timing of the treatment differs from the current practice, where integration plans are typically drawn soon after arrival. Furthermore, the elements of the treatment are sanctions and perhaps better labor market training. Unfortunately, it is hard to evaluate the relative importance of these two elements. We show below that the number of days spent in a training or subsidized work did not change in a statistically detectable way at the May 1997 threshold. However, the aim of the integration plans was to consider the individual circumstances of each immigrant more closely than before. Thus the treatment may have resulted in better matches between immigrants and policy measures.

A further possible concern is that immigrants may have disliked the treatment and thus avoided it by becoming employed before the program was implemented. If this was the case, the group of immigrants affected by the policy would be larger than those who ended up being officially treated. Hence, we would underestimate the denominator of (5.1) and the estimates of β would be biased upwards. While we do not expect this to be of major importance, a conservative interpretation is that we estimate an upper bound of the treatment effect. On the other hand, a very conservative lower bound is obtained by assuming that everyone arriving after May 1997 were affected by the policy change and thus the denominator of (5.1) would be unity. Consequently, the “reduced form” or “intention to treat” (ITT) estimate corresponding to the numerator of (5.1) provides a lower bound for the treatment effect.

Finally, we need to understand for whom the treatment effect is identified. Here we note that β measures a local average treatment effect (Imbens and Angrist, 1994; Hahn et al., 2001). That is, we identify the mean effect among those arriving on May 1st, 1997, who received an integration plan and would not have received it had they arrived earlier. We call this group

⁵More precisely, we estimate regressions $y_i = \alpha_n + \eta_n 1\{z_i > z_0\} + X_i\theta_n + k(z_i) + u_i$ and $D_i = \alpha_d + \eta_d 1\{z_i > z_0\} + X_i\theta_d + g(z_i) + v_i$ and calculate $\hat{\beta}$ as $\widehat{\eta}_n/\widehat{\eta}_d$. This is identical to the control function approach when $k(z_i)$ and $g(z_i)$ are assumed to have the same functional form. However, in our application this assumption does not seem reasonable.

“compliers at the threshold”. An important question is thus how this sub-population of compliers relates to the general immigrant population. We note that in order to end up as a complier, one had to be unemployed or to receive social assistance at least two years after arriving to Finland. Thus the compliers are likely to have below average earnings potential. Furthermore, the denominator of (5.1) is informative about the size of the complier population. Finally, we relate the means of some observed characteristics of the compliers to the entire sample following the approach discussed in the Appendix.

5.5. Data and Results

We use individual-level panel data, created by linking information from several administrative registers including population register, tax register, pension and benefit registers, student register, and the register of unemployed job-seeker maintained by the labor administration. Furthermore, National Research and Development Centre for Welfare and Health (STAKES) has provided us information on social assistance. Different data sources were combined by Statistics Finland using social security numbers.

The data are created by first drawing a 15% random sample of working age immigrants living in Finland in 1989, which is then expanded by including a 15% random sample of new immigrants arriving in each year between 1990 and 2003. For our baseline estimates, we restrict the estimation sample using three sample selection rules. First, we include only immigrants who first arrived to Finland between January 1990 and April 1999. Second, we exclude those who arrived to Finland prior to the age of 16. Finally, we exclude those who were not registered as a job seeker or received unemployment compensation or social assistance during their first three years in Finland. This is motivated by the aim of increasing the precision of the estimates by dropping those (high earning) immigrants, who did not meet the eligibility criteria. We note that this sample selection rule could affect the results if the introduction of integration plan would affect the inflow to unemployment or social assistance benefits. We return to this question in section 5.6 and show that the estimates are similar, though imprecise, also when using the full sample.

Table 5.5.1 presents descriptive statistics for the resulting data. Columns (4) and (5) report the sample means at arrival for immigrants entering Finland within four months of May 1997. The small sample size makes the

TABLE 5.5.1. Background Characteristics

	Means at Arrival (by cohort)								Jump at		Ratio of complier	
	1/90-		1/93-		1/97-		5/97-		May 1997		to sample mean	
	12/92 (1)	12/95 (2)	12/96 (3)	4/97 (4)	8/97 (5)	8/98 (6)	9/98 (7)	9/99 (8)	coef. (9)	se. (10)	coef. (10)	se. (11)
Age	30.5	31.8	32.5	34.0	33.0	32.8	32.8	32.8	-0.98	(1.02)	0.63	(0.02)***
Woman	0.46	0.52	0.57	0.55	0.58	0.60	0.54	0.54	-0.06	(0.05)	0.98	(0.02)
Single	0.36	0.31	0.29	0.22	0.23	0.29	0.26	0.26	0.00	(0.04)	0.41	(0.02)***
Has a native spouse	0.34	0.25	0.28	0.39	0.34	0.32	0.31	0.31	-0.04	(0.04)	1.97	(0.10)***
Has an immigrant spouse	0.30	0.44	0.43	0.40	0.44	0.39	0.42	0.42	0.04	(0.05)	2.69	(0.08)***
Number of children	0.68	0.89	0.72	0.83	0.90	0.75	0.70	0.70	0.24	(0.11)**	2.33	(0.07)***
Local unemployment rate	12.7	20.5	18.2	15.1	15.2	14.3	13.9	13.9	0.83	(0.39)**	0.71	(0.02)***
Lives in the Helsinki area	0.39	0.46	0.40	0.46	0.48	0.40	0.31	0.31	0.06	(0.05)	1.86	(0.08)***
<i>Region of birth</i>												
EU15/EFTA	0.05	0.05	0.07	0.05	0.09	0.07	0.09	0.09	-0.02	(0.02)	0.48	(0.07)***
New EU-members	0.09	0.15	0.10	0.09	0.06	0.06	0.07	0.07	-0.01	(0.02)	0.71	(0.09)***
form. Soviet Union	0.41	0.33	0.38	0.43	0.45	0.45	0.41	0.41	-0.01	(0.05)	1.17	(0.03)***
form. Yugoslavia	0.02	0.13	0.11	0.04	0.01	0.02	0.08	0.08	-0.02	(0.02)	1.26	(0.18)
Turkey	0.04	0.03	0.03	0.05	0.04	0.03	0.02	0.02	0.01	(0.02)	0.70	(0.16)*
Africa	0.17	0.12	0.09	0.06	0.15	0.12	0.08	0.08	0.10	(0.03)***	0.89	(0.08)
Asia	0.17	0.15	0.17	0.24	0.16	0.22	0.21	0.21	-0.04	(0.04)	1.08	(0.06)
Other/Unknown	0.06	0.03	0.04	0.04	0.04	0.03	0.03	0.03	-0.02	(0.02)	0.32	(0.08)***
<i>Legal Status</i>												
Ingrian Finn	0.10	0.25	0.17	0.11	0.19	0.13	0.08	0.08	0.08	(0.03)**	1.05	(0.07)
Family Member	0.13	0.26	0.32	0.40	0.37	0.40	0.41	0.41	-0.04	(0.05)	0.84	(0.03)***
Refugee	0.09	0.19	0.19	0.12	0.09	0.14	0.17	0.17	0.02	(0.03)	1.28	(0.08)***
Other/Unknown	0.68	0.30	0.33	0.37	0.35	0.33	0.34	0.34	-0.06	(0.04)	1.04	(0.04)
Individuals	2,004	1,888	580	167	217	594	377	377	5,827			

Note: Sample means at arrival, OLS estimates for a jump at May 1997, and complier means divided by sample means. Coefficients presented in column (8) are OLS estimates for parameter β in equation $x_i = \alpha + \beta 1\{z_i > z_0\} + \pi_1 z_i + \pi_2 z_i^2 + \pi_3 z_i 1\{z_i > z_0\} + \pi_4 z_i^2 1\{z_i > z_0\} + \epsilon$. Coefficient and bootstrapped standard errors in columns (10) and (11) are obtained using the approach discussed in the Appendix. ***, **, * indicate rejection of the null that the coefficient is zero (column 9) or one (column 11) at 1%, 5%, 10% level, respectively.

sample averages quite noisy. On balance, however, the means do not suggest that dramatic changes had occurred around the threshold date.

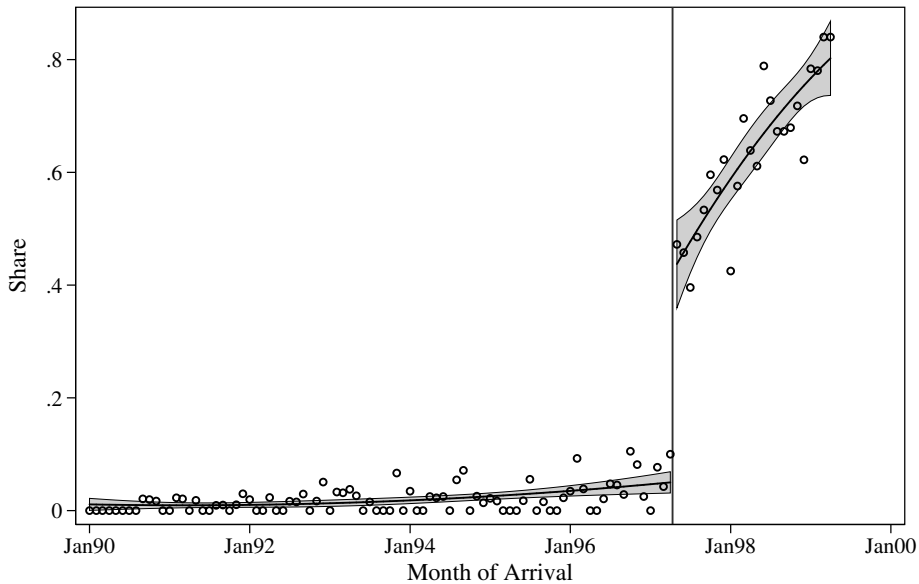
When we extend the observation period, some trends become evident. Most importantly, family unification has become more common and the share of Ingrian Finns has declined after the upsurge in early 1990s. Furthermore, unemployment has declined as the Finnish economy recovered from the severe recession of the early 1990s. It seems safe to assume that none of these changes were caused by anticipation of changes in the Finnish immigration policy. Yet, immigrants arriving at different phases of business cycle could differ in their unobserved characteristics. Note, however, that our identification is based on discontinuity. Thus the change in unobserved characteristics would need to be abrupt in May 1997 to invalidate our empirical strategy. Estimates reported at columns (8) and (9) suggest that 4 out of the 20 background characteristics considered had a statistically significant jump at the May 1997 threshold, when we model the control function as two-sided second order polynomial. While this is a bit discouraging, we note that the assumed functional form of the control function is likely to be less sensible for many of the background characteristics than for the outcome variables. Furthermore, our main results are based on regressions controlling for all observable characteristics presented in the table. We return to the last two columns below.

We next establish that the date of arrival affected the propensity for getting an integration plan. Figure 5.5.1 plots the share of immigrants in the sample, who got an integration plan at some point between May 1999 and December 2003.⁶ The dots correspond to share of treated (y-axis) by month of arrival (x-axis). The lines are the fitted values from a OLS regression corresponding to equation (5.3) without additional covariates. The control function is specified as two-sided second order polynomial. The figure strongly suggest that those arriving in May 1997 were substantially more likely to get an integration plan than those arriving in April 1997. The estimate for the magnitude of the jump is somewhat sensitive to the specification of the control function and ranges between 34 and 41 percentagepoints.⁷ The

⁶The regression results reported in Tables 5.5.2 to 5.6.3 are based on a slightly different definition, where the treatment variable is one if the person has received an integration plan *before* the beginning of the year when the outcome is measured (and zero otherwise). However, this makes little difference, since the alternative treatment definitions are identical among 99% of observations used in the regressions.

⁷Increasing the number of polynomials yield estimates of a jump of 41 (linear), 38 (quadratic), 34 (cubic), 38 (quartic) percentage points. Local linear estimators discussed in

FIGURE 5.5.1. Month of Arrival and Receiving an Integration plan



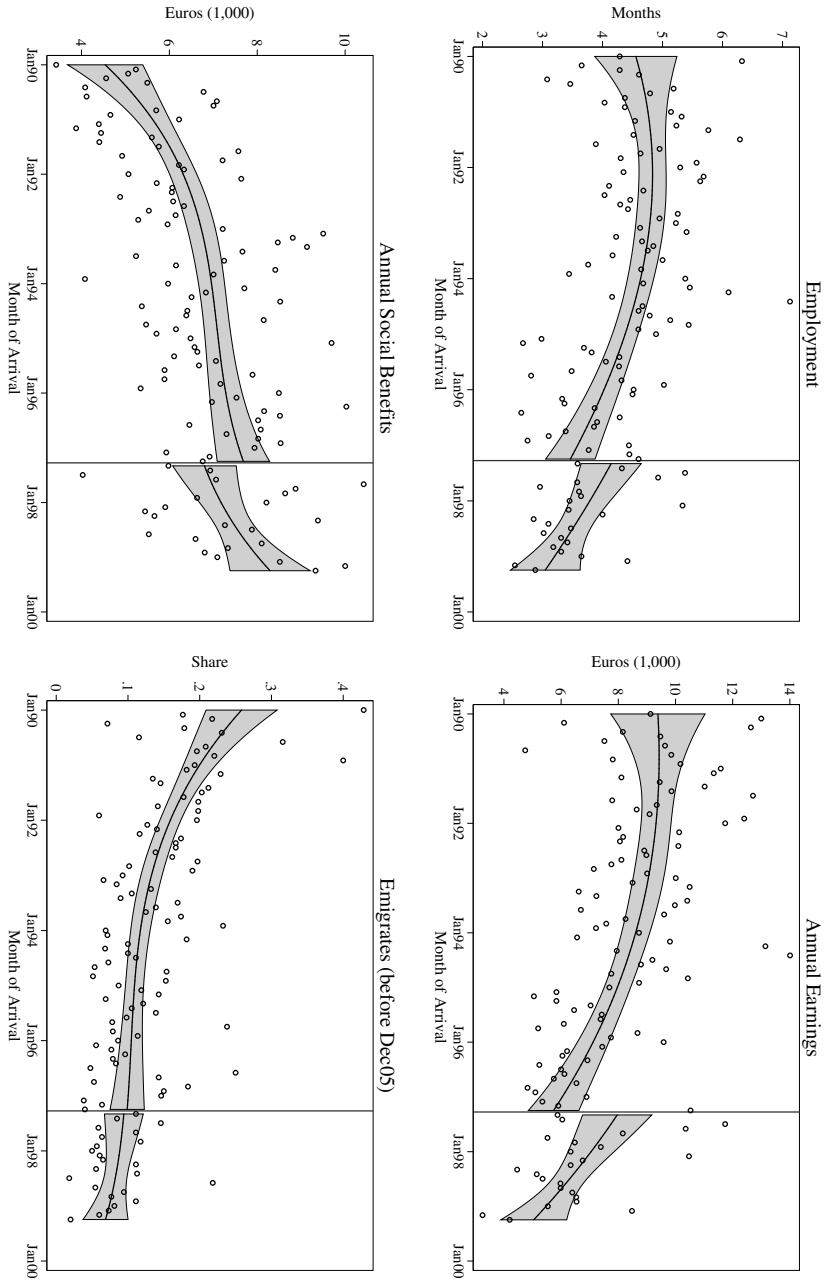
Note: Monthly means, OLS fitted values and 95% confidence interval. The control function is specified as $g(z_i) = \pi_1 z_i + \pi_2 z_i^2 + \pi_3 z_i 1\{z_i > z_0\} + \pi_4 z_i^2 1\{z_i > z_0\}$. Outcome: Receives an integration plan before the end of year 2003.

figure also reveals that the likelihood of receiving the treatment increased after the threshold date. There are two likely reasons for this later trend. First, immigrants who arrived after May 1997 and found a job before the introduction of the integration plans were not eligible. Second, there may have been administrative delays in the actual implementation of providing the plans. Interestingly, very few of those who had arrived before May 1997 exercised their right to demand an integration plan.

We next ask whether we find jumps in later labor market outcomes. Figure 5.5.2 presents average months in open employment, annual earnings and annual social benefits measured in 2003 as well as emigration rates measured at the end of year 2005. As before, the dots are raw averages by month of arrival (x-axis) and the lines are fitted values from OLS regressions, where the control function has been specified as a third order polynomial. Downward sloping lines on the top row indicate that the labor market prospects

detail in Section lead to estimates ranging between 32 percentage points (bandwidth one month) and 39 percentage points (bandwidth 22 months).

FIGURE 5.5.2. Month of Arrival and Labor Market Outcomes in 2003



Note: Monthly means, OLS fitted values and 95% confidence intervals. The control function is specified as $k(z_i) = \pi_1 z_i + \pi_2 z_i^2 + \pi_3 z_i^3$. Outcomes measured in 2003 except emigration, which is defined as not being in the population register at the end of year 2005.

of immigrants improve as they spend more time in Finland. In 2003, those who arrived to Finland in the early 1990s work and earn more and receives less social benefits than those who arrived in the late 1990s. On the other hand, earlier arrivals have had more time to emigrate and thus the lines in the bottom right panel slope downwards. These observations are well in line with previous work on immigrant assimilation to Finland (Sarvimäki, 2008).

More importantly, Figure 5.5.2 suggests that labor market outcomes among those arriving in May 1997 were substantially better in year 2003 than among those arriving in April 1997. The jumps present at the graphs correspond to the numerator of equation (5.1) and can be interpreted as intention to treat (ITT) estimates of introducing the integration plans. Thus they also estimate lower bound of the treatment effect. The point estimates suggest that, overall, the policy change increased employment by 0.7 (standard error 0.4) months in a year and annual earnings by some 2,300 (standard error 930) euros. An improvement in labor market performance also led to a decrease in annual social benefits that accounts for more than 900 (standard error 580) euros at the household level. However, we find no indications that introducing the integration plans would have affected emigration rates.

Table 5.5.2 reports the estimated jumps after controlling for differences in demographic characteristics, region of origin, legal status for residence permit as well as local unemployment rate, type of residence municipality and an indicator for living in the Helsinki area during the first year after arrival. We also report control function and 'Wald' estimates from the two-step estimation procedures discussed in Section 5.4. These correspond to equation (5.1) and can be interpreted either as the impact of an integration plan among the complier sub-population or, more conservatively, as an upper bound of this effect. In order to have a meaningful benchmark, the table also presents the mean outcomes of immigrants, who arrived to Finland in year 1997.

The results strongly suggest that integration plans improved immigrants' labor market outcomes.⁸ They also imply that these effects took about three years to materialize. Moreover, the impacts are large in magnitude. Comparison of the point estimates and the mean outcomes suggest that integration plans increased the pace of the assimilation process by roughly

⁸We have also run all regressions reported here and below using emigration indicator as the dependent variable. However, all specifications produce point estimates close to zero and are thus not reported.

TABLE 5.5.2. Impact of the Integration Plan

	Months Employed (per year)			Annual Earnings			Social Benefits		
	ITT (1)	CF (2)	Wald (3)	ITT (4)	CF (5)	Wald (6)	ITT (7)	CF (8)	Wald (9)
<i>Estimates</i>									
2000	0.26 (0.32)	0.58 (0.76)	0.65 (0.84)	456 (556)	1,042 (1,361)	1,151 (1,448)	-532 (497)	-1,071 (1,218)	-1,344 (1,303)
2001	0.61* (0.36)	1.42 (0.91)	1.54 (0.97)	810 (648)	1,859 (1,626)	2,064 (1,734)	-271 (505)	-530 (1,229)	-690 (1,326)
2002	1.16*** (0.37)	2.56*** (0.91)	2.96*** (1.01)	1,748** (685)	3,623** (1,731)	4,472** (1,874)	-1,251** (521)	-2,824** (1,299)	-3,199** (1,449)
2003	0.84** (0.38)	1.50 (0.94)	2.18** (1.03)	2,494*** (885)	5,031** (2,172)	6,438*** (2,445)	-1,270*** (448)	-2,814** (1,135)	-3,280** (1,276)
Pooled	0.72** (0.30)	1.52** (0.74)	1.82** (0.81)	1,367** (594)	2,872** (1,464)	3,476** (1,632)	-848* (434)	-1,859* (1,086)	-2,157* (1,173)
<i>Means of the 1997 cohort</i>									
2000		2.43			4,070			9,852	
2001		3.25			5,610			8,543	
2002		3.78			6,564			8,383	
2003		4.16			7,661			7,172	
Pooled		3.40			5,975			8,488	

Note: Reduced form estimates (ITT), control function estimates (CF) and ratios of reduced form estimates (Wald), see Section 5.4 for discussion. Standard errors (in parentheses) bootstrapped with 1,000 replications (individual-level block sampling in the pooled sample). Rows correspond to year of measuring the outcome. The control function of the first-stage is modeled as $g(z_i) = \pi_1 z_i + \pi_2 z_i^2 + \pi_3 z_i 1\{z_i > z_0\} + \pi_4 z_i^2 1\{z_i > z_0\}$ and the control function for the second-stage as $k(z_i) = \delta_1 z_i + \delta_2 z_i^2 + \delta_3 z_i^3$. Controlling for age, age squared, gender, region of origin, legal status, local unemployment rate at arrival, type of municipality (city, semi-rural, rural) at arrival, lives in the Helsinki region (Uusimaa) at arrival, marital status, indicators for having children younger than 3-years old, 7-years old and 18-years old in the household, interactions between family characteristics and gender.

one to two years. For instance, the point estimates for months in employment in year 2003 vary between 0.8 months and 2.2 months. Note that interpreting ITT estimates as the lower bound of the treatment effect is very conservative. Thus the true impact for those who got the integration plan is likely to be in the upper part of this range. Also the estimates for annual earnings and social benefits suggest large improvements.

Given such large effects, one would like to know how they came about. Interestingly, we find no indication that getting an integration plan would have lead to more training or subsidized work. Table 5.5.3 reports the number of days immigrants spent in three types of active labor market policies during their first four years in Finland. Nothing suggest that those arriving after May 1997 would have been targeted with more training or subsidized employment.⁹ Unfortunately, we are not able to study whether the integration plans increased the quality of matches between immigrants and policy measures. At minimum, however, these results are consistent with the hypothesis that an important part of the treatment was making social benefits conditional on following the integration plan. It seems reasonable to assume that this conditionality increased incentives to find employment.

Finally, we ask for whom these effects are identified. To do this, we return to Table 5.5.1 and Figure 5.5.1. Note that the jump of the share of treated in the figure reveals that compliers make up more than a third of the individuals in the estimation sample. Furthermore, column (10) of Table 5.5.1 contrasts the characteristics of the compliers in comparison to the entire sample. More precisely it reports estimates for the ratio of expected values of binary variables among compliers divided by the expected values for all immigrants in the estimation sample. For continuous variables, the figures correspond to the likelihoods of the variable having a value above sample median. These estimates imply that compliers differed from average immigrants in many ways. In particular, they were younger, had more children, were less likely to be single and more likely to come from the former Soviet Union, to be refugees and to live outside the Helsinki area.

⁹We acknowledge that some policy measures may be missing from our data. We know from other sources that comparable training, i.e. training offered by other parties than the labor administration, has increased substantially. For instance, in the Helsinki region, the number of participants in comparable training has reached the participant levels in training provided by the labor administration (Ministry of Labour, 2005). However, the share of these training courses is not likely to have increased rapidly for those arriving to Finland in May 1997.

TABLE 5.5.3. Policy Measures (during the first four years in Finland)

	Means at Arrival (by cohort)								Jump at May 1997
	1/90- 12/92 (1)	1/93- 12/95 (2)	1/96- 12/96 (3)	1/97- 4/97 (4)	5/97- 8/97 (5)	9/97- 8/98 (6)	9/98 4/99 (7)	coef. se. (8)	
Integration Plan	0.00	0.00	0.03	0.05	0.44	0.63	0.72	0.38	(0.04)***
Days in training	177	264	294	313	289	304	281	6.1	(25.67)
...preparatory training	83	164	180	173	184	191	173	7.2	(13.17)
...vocational training	39	40	40	47	33	39	40	11.3	(17.23)
...subsidized jobs	55	60	74	93	72	73	68	-12.4	(9.86)
Individuals	2,004	1,888	580	167	217	594	377	5,827	

Note: Sample means during the first four years in Finland by arrival cohort and estimates for a jump at May 1997. Coefficients presented in column (8) are OLS estimates for parameter β in equation $x_i = \alpha + \beta 1_{\{z_i > z_0\}} + \delta_1 z_i + \delta_1 z_i^2 + \delta_1 z_i^3 + \epsilon_{it}$. Standard errors are robust to clustering at individual level.

5.6. Robustness Checks

5.6.1. Pre-Treatment Outcomes. We begin testing the robustness of our results with an “falsification test” based on pre-treatment outcomes. Recall that immigrants arriving after May 1997 became obliged to get an

TABLE 5.6.1. Reduced Form Estimates for Pre- and Post-Treatment Outcomes

	Months employed	Annual Earnings	Social Benefits
	(1)	(2)	(3)
1998	0.21 (0.33)	-144 (614)	430 (687)
1999	-0.33 (0.39)	-806 (705)	-75 (684)
2000	0.06 (0.42)	336 (850)	-782 (684)
2001	0.71 (0.48)	957 (925)	-843 (654)
2002	0.85* (0.49)	1,254 (1,031)	-1,419** (698)
2003	0.55 (0.53)	1,531 (1,253)	-800 (620)

Note: Reduced form estimates and robust standard errors (in parentheses). Sample: excluding those arriving to Finland after December 1997.

integration plan, but only after May 1999. Thus the differences between those arriving before and after May 1997 should materialize only *after* the Integration Act was passed. Thus we can assess the plausibility of the identification strategy by studying whether outcomes measured in 1998 jump at the May 1997 threshold.

Table 5.6.1 reports the reduced form estimates for 1998 outcomes using a similar specification as our baseline estimates. However, since we now can use only immigrants who arrived before January 1998, the estimation sample differs from that used above. Thus we report reduced form estimates for all years from regressions using this more limited sample.

We find no evidence of a jump in pre-treatment outcomes at the May 1997 arrival cohort. Furthermore, the estimates are small for all 1998 outcomes. While the point estimates for later years are close to those obtained from the full baseline sample, the estimates are now more imprecise and thus only few are statistically significantly different from zero.

5.6.2. Jumps at Non-Discontinuity Points. Another way to scrutinize the baseline results is to introduce arbitrary discontinuities in the data and test for their significance. To do this, we create “placebo” thresholds to the data and ask if outcomes measured four to six years later differ between

those arriving before and after these dates. We test for jumps at each month of arrival between January 1993 and December 1998.

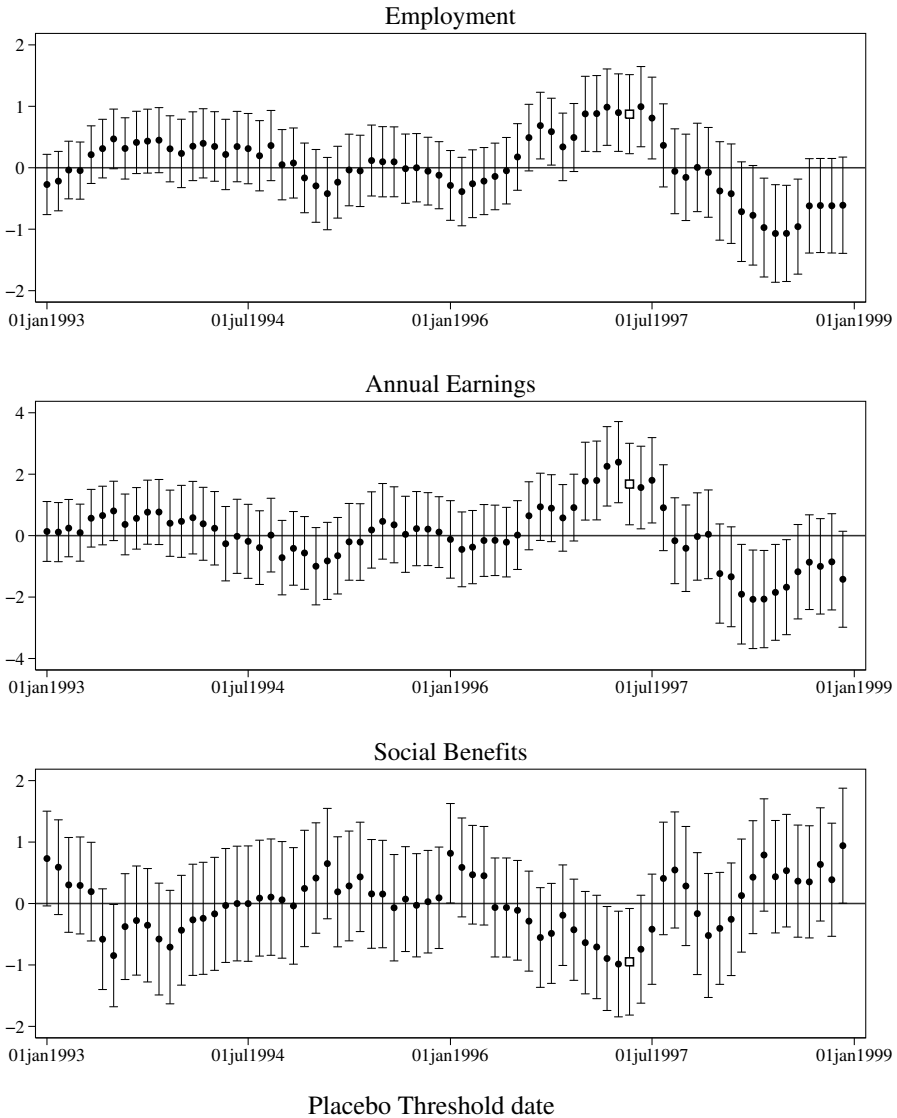
The results are reported in Figure 5.6.1. They strongly suggests that our results do not follow from some systematic error in the specification. We get significant estimates only around the true threshold of May 1997 (marked by an empty square). Note that we should expect to see similar estimates for placebo thresholds close to May 1997 as they can be considered as measuring the true threshold with measurement error. In contrast, estimates for those placebo thresholds, which are not very close to the actual threshold are rarely significant and there are both positive and negative point estimates for each outcome.

5.6.3. Sensitivity to the Sample Selection Rules. Our baseline results are obtained from a sample where we have excluded those immigrants who did not register as unemployed job seekers or received unemployment benefits or social assistance during their first three years in Finland. While this sample selection rule should allow us to focus on the relevant population and thus to improve the precision of the estimates, it also raises two possible concerns. First, if the integration plan has a threat effect, it may alter the inflow to unemployment or social assistance.¹⁰ Second, our data records social assistance paid to the immigrant and to his or her possible spouse. However, we do not observe social assistance paid to the parents in any year. Since one inclusion criteria is based on actually paid social assistance, this sample selection rule excludes all grown-up children who are entitled to an integration plan, but who do not register as job seekers.

Table 5.6.2 examines the sensitivity of the baseline results on the sample selection rules. The first two columns give sample sizes. We find that that excluding immigrants who have no unemployment or social assistance record, we lose more than a quarter from the original sample. However, the estimate presented in column (3) suggest that the likelihood of meeting this sample selection criteria did not jump at the May 1997 threshold. Finally, columns (4) to (7) report the estimates for the full sample and the baseline using the same specification as above. The point estimates point towards qualitatively similar conclusions regardless of the estimation sample used. However, the

¹⁰ *A priori* the direction of this effect is not clear. If immigrants in the weakest job market positions do not register as job seekers or apply for social assistance, our estimates would overestimate the true impact. The contrary happens if mainly the immigrants with the best job prospects would be affected.

FIGURE 5.6.1. Jumps at non-discontinuity points



Note: Reduced form (OLS) estimates and 95% confidence intervals of jumps at non-discontinuity points. X-axis: Placebo threshold date. Specification: Pooling together data 4–6 years after the placebo threshold date and controlling for background characteristics. Confidence intervals are robust to clustering at individual level. The square marker corresponds to the estimate for the real threshold of May, 1997.

TABLE 5.6.2. Sensitivity to the sample selection rule

	Sample size		Jump at	CF estimate (pooled)		
	Ind. (1)	Obs. (2)	May 1997 (3)	Emp. (4)	Earnings (5)	Benefits (6)
All data	8,754	34,934	.	0.88 (0.74)	4,258 (2,870)	-1,289 (1,224)
Baseline sample	5,831	23,313	0.010 (0.016)	1.52* (0.78)	2,872* (1,475)	-1,859* (1,072)

Note: Column (3) reports the OLS estimates for parameter β in equation $e_i = \alpha + \beta 1\{z_i > z_0\} + \pi_1 z_i + \pi_2 z_i^2 + \pi_3 z_i 1\{z_i > z_0\} + \pi_4 z_i^2 1\{z_i > z_0\} + X_i \beta + \epsilon_i$, where e_i is an indicator variable taking value one if the person is in the baseline sample and zero otherwise. Figures in columns (4) to (6) are control function estimates of the impact of an integration plan for pooled 2000–2003 outcomes. The control function of the first-stage is modeled as $g(z_i) = \pi_1 z_i + \pi_2 z_i^2 + \pi_3 z_i 1\{z_i > z_0\} + \pi_4 z_i^2 1\{z_i > z_0\}$ and the control function for the second-stage as $k(z_i) = \delta_1 z_i + \delta_2 z_i^2 + \delta_3 z_i^3$. Standard errors (in parentheses) are based on individual level block bootstrapping with 1,000 replications.

estimates based on the full sample more imprecise and thus not statistically significant.

5.6.4. Sensitivity to Control Function Specification and Local Linear Estimates. One shortcoming of the control function approach is that consistency of the estimates requires correct specification of the control functions. Unfortunately, the true form of these functions are unknown. To assess the sensitivity of our results, Table 5.6.3 reports the main estimates using five alternative parametrizations. The estimates are remarkably stable across these specifications.

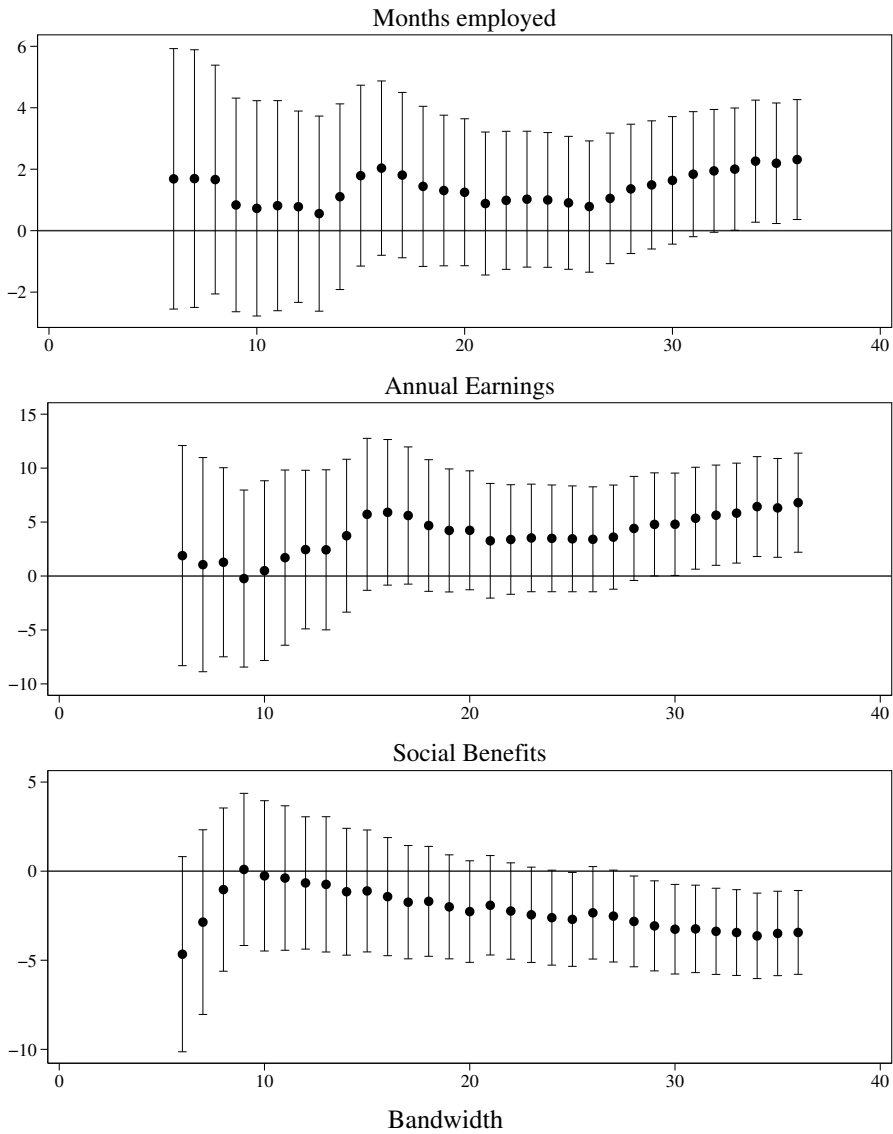
While this stability is reassuring, the best option would be to use a fully nonparametric approach. In principle, any nonparametric estimator could be used. However, the unusual feature of the regression discontinuity estimation problem is that we are interested in single boundary points. Thus standard kernel regression, series regression and sieve methods have poor finite sample properties. Hahn et al. (2001) propose using a local linear regression to overcome these problems. They also show that one version of this estimator can be by running two-stage least squares regressions using a subset of obser-

TABLE 5.6.3. Sensitivity to the Control Function Specification (pooled 2000–2003 outcomes)

Control Function Specification	Months Employed (per year)			Annual Earnings			Social Benefits		
	ITT (1)	CF (2)	Wald (3)	ITT (4)	CF (5)	Wald (6)	ITT (7)	CF (8)	Wald (9)
Quadratic	0.68** (0.30)	1.18* (0.65)	1.83** (0.88)	1,323** (620)	2,297* (1,360)	3,538** (1,799)	-662 (415)	-742 (922)	-1,770 (1,203)
Cubic	0.72** (0.29)	1.59** (0.73)	1.92** (0.85)	1,367** (590)	2,991** (1,493)	3,655** (1,756)	-848** (421)	-1,847* (1,085)	-2,268* (1,258)
Quartic	0.62* (0.34)	1.40* (0.83)	1.66* (0.99)	1,224* (656)	2,686* (1,542)	3,273* (1,930)	-762* (445)	-1,660 (1,104)	-2,037 (1,326)
Lin. w/ int.	0.51* (0.27)	1.18* (0.66)	1.37* (0.79)	969* (557)	2,209 (1,350)	2,590 (1,629)	-876** (376)	-2,007** (928)	-2,341** (1,140)
Quad. w/ int.	0.74** (0.35)	1.92** (0.96)	1.97* (1.08)	1,433* (754)	3,723* (1,989)	3,831* (2,245)	-868 (535)	-2,066 (1,467)	-2,321 (1,640)

Note: Reduced form estimates (ITT), control function estimates (CF) and ratios of reduced form estimates (Wald), see Section 5.4 for discussion. Standard errors (in parentheses) are based on individual level block bootstrapping with 1,000 replications. Rows correspond to different second-stage modeling assumptions of the control function $k(z_i)$. The control function of the first-stage is modeled as $g(z_i) = \pi_1 z_i + \pi_2 z_i^2 + \pi_3 z_i^3 + \pi_4 z_i^4 1\{z_i > z_0\} + \pi_4 z_i^2 1\{z_i > z_0\}$ in all specifications. Controlling for age, age squared, gender, region of origin, legal status, local unemployment rate at arrival, type of municipality (city, semi-rural, rural) at arrival, lives in the Helsinki region (Uusimaa) at arrival, marital status, indicators for having children younger than 3-years old, 7-years old and 18-years old in the household, interactions between family characteristics and gender.

FIGURE 5.6.2. Local Linear Estimates



Note: Local linear (2SLS) estimates and 95% confidence intervals of the impact on an Integration plan. X-axis: bandwidth (months from May 1997). Outcomes measured in 2003.

variations close to the threshold and the indicator $1\{z_i \geq z_0\}$ as an excluded instrument.

Ideally the evaluation would use observations infinitely close to the threshold. In practice, we need to widen the observation window in order to obtain

sufficiently large sample sizes for meaningful analyzes. There is no unambiguous rule for this choice (Ludwig and Miller, 2005; Imbens and Lemieux, 2008). Thus Figure 5.6.2 reports point estimates and 95% confidence intervals for 2SLS estimates when we gradually increase the bandwidth from 6 to 36 months. The estimates are for 2003 outcomes and control for the same variables as the control function estimates. Results for other years and for specifications without covariates are qualitatively similar.

The figure illustrates that in this application one would need to use very wide bandwidths in order to obtain sufficient statistical power to reveal even very large effects. However, using such bandwidths would raise the question on whether the approach can truly be considered nonparametric. Thus we conclude that given the sample size, the most reasonable approach is to use parametric control functions. In any case, the point estimates presented in Figure 5.6.2 lead to similar conclusions as the baseline estimates.

5.7. Conclusions

In this paper, we have provided evidence on that individualized integration plans substantially increased labor market performance and decreased welfare dependency among immigrants to Finland. These findings are based on a quasi-experimental setting and are robust to a variety of robustness checks. Thus we argue that they are very likely to be internally valid.

With respect to their external validity, it is important to understand for whom the setup allows us to measure the effect and what the treatment is. The answer to the former is clear: we estimate the impact for a sub-population of immigrants who arrived to Finland in May 1997 and were unemployed or claimed social assistance two years later. This sub-population corresponds to a large share of Finland's adult immigrant population in late 1990s.

Defining precisely the treatment is harder. We know that it involved receiving an integration plan at least two years after arrival. This plan consisted of an individualized sequence of active labor market policy measures and an obligation to follow that plan with non-compliance sanctioned by substantial withdrawal of welfare benefits. Furthermore, we find that those arriving to Finland after May 1997 spent the same amount of days in training during their first four years in the country than those arriving earlier. Thus the policy may have an impact either through better matches between immigrants and the policy measures or through sanctioning non-compliance

(or a combination of the two). Unfortunately, the data does not allow us to distinguish between these two channels.

These limitations should be acknowledged when extrapolating from our result to other populations and other versions of integration policies. Yet, the results are the first to provide robust evidence on that integration policies can be efficient. Given the high profile these programs have on the policy agenda of many countries, and the controversy surrounding them, it seems fair to argue that these results provide important insights for policy makers and researchers alike.

References

- ANGRIST, J. D. AND J.-S. PISCHKE (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.
- ANTECOL, H., P. KUHN, AND S. TREJO (2003): "Immigration Policy and the Skills of Immigrants to Australia, Canada, and the United States," *Journal of Human Resources*, 38, 192–218.
- BLUNDELL, R. (2002): "Welfare-to-Work: Which Policies Work and Why?" *Proceedings of the British Academy*, 117, 477–524.
- BORJAS, G. (1993): "Immigration Policy, National Origin, and Immigrant Skills: A Comparison of Canada and the United States," in *Small Differences That Matter: Labor Markets and Income Maintenance in Canada and the United States*, ed. by D. Card and R. Freeman, University of Chicago Press.
- (2002): "Welfare Reform and Immigrant Participation in Welfare Programs," *International Migration Review*, 36, 1093–1123.
- CARRERA, S. (2006): "A Comparison of Integration Programs in the EU: Trends and Weaknesses," Challenge Papers No. 1, Centre for European Policy Studies, Brussels.
- COHEN, S. AND Z. ECKSTEIN (2008): "Labor Mobility of Immigrants: Training, Experience, Language and Opportunities," *forthcoming in International Economic Review*.
- CONSTANT, A. AND K. ZIMMERMANN (2005): "Immigrant Performance and Selective Immigration Policy: A European Perspective," *National Institute Economic Review*, 194, 94.
- EDIN, P., P. FREDRIKSSON, AND O. ÅSLUND (2004): "Settlement policies and the economic success of immigrants," *Journal of Population Economics*, 17, 133–155.
- GOVERNMENT REPORT ON IMPLEMENTATION OF THE INTEGRATION ACT (2003), Publication of Labor Administration 317. Helsinki: Ministry of Labor.
- HAHN, J., P. TODD, AND W. KLAAUW (2001): "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69, 201–209.
- HANSON, G. (2005): *Why Does Immigration Divide America?: Public Finance and Political Opposition to Open Borders*, Peterson Institute.

- HATTON, T. AND J. WILLIAMSON (2006): *Migration and the World Economy: Two Centuries for Policy and Performance*, Cambridge, Massachusetts: MIT Press.
- IMBENS, G. AND J. ANGRIST (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 467–467.
- IMBENS, G. AND T. LEMIEUX (2008): "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 142, 615–635.
- JOPPKE, C. (2007): "Beyond national models: Civic integration policies for immigrants in Western Europe," *West European Politics*, 30, 1–22.
- LUDWIG, J. AND D. MILLER (2005): "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design," Working Paper 11702. NBER.
- MINISTRY OF LABOUR (2005): "Yhteistyössä kotouttamiskoulutusta toteuttamaan - Työ- ja opetushallinnon työryhmän raportti," Publication of Labor Administration 317. Helsinki: Ministry of Labor.
- MOFFITT, R. (2002): "Welfare Programs and Labor Supply," in *Handbook of Public Economics*, ed. by A. Auerbach and M. Feldstein, vol. 4, 2327–2430.
- ROSHOLM, M. AND R. VEJLIN (2007): "Reducing Income Transfers to Refugee Immigrants: Does Starthelp Help You Start?" IZA Discussion Paper No. 2720.
- SARVIMÄKI, M. (2008): "Assimilation to a Welfare State: Labor Market Performance and Use of Social Benefits by Immigrants to Finland," VATT Discussion Paper 454.
- ÅSLUND, O. AND P. JOHANSSON (2006): "Virtues of SIN. Effects of an immigrant workplace introduction program," IFAU Working Paper 2006:7.
- VAN DER KLAAUW, W. (2002): "Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach," *International Economic Review*, 43, 1249–1287.

Appendix

Characterizing the complier population. We follow the approach discussed in Angrist and Pischke (2009). That is, for binary observed characteristic x_i , we can estimate the ratio

$$\begin{aligned} \frac{\mathbb{E}[D_i|z_i \geq z_0, x_i = 1] - \mathbb{E}[D_i|z_i < z_0, x_i = 1]}{\mathbb{E}[D_i|z_i \geq z_0] - \mathbb{E}[D_i|z_i < z_0]} &= \\ \frac{P(D_{i,z_i \geq z_0} > D_{i,z_i < z_0} | x_i = 1)}{P(D_{i,z_i \geq z_0} > D_{i,z_i < z_0})} &= \\ \frac{P(x_i = 1 | D_{i,z_i \geq z_0} > D_{i,z_i < z_0})}{P(x_i = 1)} \end{aligned}$$

where $D_{i,z_i \geq z_0}$ is the potential treatment status of individual i if she enters the country after the threshold date and $D_{i,z_i < z_0}$ is her potential treatment status if she enters the country before the threshold date (the compliers are those for whom $D_{i,z_i \geq z_0} = 1$ and $D_{i,z_i < z_0} = 0$, i.e. those who would get the treatment if they enter the country after the threshold date, but not if they would arrive earlier). Note that the numerator of the expression above is the expected value of a binary variable x_i among the compliers and the denominator is the expected value of x_i among the entire population.

A-SARJA: VÄITÖSKIRJOJA - DOCTORAL DISSERTATIONS. ISSN 1237-556X.

- A:287. TIMO JÄRVENSIVU: Values-driven management in strategic networks: A case study of the influence of organizational values on cooperation. 2007.
ISBN-10: 952-488-081-4, ISBN-13: 978-952-488-081-7.
- A:288. PETRI HILLI: Riskinhallinta yksityisen sektorin työeläkkeiden rahoituksessa. 2007.
ISBN-10: 952-488-085-7, ISBN-13: 978-952-488-085-5.
E-version: ISBN 978-952-488-110-4.
- A:289. ULLA KRUHSE-LEHTONEN: Empirical Studies on the Returns to Education in Finland. 2007. ISBN 978-952-488-089-3, E-version ISBN 978-952-488-091-6.
- A:290. IRJA HYVÄRI: Project Management Effectiveness in Different Organizational Conditions. 2007. ISBN 978-952-488-092-3, E-version: 978-952-488-093-0.
- A:291. MIKKO MÄKINEN: Essays on Stock Option Schemes and CEO Compensation. 2007.
ISBN 978-952-488-095-4.
- A:292. JAAKKO ASPARA: Emergence and Translations of Management Interests in Corporate Branding in the Finnish Pulp and Paper Corporations. A Study with an Actor-Network Theory Approach. 2007. ISBN 978-952-488-096-1, E-version: 978-952-488-107-4.
- A:293. SAMI J. SARPOLA: Information Systems in Buyer-supplier Collaboration. 2007.
ISBN 978-952-488-098-5.
- A:294. SANNA K. LAUKKANEN: On the Integrative Role of Information Systems in Organizations: Observations and a Proposal for Assessment in the Broader Context of Integrative Devices. 2006. ISBN 978-952-488-099-2.
- A:295. CHUNYANG HUANG: Essays on Corporate Governance Issues in China. 2007.
ISBN 978-952-488-106-7, E-version: 978-952-488-125-8.
- A:296. ALEKSI HORSTI: Essays on Electronic Business Models and Their Evaluation. 2007.
ISBN 978-952-488-117-3, E-version: 978-952-488-118-0.
- A:297. SARI STENFORS: Strategy tools and strategy toys: Management tools in strategy work. 2007. ISBN 978-952-488-120-3, E-version: 978-952-488-130-2.
- A:298. PÄIVI KARHUNEN: Field-Level Change in Institutional Transformation: Strategic Responses to Post-Socialism in St. Petersburg Hotel Enterprises. 2007.
ISBN 978-952-488-122-7, E-version: 978-952-488-123-4.
- A:299. EEVA-KATRI AHOLA: Producing Experience in Marketplace Encounters: A Study of Consumption Experiences in Art Exhibitions and Trade Fairs. 2007.
ISBN 978-952-488-126-5.
- A:300. HANNU HÄNNINEN: Negotiated Risks: The Estonia Accident and the Stream of Bow Visor Failures in the Baltic Ferry Traffic. 2007. ISBN 978-952-499-127-2.

- A:301. MARIANNE KIVELÄ: Dynamic Capabilities in Small Software Firms. 2007. ISBN 978-952-488-128-9.
- A:302. OSMO T.A. SORONEN: A Transaction Cost Based Comparison of Consumers' Choice between Conventional and Electronic Markets. 2007. ISBN 978-952-488-131-9.
- A:303. MATTI NOJONEN: Guanxi – The Chinese Third Arm. 2007. ISBN 978-952-488-132-6.
- A:304. HANNU OJALA: Essays on the Value Relevance of Goodwill Accounting. 2007. ISBN 978-952-488-133-3, E-version: 978-952-488-135-7.
- A:305. ANTTI KAUKHANEN: Essays on Empirical Personnel Economics. 2007. ISBN 978-952-488-139-5.
- A:306. HANS MÄNTYLÄ: On "Good" Academic Work – Practicing Respect at Close Range. 2007. ISBN 978,952-488-1421-8, E-version: 978-952-488-142-5.
- A:307. MILLA HUURROS: The Emergence and Scope of Complex System/Service Innovation. The Case of the Mobile Payment Services Market in Finland. 2007. ISBN 978-952-488-143-2
- A:308. PEKKA MALO: Higher Order Moments in Distribution Modelling with Applications to Risk Management. 2007. ISBN 978-952-488-155-5, E-version: 978-952-488-156-2.
- A:309. TANJA TANAYAMA: Allocation and Effects of R&D Subsidies: Selection, Screening, and Strategic Behavior. 2007. ISBN 978-952-488-157-9, E-version: 978-952-488-158-6.
- A:310. JARI PAULAMÄKI: Kauppiasyrittäjän toimintavapaus ketjuyrityksessä. Haastattelututkimus K-kaupiaan kokemasta toimintavapaudesta agenttiteorian näkökulmasta. 2008. Korjattu painos. ISBN 978-952-488-246-0, E-version: 978-952-488-247-7.
- A:311. JANNE VIHINEN: Supply and Demand Perspectives on Mobile Products and Content Services. ISBN 978-952-488-168-5.
- A:312. SAMULI KNÜPFER: Essays on Household Finance. 2007. ISBN 978-952-488-178-4.
- A:313. MARI NYRHINEN: The Success of Firm-wide IT Infrastructure Outsourcing: an Integrated Approach. 2007. ISBN 978-952-488-179-1.
- A:314. ESKO PENTTINEN: Transition from Products to Services within the Manufacturing Business. 2007. ISBN 978-952-488-181-4, E-version: 978-952-488-182-1.
- A:315. JARKKO VESA: A Comparison of the Finnish and the Japanese Mobile Services Markets: Observations and Possible Implications. 2007. ISBN 978-952-488-184-5.
- A:316. ANTTI RUOTOISTENMÄKI: Condition Data in Road Maintenance Management. 2007. ISBN 978-952-488-185-2, E-version: 978-952-488-186-9.
- A:317. NINA GRANQVIST: Nanotechnology and Nanolabeling. Essays on the Emergence of New Technological Fields. 2007. ISBN 978-952-488-187-6, E-version: 978-952-488-188-3.
- A:318. GERARD L. DANFORD: INTERNATIONALIZATION: An Information-Processing Perspective. A Study of the Level of ICT Use During Internationalization. 2007. ISBN 978-952-488-190-6.

- A:319. TIINA RITVALA: Actors and Institutions in the Emergence of a New Field: A Study of the Cholesterol-Lowering Functional Foods Market. 2007. ISBN 978-952-488-195-1.
- A:320. JUHA LAAKSONEN: Managing Radical Business Innovations. A Study of Internal Corporate Venturing at Sonera Corporation. 2007. ISBN 978-952-488-201-9, E-version: 978-952-488-202-6.
- A:321. BRETT FIFIELD: A Project Network: An Approach to Creating Emergent Business. 2008. ISBN 978-952-488-206-4, E-version: 978-952-488-207-1.
- A:322. ANTTI NURMI: Essays on Management of Complex Information Systems Development Projects. 2008. ISBN 978-952-488-226-2.
- A:323. SAMI RELANDER: Towards Approximate Reasoning on New Software Product Company Success Potential Estimation. A Design Science Based Fuzzy Logic Expert System. 2008. ISBN 978-952-488-227-9.
- A:324. SEPPO KINKKI: Essays on Minority Protection and Dividend Policy. 2008. ISBN 978-952-488-229-3.
- A:325. TEEMU MOILANEN: Network Brand Management: Study of Competencies of Place Branding Ski Destinations. 2008. ISBN 978-952-488-236-1.
- A:326. JYRKI ALI-YRKKÖ: Essays on the Impacts of Technology Development and R&D Subsidies. 2008. ISBN 978-952-488-237-8.
- A:327. MARKUS M. MÄKELÄ: Essays on software product development. A Strategic management viewpoint. 2008. ISBN 978-952-488-238-5.
- A:328. SAMI NAPARI: Essays on the gender wage gap in Finland. 2008. ISBN 978-952-488-243-9.
- A:329. PAULA KIVIMAA: The innovation effects of environmental policies. Linking policies, companies and innovations in the Nordic pulp and paper industry. 2008. ISBN 978-952-488-244-6.
- A:330. HELI VIRTA: Essays on Institutions and the Other Deep Determinants of Economic Development. 2008. ISBN 978-952-488-267-5.
- A:331. JUKKA RUOTINEN: Essays in trade in services difficulties and possibilities. 2008. ISBN 978-952-488-271-2, E-version: ISBN 978-952-488-272-9.
- A:332. IIKKA KORHONEN: Essays on commitment and government debt structure. 2008. ISBN 978-952-488-273-6, E-version: ISBN 978-952-488-274-3.
- A:333. MARKO MERISAVO: The interaction between digital marketing communication and customer loyalty. 2008. ISBN 978-952-488-277-4, E-version 978-952-488-278-1.
- A:334. PETRI ESKELINEN: Reference point based decision support tools for interactive multiobjective optimization. 2008. ISBN 978-952-488-282-8.
- A:335. SARI YLI-KAUHALUOMA: Working on technology: a study on collaborative R&D work in industrial chemistry. 2008. ISBN 978-952-488-284-2

- A:336. JANI KILPI: Sourcing of availability services - case aircraft component support. 2008. ISBN 978-952-488-284-2, 978-952-488-286-6 (e-version).
- A:337 HEIDI SILVENNOINEN: Essays on household time allocation decisions in a collective household model. 2008. ISBN 978-952-488-290-3, ISBN 978-952-488-291-0 (e-version).
- A:338 JUKKA PARTANEN: Pk-yrityksen verkostokyvyykkydet ja nopea kasvu - case: Tiede- ja teknologiavetoiset yritykset. 2008. ISBN 978-952-488-295-8.
- A:339 PETRUS KAUTTO: Who holds the reins in Integrated Product Policy? An individual company as a target of regulation and as a policy maker. 2008. ISBN 978-952-488-300-9, 978-952-488-301-6 (e-version).
- A:340 KATJA AHONIEMI: Modeling and Forecasting Implied Volatility. 2009. ISBN 978-952-488-303-0, E-version: 978-952-488-304-7.
- A:341 MATTI SARVIMÄKI: Essays on Migration. 2009. ISBN 978-952-488-305-4, 978-952-488-306-1 (e-version).
- B-SARJA: TUTKIMUKSIA - RESEARCH REPORTS. ISSN 0356-889X.
- B:77. MATTI KAUTTO – ARTO LINDBLOM – LASSE MITRONEN: Kaupan liiketoiminta-osaaminen. 2007. ISBN 978-952-488-109-8.
- B:78. NIILLO HOME: Kauppiasyrittäjyys. Empiirinen tutkimus K-ruokakauppioiden yrittäjyysasenteista. Entrepreneurial Orientation of Grocery Retailers – A Summary. ISBN 978-952-488-113-5, E-versio: ISBN 978-952-488-114-2.
- B:79. PÄIVI KARHUNEN – OLENA LESYK – KRISTO OVASKA: Ukraina suomalaisyritysten toimintaympäristönä. 2007. ISBN 978-952-488-150-0, E-versio: 978-952-488-151-7.
- B:80. MARIA NOKKONEN: Näkemyksiä pörssiyrityiden hallitusten sukupuolikiintiöistä. Retorinen diskurssianalyysi Helsingin Sanomien verkkokeskusteluista. Nasta-projekti. 2007. ISBN 978-952-488-166-1, E-versio: 978-952-488-167-8.
- B:81. PIIA HELISTE – RIITTA KOSONEN – MARJA MATTILA: Suomalaisyritykset Baltiassa tänään ja huomenna: Liiketoimintanormien ja -käytäntöjen kehityksestä. 2007. ISBN 978-952-488-177-7, E-versio: 978-952-488-183-8.
- B:82. OLGA MASHKINA – PIIA HELISTE – RIITTA KOSONEN: The Emerging Mortgage Market in Russia: An Overview with Local and Foreign Perspectives. 2007. ISBN 978-952-488-193-7, E-version: 978-952-488-194-4.
- B:83. PIIA HELISTE – MARJA MATTILA – KRZYSZTOF STACHOWIAK: Puola suomalaisyritysten toimintaympäristönä. 2007. ISBN 978-952-488-198-2, E-versio: 978-952-488-199-9.
- B:84. PÄIVI KARHUNEN – RIITTA KOSONEN – JOHANNA LOGRÉN – KRISTO OVASKA: Suomalaisyritysten strategiat Venäjän muuttuvassa liiketoimintaympäristössä. 2008. ISBN 978-953-488-212-5, E-versio: 978-952-488-241-5.

- B:85. MARJA MATTILA – EEVA KEROLA – RIITTA KOSONEN: Unkari suomalaisyritysten toimintaympäristönä. 2008. ISBN 978-952-488-213-2, E-versio: 978-952-488-222-4.
- B:86. KRISTIINA KORHONEN – ANU PENTTILÄ – MAYUMI SHIMIZU – EEVA KEROLA – RIITTA KOSONEN: Intia suomalaisyritysten toimintaympäristönä. 2008. ISBN 978-952-488-214-9, E-versio: 978-952-488-283-5
- B:87. SINIKKA VANHALA – SINIKKA PESONEN: Työstä nauttien. SEFE:en kuuluvien nais- ja miesjohtajien näkemyksiä työstään ja urastaan. 2008. ISBN 978-952-488-224-8, E-versio: 978-952-488-225-5.
- B:88. POLINA HEININEN – OLGA MASHKINA – PÄIVI KARHUNEN – RIITTA KOSONEN: Leningradin lääni yritysten toimintaympäristönä: pk-sektorin näkökulma. 2008. ISBN 978-952-488-231-6, E-versio: 978-952-488-235-4.
- B:89. Ольга Машкина – Полина Хейнинен: Влияние государственного сектора на развитие малого и среднего предпринимательства в Ленинградской области: взгляд предприятий. 2008. ISBN 978-952-488-233-0, E-version: 978-952-488-240-8.
- B:90. MAI ANTTILA – ARTO RAJALA (Editors): Fishing with business nets – keeping thoughts on the horizon Professor Kristian Möller. 2008. ISBN 978-952-488-249-1, E-version: 978-952-488-250-7.
- B:91. RENÉ DE KOSTER – WERNER DELFMANN (Editors): Recent developments in supply chain management. 2008. ISBN 978-952-488-251-4, E-version: 978-952-488-252-1.
- B:92. KATARIINA RASILAINEN: Valta orkesterissa. Narratiivinen tutkimus soittajien kokemuksista ja näkemyksistä. 2008. ISBN 978-952-488-254-5, E-versio: 978-952-488-256-9.
- B:93. SUSANNA KANTELINEN: Opiskelen, siis koen. Kohti kokevan subjektin tunnistavaa korkeakoulututkimusta. 2008. ISBN 978-952-488-257-6, E-versio: 978-952-488-258.
- B:94. KATRI KARJALAINEN – TUOMO KIVIOJA – SANNA PELLAVA: Yhteishankintojen kustannusvaikutus. Valtion hankintatoimen kustannussäästöjen selvittäminen. 2008. ISBN 978-952-488-263-7, E-versio: ISBN 978-952-488-264-4.
- B:95. ESKO PENTTINEN: Electronic Invoicing Initiatives in Finland and in the European Union – Taking the Steps towards the Real-Time Economy. 2008. ISBN 978-952-488-268-2, E-versio: ISBN 978-952-488-270-5.
- B:96. LIISA UUSITALO (Editor): Museum and visual art markets. 2008. ISBN 978-952-488-287-3, E-version: ISBN 978-952-488-288-0.
- B:97. EEVA-LIISA LEHTONEN: Pohjoismaiden ensimmäinen kauppatieteiden tohtori Vilho Paavo Nurmilahti 1899-1943. 2008. ISBN 978-952-488-292-7, E-versio: ISBN 978-952-488-293-4.
- B:98. ERJA KETTUNEN – JYRI LINTUNEN – WEI LU – RIITTA KOSONEN: Suomalaisyritysten strategiat Kiinan muuttuvassa toimintaympäristössä. 2008 ISBN 978-952-488-234-7, E-versio: ISBN 978-952-488-297-2.
- B:99. SUSANNA VIRKKULA – EEVA-KATRI AHOLA – JOHANNA MOISANDER – JAAKKO ASPARA – HENRIKKI TIKKANEN: Messut kuluttajia osallistavan markkinakulttuurin fasilitaattorina: messukokemuksen rakentuminen Venemessuilla. 2008. ISBN 978-952-488-298-9, E-versio: ISBN 978-952-488-299-6

- N:63. SOILE MUSTONEN – ANNE GUSTAFSSON-PESONEN: Oppilaitosten yrittäjyyskoulutuksen kehittämishanke 2004–2006 Etelä-Savon alueella. Tavoitteiden, toimenpiteiden ja vaikuttavuuden arviointi. 2007. ISBN: 978-952-488-086-2.
- N:64. JOHANNA LOGRÉN – VESA KOKKONEN: Pietarissa toteutettujen yrittäjäkoulutusohjelmien vaikuttavuus. 2007. ISBN 978-952-488-111-1.
- N:65. VESA KOKKONEN: Kehity esimiehenä – koulutusohjelman vaikuttavuus. 2007. ISBN 978-952-488-116-6.
- N:66. VESA KOKKONEN – JOHANNA LOGRÉN: Kaupallisten avustajien – koulutusohjelman vaikuttavuus. 2007. ISBN 978-952-488-116-6.
- N:67. MARKKU VIRTANEN: Summary and Declaration. Of the Conference on Public Support Systems of SME's in Russia and Other North European Countries. May 18 – 19, 2006, Mikkeli, Finland. 2007. ISBN 978-952-488-140-1.
- N:68. ALEKSANDER PANFILO – PÄIVI KARHUNEN: Pietarin ja Leningradin läänin potentiaali kaakkoissuomalaisille metallialan yrityksille. 2007. ISBN 978-952-488-163-0.
- N:69. ALEKSANDER PANFILO – PÄIVI KARHUNEN – VISA MIETTINEN: Pietarin innovaatiojärjestelmä jayhteistyöpotentiaali suomalaisille innovaatiotoimijoille. 2007. ISBN 978-952-488-164-7.
- N:70. VESA KOKKONEN: Perusta Oma Yritys – koulutusohjelman vaikuttavuus. 2007. ISBN 978-952-488-165-4.
- N:71. JARI HANDELBERG – MIKKO SAARIKIVI: Tutkimus Miktech Yrityshautomon yritysten näkemyksistä ja kokemuksista hautomon toiminnasta ja sen edelleen kehittamisestä. 2007. ISBN 978-952-488-175-3.
- N:72. SINIKKA MYNTTINEN – MIKKO SAARIKIVI – ERKKI HÄMÄLÄINEN: Mikkelin Seudun yrityspalvelujen henkilökunnan sekä alueen yrittäjien näkemykset ja suhtautuminen mentorointiin. 2007. ISBN 978-952-488-176-0.
- N:73. SINIKKA MYNTTINEN: Katsaus K-päivittäistavarakauppaan ja sen merkitykseen Itä-Suomessa. 2007. ISBN 978-952-488-196-8.
- N:74. MIKKO SAARIKIVI: Pk-yritysten kansainvälistymisen sopimukset. 2008. ISBN 978-952-488-210-1.
- N:75. LAURA TUUTTI: Uutta naisjohtajuutta Delfoi Akatemiasta – hankkeen vaikuttavuus. 2008. ISBN 978-952-488-211-8.
- N:76. LAURA KEHUSMAA – JUSSI KÄMÄ – ANNE GUSTAFSSON-PESONEN (ohjaaja): StuNet -Business Possibilities and Education - hankkeen arviointi. 2008. ISBN 978-952-488-215-6.
- N:77. PÄIVI KARHUNEN – ERJA KETTUNEN – VISA MIETTINEN – TIINAMARI SIVONEN: Determinants of knowledge-intensive entrepreneurship in Southeast Finland and Northwest Russia. 2008. ISBN 978-952-488-223-1.

- N:78. ALEKSANDER PANFILO – PÄIVI KARHUNEN – VISA MIETTINEN: Suomalais-venäläisen innovaatioyhteistyön haasteet toimijanäkökulmasta. 2008. ISBN 978-952-488-232-3.
- N:79. VESA KOKKONEN: Kasva Yrittäjäksi – koulutusohjelman vaikuttavuus. 2008. ISBN 978-952-488-248-4.
- N:80. VESA KOKKONEN: Johtamisen taidot - hankkeessa järjestettyjen koulutusohjelmien vaikuttavuus. 2008. ISBN 978-952-488-259-0.
- N:81. MIKKO SAARIKIVI: Raportti suomalaisten ja brittiläisten pk-yritysten yhteistyön kehittämisestä uusiutuvan energian sektorilla. 2008. ISBN 978-952-488-260-6.
- N:82. MIKKO SAARIKIVI – JARI HANDELBERG – TIMO HOLMBERG – ARI MATILAINEN: Selvitys lujitemuovikomposiittituotteiden mahdollisuuksista rakennusteollisuudessa. 2008. ISBN 978-952-488-262-0.
- N:83. PÄIVI KARHUNEN – SVETLANA LEDYAEVA – ANNE GUSTAFSSON-PESONEN – ELENA MOCHNIKOVA – DMITRY VASILENKO: Russian students' perceptions of entrepreneurship. Results of a survey in three St. Petersburg universities. Entrepreneurship development –project 2. 2008. ISBN 978-952-488-280-4.

W-SARJA: TYÖPAPEREITA - WORKING PAPERS . ISSN 1235-5674.
ELECTRONIC WORKING PAPERS, ISSN 1795-1828.

- W:412. LOTHAR THIELE – KAISA MIETTINEN – PEKKA J. KORHONEN – JULIAN MOLINA: A Preference-Based Interactive Evolutionary Algorithm for Multiobjective Optimization. 2007. ISBN 978-952-488-094-7.
- W:413. JAN-ERIK ANTIPIN – JANI LUOTO: Are There Asymmetric Price Responses in the Euro Area? 2007. ISBN 978-952-488-097-8.
- W:414. SAMI SARPOLA: Evaluation Framework for VML Systems. 2007. ISBN 978-952-488-097-8.
- W:415. SAMI SARPOLA: Focus of Information Systems in Collaborative Supply Chain Relationships. 2007. ISBN 978-952-488-101-2.
- W:416. SANNA LAUKKANEN: Information Systems as Integrative Infrastructures. Information Integration and the Broader Context of Integrative and Coordinative Devices. 2007. ISBN 978-952-488-102-9.
- W:417. SAMULI SKURNIK – DANIEL PASTERNAK: Uusi näkökulma 1900-luvun alun murroskauteen ja talouden murrosvaiheiden dynamiikkaan. Liikemies Moses Skurnik osakesijoittajana ja -välittäjänä. 2007. ISBN 978-952-488-104-3.
- W:418. JOHANNA LOGRÉN – PIIA HELISTE: Kymenlaakson pienten ja keskisuurten yritysten Venäjä-yhteistyöpotentiaali. 2001. ISBN 978-952-488-112-8.
- W:419. SARI STENFORS – LEENA TANNER: Evaluating Strategy Tools through Activity Lens. 2007. ISBN 978-952-488-120-3.
- W:420. RAIMO LOVIO: Suomalaisen monikansallisten yritysten kotimaisen sidoksen heikkeneminen 2000-luvulla. 2007. ISBN 978-952-488-121-0.

- W:421. PEKKA J. KORHONEN – PYRY-ANTTI SIITARI: A Dimensional Decomposition Approach to Identifying Efficient Units in Large-Scale DEA Models. 2007. ISBN 978-952-488-124-1.
- W:422. IRYNA YEVSEYEVA – KAISA MIETTINEN – PEKKA SALMINEN – RISTO LAHDELMA: SMAA-Classification - A New Method for Nominal Classification. 2007. ISBN 978-952-488-129-6.
- W:423. ELINA HILTUNEN: The Futures Window – A Medium for Presenting Visual Weak Signals to Trigger Employees' Futures Thinking in Organizations. 2007. ISBN 978-952-488-134-0.
- W:424. TOMI SEPPÄLÄ – ANTTI RUOTOISTENMÄKI – FRIDTJOF THOMAS: Optimal Selection and Routing of Road Surface Measurements. 2007. ISBN 978-952-488-137-1.
- W:425. ANTTI RUOTOISTENMÄKI: Road Maintenance Management System. A Simplified Approach. 2007. ISBN 978-952-488-1389-8.
- W:426. ANTTI PIRJETÄ – VESA PUTTONEN: Style Migration in the European Markets 2007. ISBN 978-952-488-145-6.
- W:427. MARKKU KALLIO – ANTTI PIRJETÄ: Incentive Option Valuation under Imperfect Market and Risky Private Endowment. 2007. ISBN 978-952-488-146-3.
- W:428. ANTTI PIRJETÄ – SEPPO IKÄHEIMO – VESA PUTTONEN: Semiparametric Risk Preferences Implied by Executive Stock Options. 2007. ISBN 978-952-488-147-0.
- W:429. OLLI-PEKKA KAUPPILA: Towards a Network Model of Ambidexterity. 2007. ISBN 978-952-488-148-7.
- W:430. TIINA RITVALA – BIRGIT KLEYMANN: Scientists as Midwives to Cluster Emergence. An Interpretative Case Study of Functional Foods. 2007. ISBN 978-952-488-149-4.
- W:431. JUKKA ALA-MUTKA: Johtamiskyvykkyyden mittaaminen kasvuyrityksissä. 2007. ISBN 978-952-488-153-1.
- W:432. MARIANO LUQUE – FRANCISCO RUIZ – KAISA MIETTINEN: GLIDE – General Formulation for Interactive Multiobjective Optimization. 2007. ISBN 978-952-488-154-8.
- W:433. SEPPO KINKKI: Minority Protection and Information Content of Dividends in Finland. 2007. ISBN 978-952-488-170-8.
- W:434. TAPIO LAAKSO: Characteristics of the Process Supersede Characteristics of the Debtor Explaining Failure to Recover by Legal Reorganization Proceedings. 2007. ISBN 978-952-488-171-5.
- W:435. MINNA HALME: Something Good for Everyone? Investigation of Three Corporate Responsibility Approaches. 2007. ISBN 978-952-488-189.
- W:436. ARTO LAHTI: Globalization, International Trade, Entrepreneurship and Dynamic Theory of Economics.The Nordic Resource Based View. Part One. 2007. ISBN 978-952-488-191-3.
- W:437. ARTO LAHTI: Globalization, International Trade, Entrepreneurship and Dynamic Theory of Economics.The Nordic Resource Based View. Part Two. 2007. ISBN 978-952-488-192-0.

- W:438. JANI KILPI: Valuation of Rotable Spare Parts. 2007. ISBN 978-952-488-197-5.
- W:439. PETRI ESKELINEN – KAISA MIETTINEN – KATHRIN KLAMROTH – JUSSI HAKANEN: Interactive Learning-oriented Decision Support Tool for Nonlinear Multiobjective Optimization: Pareto Navigator. 2007. ISBN 978-952-488-200-2.
- W:440. KALYANMOY DEB – KAISA MIETTINEN – SHAMIK CHAUDHURI: Estimating Nadir Objective Vector: Hybrid of Evolutionary and Local Search. 2008. ISBN 978-952-488-209-5.
- W:441. ARTO LAHTI: Globalisaatio haastaa pohjoismaisen palkkatalousmallin. Onko löydettävissä uusia aktiivisia toimintamalleja, joissa Suomi olisi edelleen globalisaation voittaja? 2008. ISBN 978-952-488-216-3.
- W:442. ARTO LAHTI: Semanttinen Web – tulevaisuuden internet. Yrittäjien uudet liiketoimintamahdollisuudet. 2008. ISBN 978-952-488-217-0.
- W:443. ARTO LAHTI: Ohjelmistoteollisuuden globaali kasvustrategia ja immateriaalioikeudet. 2008. ISBN 978-952-488-218-7.
- W:444. ARTO LAHTI: Yrittäjän oikeusvarmuus globalisaation ja byrokratisoitumisen pyörteissä. Onko löydettävissä uusia ja aktiivisia toimintamalleja yrittäjien syrjäytymisen estämiseksi? 2008. ISBN 978-952-488-219-4.
- W:445. PETRI ESKELINEN: Objective trade-off rate information in interactive multiobjective optimization methods – A survey of theory and applications. 2008. ISBN 978-952-488-220-0.
- W:446. DEREK C. JONES – PANU KALMI: Trust, inequality and the size of co-operative sector – Cross-country evidence. 2008. ISBN 978-951-488-221-7.
- W:447. KRISTIINA KORHONEN – RIITTA KOSONEN – TIINAMARI SIVONEN – PASI SAUKKONEN: Pohjoiskarjalaisten pienten ja keskisuurten yritysten Venäjä-yhteistyöpotentiaali ja tukitarpeet. 2008. ISBN 978-952-488-228-6.
- W:448. TIMO JÄRVENSIVU – KRISTIAN MÖLLER: Metatheory of Network Management: A Contingency Perspective. 2008. ISBN 978-952-488-231-6.
- W:449. PEKKA KORHONEN: Setting “condition of order preservation” requirements for the priority vector estimate in AHP is not justified. 2008. ISBN 978-952-488-242-2.
- W:450. LASSE NIEMI – HANNU OJALA – TOMI SEPPÄLÄ: Misvaluation of takeover targets and auditor quality. 2008. ISBN 978-952-488-255-2.
- W:451. JAN-ERIK ANTIPIN – JANI LUOTO: Forecasting performance of the small-scale hybrid New Keynesian model. 2008. ISBN 978-952-488-261-3.
- W:452. MARKO MERISAVO: The Interaction between Digital Marketing Communication and Customer Loyalty. 2008. ISBN 978-952-488-266-8.
- W:453. PETRI ESKELINEN – KAISA MIETTINEN: Trade-off Analysis Tool with Applicability Study for Interactive Nonlinear Multiobjective Optimization. 2008. ISBN 978-952-488-269-9.

- W:454. SEPPÖ IKÄHEIMO – VESA PUTTONEN – TUOMAS RATILAINEN: Antitakeover provisions and performance – Evidence from the Nordic countries. 2008. ISBN 978-952-488-275-0.
- W:455. JAN-ERIK ANTIPIN: Dynamics of inflation responses to monetary policy in the EMU area. 2008. ISBN 978-952-488-276-7.
- W:456. KIRSI KOMMONEN: Narratives on Chinese colour culture in business contexts. The Yin Yang Wu Xing of Chinese values. 2008. ISBN 978-952-488-279-8.
- W:457. MARKKU ANTTONEN – MIKA KUISMA – MINNA HALME – PETRUS KAUTTO: Materiaalitehokkuuden palveluista ympäristömyötäistä liiketoimintaa (MASCO2). 2008. ISBN 978-952-488-279-8.

Z-SARJA: HELSINKI SCHOOL OF ECONOMICS.
CENTRE FOR INTERNATIONAL BUSINESS RESEARCH. CIBR WORKING PAPERS. ISSN 1235-3931.

- Z:16. PETER GABRIELSSON – MIKA GABRIELSSON: Marketing Strategies for Global Expansion in the ICT Field. 2007. ISBN 978-952-488-105-0.
- Z:17. MIKA GABRIELSSON – JARMO ERONEN – JORMA PIETALA: Internationalization and Globalization as a Spatial Process. 2007. ISBN 978-952-488-136-4.

Kaikkia Helsingin kauppakorkeakoulun julkaisusarjassa ilmestyneitä julkaisuja voi tilata osoitteella:

KY-Palvelu Oy
Kirjakauppa
Runeberginkatu 14-16
00100 Helsinki
Puh. (09) 4313 8310, fax (09) 495 617
Sähköposti: kykirja@ky.hse.fi

Helsingin kauppakorkeakoulu
Julkaisutoimittaja
PL 1210
00101 Helsinki
Puh. (09) 4313 8579, fax (09) 4313 8305
Sähköposti: julkaisu@hse.fi

All the publications can be ordered from

Helsinki School of Economics
Publications officer
P.O.Box 1210
FIN-00101 Helsinki
Phone +358-9-4313 8579, fax +358-9-4313 8305
E-mail: julkaisu@hse.fi