The Long-Term Effects of Forced Migration*

Matti Sarvimäki Roope Uusitalo Markus Jäntti

May 2010

Abstract

We examine the long-term effects of internal displacement using individual-level panel data. After World War II, Finland evacuated 430,000 people from areas ceded to the Soviet Union and resettled them to the remaining parts of the country. A quarter of a century later, the displaced earned substantially more than plausible control groups. The effect was limited to men who lived in rural areas before the war. We explain these income gains by a faster transition from traditional to modern occupations and from rural to urban areas.

JEL classification: J61, O15, R23
Keywords: Displaced persons, migration, regional labor markets

*This is a revised version of IZA DP 4003 / SERC DP 15. We thank Josh Angrist, David Card, Christian Dustmann, Steve Gibbons, Albrecht Glitz, Kristiina Huttunen, Guy Michaels, Tuomas Pekkarinen, Marie Schultz-Nielsen, Daniel Sturm, Olof Åslund and seminar participants at Fordham, HECER, LSE, NYU, UCL, Uppsala, the Econometric Society, the European Society for Population Economics, the Nordic Summer Institute in Labor Economics, the Royal Economic Society, the Society of Labor Economists, the Nordic Migration Workshop and the Finnish Economic Association for helpful comments and Tiina Hytönen, Nina Intonen and Anni-Mari Karvinen for able research assistance. Financial support from the NORFACE project "Migration: Integration, Impact and Interaction" and the Yrjö Jahnsson Foundation is gratefully acknowledged. Sarvimäki also acknowledges financial support from the Jenny and Antti Wihuri Foundation. *Sarvimäki* (corresponding author): Government Institute for Economic Research and SERC/CEP, London School of Economics, m.sarvimaki@lse.ac.uk. *Uusitalo*: Government Institute for Economic Research, IFAU and IZA, roope.uusitalo@vatt.fi. *Jäntti*: Stockholm University, markus.jantti@iki.fi.

1 Introduction

At the end of year 2007, some 26 million people were internally displaced by armed conflicts and another 25 million by natural disasters (UNHCR, 2008). In addition, there were roughly 16 million international refugees and an undocumented number of people forced to migrate due to policies aimed at enhancing economic development. For example, dam construction alone has displaced between 40 and 80 million people in the past half a century (World Commission on Dams, 2000).

Forced migration is likely to have important consequences for those who are displaced as well as for the sending and the receiving areas. Yet, we know little about the magnitude—or even the direction—of these effects. To a large extent, the lack of evidence follows from the rarity of high-quality data on a displaced population and an appropriate control group.

In this paper, we examine the long-term effects of internal displacement among people who were forced to migrate in Finland in the 1940s. After World War II, Finland ceded its eastern parts to the Soviet Union and evacuated the entire population living in these regions. In total, 430,000 individuals—11 percent of the 1940 population—were resettled to the remaining parts of the country and were paid a partial compensation for their lost property. In particular, displaced farmers were given land and assistance to establish small new farms.

We focus on the effects on long-term income and on the transitions between occupations and geographical areas. To perform the analysis, we have access to individual-level longitudinal data recording demographic characteristics and labor market outcomes in 1939, 1950 and in the early 1970s. These data allow us to follow more than 20,000 people who remain of working age throughout the period under study. All information for the displaced and non-displaced populations come from the same sources and is thus strictly comparable.

We evaluate the effect of being displaced by comparing the post-war labor market outcomes between those who lived in the ceded areas before the war and those who lived in the areas that remained part of Finland.

Figure 1 illustrates our approach using data on men who lived in rural areas before the war. It plots the log income in 1971 against the distance from the municipality of residence in 1939 to the post-war border. Average incomes tend to be higher among men who lived in Western Finland before the war. However, the men who lived east of the new border before the war (and were therefore displaced) earned substantially more in 1971 than those who lived west of the border before the war. We will argue that this discontinuous jump at the post-war border represents a causal effect of forced migration.

The validity of this argument hinges on the assumption that the displaced population does not differ from the rest of the population in unobservable ways that would affect long-term labor market outcomes. We provide several pieces of evidence supporting this identifying assumption. First, a review of historical studies suggests that the location of the new border was not determined by regional economic differences. Second, we subject the estimates to several robustness checks and falsification exercises. We show that there were no large differences in socioeconomic status, sector of employment or local authority level per capita incomes before the war and that our conclusions are not sensitive for controlling for the remaining differences. We report estimates from ordinary least-squares and regression-discontinuity specifications using both a sample of the entire Finnish population and a sub-sample of individuals who lived before the war within 100 kilometers of the post-war border. Furthermore, we experiment with alternative approaches in measuring income. All these approaches yield qualitatively similar findings.

Our main result is that forced migration increased the long-term income among men who lived in rural areas before the war. We find no significant impact on the income of women or on the income of the urban population. We attribute the income gains to faster transition towards the modern economy. The estimates suggest that the displacement increased the likelihood of moving from rural areas to cities and from agriculture to the other sectors. The displacement also increased regional mobility even after the resettlement had been completed. We speculate that this may reflect a decrease in mobility costs due to weakening of economic and social ties, and review survey evidence from the late 1940s supporting this hypothesis.

These findings are in contrast with previous results on the economic impacts of forced migration. Ibáñez and Vélez (2008) and Kondylis (2007, 2008) report large negative effects of conflict-induced displacement in Colombia, Bosnia-Hertzegovina, and Rwanda. Similarly, Vigdor (2007) and Groen and Polivka (2008) find that being evacuated from the New Orleans area due to Hurricane Katrina had a negative impact on employment and income. Interestingly, however, Sacerdote (2008) concludes that the students who had to move from poorly performing New Orleans schools to better schools elsewhere improved their test scores after the hurricane.

In comparison to the previous studies, our research design provides several attractive features. First, we are able to examine a period extending over three decades and therefore to measure truly long-term effects. The long period may also explain the differences between our results and those from the previous studies. Second, we study permanent displacement. Thus our estimates are easier to interpret than those arising from conflicts or natural disasters that lead to a relatively short period of absence and return to the partly destroyed home area. Third, we examine a situation where the entire population living in a certain area is forced to move and thus avoid estimation problems arising from self-selection into migration.

The last feature makes the results relevant also for the more general literature on the effects of mobility. A key result of the previous research is that migrants are a self-selected subgroup of the population.[1] As a consequence, correlations between migrant status and outcomes of interest in a typical observational dataset cannot be interpreted as a causal relationship. In our case, we are able to estimate the impact of forced migration for an essentially randomly chosen individual in the sample.

The rest of this paper is organized as follows. The next section provides the historical background and the details of the implemented settlement

[1] Borjas (1987), Chiquiar and Hanson (2005) and Banerjee and Newman (1998) present formal models on selective migration. The empirical literature on the topic is too large to be summarized here. Recent contributions include, but are not limited to, Chiquiar and Hanson (2005), Grogger and Hanson (2008), Miguel and Hamory (2009), Gibson and McKenzie (forthcoming), Jaeger et al. (forthcoming), McKenzie and Rapoport (forthcoming) and McKenzie et al. (forthcoming).

policy. Section 3 describes the data and Section 4 discusses our empirical approach. Section 5 presents the results and robustness checks. We discuss the implications of the results in Section 6 and present our conclusions in Section 7.

2 The Historical Background

In August 1939, just before the outbreak of the Second World War, Germany and the Soviet Union signed a non-aggression pact that included a secret clause in which Eastern Europe was divided between the two powers.[2] Finland was consigned to the Soviet sphere of influence. Shortly after, Finnish negotiators were summoned to Moscow. The Soviet Union demanded that, in order to secure the defense of Leningrad, parts of Finnish territory in the southeast should be ceded to the Soviet Union and that Soviet naval bases should be established in Finland. The talks broke down in mid-November and a few weeks later the Red Army attacked Finland. In the peace treaty ending the hostilities 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. In July 1940, the Finnish Parliament enacted an Emergency Settlement Act (*Pika-asutuslaki*) guiding the resettlement policy.

The resettlement was still in an early stage in June 1941 when Finland joined Germany in its attack on the Soviet Union. By the end of August, the Finnish troops had reoccupied the ceded areas and in early December, the Parliament declared them re-united with the rest of Finland. Roughly two thirds of all displaced persons—and almost all the displaced farmers—returned to their pre-war homes (Pihkala, 1952; Waris et al., 1952).

After two and half years of trench warfare, the Soviet Union launched a massive attack on June 9th, 1944. The Finnish troops were pushed towards the west and the entire defense was close to collapsing. However, the Finns managed to stop the Red Army in early July. By then, the Soviet troops were more urgently needed on the Baltic front. Peace negotiations were started in August, and an armistice was signed in September. The armistice that

[2] Unless otherwise stated, this section draws from Kirby (2006).

was later ratified in the Paris Peace Treaty with the Allied Powers in 1947 restored the 1940 borders with some additional areas ceded to the Soviet Union. The entire population living in the ceded area was again evacuated and resettled. Finland also agreed to pay USD300 million in war reparations and to expel German troops from its territory.

Figure 2 presents the pre-war and the post-war borders. As we explain in detail in the next section, we will use those living west of the new border before the war as a control group for those living in the ceded areas. The validity of this approach would be called into question if the new border had been determined by regional economic differences. However, this does not seem to be the case. The 1944 border followed roughly the border set in the treaty of Nystad in 1721. Thus the location of the 1721 border determined whether the persons living in Eastern Finland were displaced or not.[3]

Importantly, there were many historical borders to choose from. Finland was part of Sweden until 1809 and the Swedish-Russian border had been moved several times. Recent archive findings indicate that the US initially planned to propose to their Soviet allies a peace treaty with Finland based on the 1920 borders—roughly one hundred kilometers east of the current border. When the Soviet Union offered peace talks in March 1944, it was preparing to negotiate on the basis of 1743 borders—roughly sixty kilometers west of the current border. However, when the peace talks began in August 1944, the unexpected success of the Finnish troops had improved Finland's position in the negotiations and thus moderated the Soviet demands. (Rentola, 2001)

It seems reasonable to think that in terms of individual-level earnings potential, the location of the 1944 border can be considered as good as randomly assigned. The new border split the historical province of Karelia in half. Areas close to the post-WWII border had been part of the same country for two hundred years, and areas further west for 135 years, belonging first to the Russian Empire and then to Finland after she became independent in

[3]In addition to ceding the southeastern part of the country, Finland ceded a sparsely populated Petsamo area in the North and leased the Porkkala Peninsula near Helsinki for a Soviet naval base for fifty years. However, improvements in international relations and changes in military technology led the Soviet Union to return Porkkala to Finland in 1956. Persons living in the Porkkala area are excluded from our sample.

1917. Furthermore, Figure 2 illustrates that average incomes were roughly similar before the war on both sides of the post-war border. In Section 5, we also show that the average pre-war income in the ceded area was close to the national average.

Resettling the 430,000 displaced persons was a major challenge. Finland was still a developing country that had won independence just two decades earlier, gone through a short but brutal civil war in 1918 and then evolved into a fairly well functioning democracy. One half of the population was working in agriculture, typically owning small farms and working as hired labor in forest work during the winter. According to Maddison (2010), GDP per capita was 3,589 International Geary-Khamis 1990 dollars in 1938.[4] Furthermore, the war had left Finland with approximately 95,000 dead and 228,000 injured out of a 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 the war reparations. For example, a quarter of the Finnish commercial fleet was handed over to the Soviet Union. Altogether the war reparations took about 15 percent of the government budget between 1945 and 1949 (Tarkka, 1988).

Despite this grave economic situation, the Parliament approved a series of laws in 1945 that offered compensation for the property lost due to the displacement. The rate of the compensation varied from full reimbursement for small losses to compensations of only ten percent for very large losses. Those who had lost property filed a claim reporting their losses and, on average, roughly forty percent of the claimed property was compensated (Waris et al., 1952). The urban population received their compensations mostly in the form of inflation-indexed government bonds for which a liquid secondary market quickly emerged. Those who had owned or rented land in the ceded areas were given agricultural and forest land.

The resettlement was financed by levying a massive tax on capital. Land for the settlers was primarily taken from the state, the local authorities (municipalities) and the church, but the required amount far exceeded the

[4]In comparison, Maddison (2010) reports similar GDP per capita figures for Morocco, Algeria, Moldova, Serbia/Montenegro/Kosovo, Jamaica, Egypt and Cuba in 2008.

capacity of the public sector. Thus, roughly two thirds of the cultivated fields, one half of the land that could be cleared for cultivation and a third of forest land was seized from private owners. Land were seized using an explicit progressive expropriation schedule. Similarly, a heavy progressive tax was set on other forms of property.[5]

The implementation of the Land Acquisition Act was entrusted to the Department of Land Settlement in the Ministry of Agriculture. Each ceded village was settled to a designated target 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 and very few were allocated to the Swedish-speaking municipalities on the Western and Southern coasts.

The non-agrarian population was free to settle wherever they could find accommodation and employment. Cities in Eastern Finland received flows of displaced persons that constituted almost ten percent of the population, while cities further west and cities with the most severe housing problems received much less.[6] While those who had not owned land were not explicitly allocated, the settlement plan influenced their migration also 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).

[5]The schedule for farm land required private land owners to cede up to 80 percent of their land holdings depending on the size of their farms. No land was expropriated from farms smaller than 25 hectares. Landowners were compensated with government bonds yielding 4 percent nominal interest. Inflation eventually wiped out about four fifths of their value. However, the bonds could be used for paying the Property Expatriation Tax, which was collected from all capital owners. Pihkala (1952) discusses the land acquisition policy in detail and argues that landowners did not suffer more than other owners of property.

[6]The share of the displaced population in urban areas in 1948 varied from nine percent in Mikkeli, eight percent in Jyväskylä and Lahti (all located in Central or Eastern Finland) to three percent in Pori (on the west coast) and less than half a percent in Pietarsaari (a Swedish-speaking town on the western coast).

3 Data

Our primary source of data is the 1950 population census. Data were collected by personal interviews and the information on 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 manually inserted into a database. The resulting sample contained about 114,000 dwelling units with 411,629 persons from 392 of municipalities (out of a total of 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 (introduced in 1964) of each person. Thus, in order to end up in the final sample, the person had to be alive and be living in Finland in the mid-1960s. Social security numbers allowed merging the 1950 census data with information from the 1970 census and the 1971 tax records.[7] Statistics Finland (1998) provides a detailed discussion of these 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, it also contains retrospective information concerning municipality of residence 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 micro-data on the displaced and the non-displaced persons from the pre- and post-displacement periods. The same survey instruments were administered to both groups and hence all the information is fully comparable.

Linking the 1950 census to later data further increases the amount of available information. Most importantly, the 1970 census file includes tax

[7] Ideally 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.

records data from 1971. This provides an opportunity to evaluate the long-term effects of migration on income more than a quarter of a century after the war had ended. Furthermore, the data include information on the residence municipality and industry of employment, which allow us to study transitions from the traditional to the modern economy.

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 using the 1971 microdata and annual tables of taxable income from local tax boards. The procedures are discussed in detail in the Appendix.

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. These cohorts were 14–32 years old in 1939 and 46–64 years old in 1971. Moreover, we exclude observations for which the municipality of residence in 1939 is unknown (360 persons) and those who lived in partly ceded municipalities (971 persons) or on the Åland Islands (176 persons). This leaves us with information on 22,896 individuals of whom 2,277 were displaced.

4 Empirical Strategy

We 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. The basic equation that we estimate is

$$y_{ijt} = \alpha D_i + \mathbf{X_{0i}}\beta + \varepsilon_{ijt} \tag{1}$$

where y_{ijt} is the outcome of interest for individual i living in location j at time t, D_i is an indicator variable taking value one if the person was living in the ceded area before the war and zero otherwise, $\mathbf{X_{0i}}$ 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

$$\varepsilon_{ijt} = u_{jt} + \nu_{it} \tag{2}$$

where u_{jt} captures the unobserved "quality" of the labor market j at time t and ν_{it} represents individual-specific unobserved factors.

Our identifying assumption is that, once we condition on the observed pre-war characteristics, displacement status is uncorrelated with unobserved individual characteristics

$$Cov\left(D_i, \nu_{it} | \mathbf{X_{0i}}\right) = 0. \tag{3}$$

Since the location of the new border was determined by an interaction of battles and historical borders, this assumption seems plausible. Those who happened to live east of the new border 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. Below, we provide two types of evidence suggesting that this does not drive the results. First, the available data suggest that there were few important pre-war differences between the future displaced and non-displaced persons. Second, the conclusions are not sensitive to controlling for pre-war observable characteristics, to limiting the sample to a 100 kilometer region on both sides of the new border (see Figure 2) or using a regression-discontinuity specification (discussed below).

The key to interpreting the estimates is to note that the correlation between displacement status (D_i) and post-war labor market quality (u_{jt}) is likely to be positive. The reason is that the resettlement moved individuals to new labor markets and occupations. While the displaced had limited opportunities to choose their initial destination, the authorities assigned more displaced persons to more prosperous municipalities. Furthermore, the displacement appears to have hastened the transition from traditional to mod-

ern economy and increased geographical mobility even after the resettlement. Internal migration later on is likely to be correlated with job opportunities. If the displaced had lower costs of moving, 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

$$\text{plim } \widehat{\alpha} = \pi_D + \theta \tag{4}$$

where π_D is the partial correlation between displacement status and labor market quality and the parameter θ captures the direct effects of being displaced. That is, the impact of forced migration works through two channels. The first is moving to better labor markets. The second is "other" effects due to, for example, loss of location-specific human capital, trauma or loss of property.[8]

Given the assumption (3), $\widehat{\alpha}$ has a causal interpretation as it measures the difference in expected outcomes between two well-defined counter-factual states. However, it is important to acknowledge that large-scale migration flows affected the entire population of post-war Finland. Hence, the estimates should be interpreted using a thought experiment where one would manipulate the displacement status of a single individual, while 430,000 others are still forced to migrate.

We also report regression-discontinuity (RD) estimates, where the specification includes the distance of the pre-war residence municipality to the post-war border (and its interaction with the displacement status). The interpretation of these estimates also warrants some discussion. The motivation for using the RD specification is that economic opportunities or unobserved earnings potential might have differed across regions. In partic-

[8]We interpret labor markets as an interaction between location and occupation, To derive probability limit (4), suppose that the data generating process is $y_{ijt} = \theta D_i + \mathbf{X_{0i}}\beta + u_{jt} + \nu_{it}$. Equation (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 $\mathbf{X_{0i}}$. Note that the likely presence of the direct effects, θ, prevents us from using the displacement status as an instrument for later migrations. Thus, the research design does not allow for estimating returns to voluntary mobility.

ular, Figure 1 suggests that those living in the western part of the country may have had some advantages in comparison to those living in the east. If true, our baseline estimates would be biased downwards. However, the caveat to the RD approach is that the control areas close to the new border may have been disproportionately affected by the shift of the border. For instance, Redding and Sturm (2008) find that the division of Germany led to a decline of West German cities close to the East-West German border. If something similar happened in the eastern parts of Finland, the local comparisons at the new border would lead to a larger (positive) effect of forced migration than our baseline estimates.

We will also discuss estimates from specifications adding post-war labor market fixed effects to the equation (1). Under strong assumptions, these specifications would identify the direct effects of being displaced (θ in equation 4). However, these assumptions are likely to be violated and thus we do not interpret this set of estimates as measuring a causal relationship.[9] Nevertheless, they contain useful information. Namely, they can be interpreted as average differences between a displaced person and a non-displaced person who were similar in their pre-war characteristics and who lived in similar labor markets in 1970.

In all regressions, inference is based on standard errors that are clustered at the 1939 residence municipality level. This choice of the level of clustering is motivated by the assumption that persons coming from the same areas might have been exposed to common shocks, in particular as the settlement plan affected individuals based on their pre-war residence municipality. Fur-

[9]The causal interpretation would require that the displacement status and individual unobservable factors were uncorrelated *conditional on labor market quality*, $Cov\left(D_i, \nu_{it} | \mathbf{X_{0i}}, u_{jt}\right) = 0$. This is a stronger assumption than our identifying assumption, $Cov\left(D_i, \nu_{it} | \mathbf{X_{0i}}\right) = 0$, as it would require either that displacement status and labor market quality are not correlated, $Cov(D_i, u_{jt} | \mathbf{X}) = 0$, or that unobserved individual characteristics and labor market quality are uncorrelated, $Cov(u_{jt}, \nu_{it} | \mathbf{X}) = 0$. We have already argued that the first condition does not hold. The second condition would imply that individuals do not sort into localities based on their unobservable characteristics. This is also unlikely to be true. For instance, Jaeger et al. (forthcoming) find that individuals who are more willing to take risks are more likely to migrate between labor markets in Germany, and Miguel and Hamory (2009) report a positive relationship between cognitive ability and subsequent migration in Kenya.

thermore, distance to the post-war border is measured using municipality midpoints and hence the RD estimates are based on a discrete assignment variable. Lee and Card (2008) show that this introduces a group structure in the standard errors. Since the discretization of the assignment variable occurs at the 1939 residence municipality level, our choice of the clustering level also adjusts the standard errors to this group structure.

5 Results

This section reports our results. In the first subsection, we show that the pre-war differences between the displaced persons and the rest of the population were small and can be controlled for by using standard linear regression models. We then present estimates suggesting that forced migration increased income among men living in rural areas before the war. For other groups, we do not find statistically significant effects. In the last subsection, we argue that these effects can be explained by faster urbanization and transition from agricultural to modern occupations.

5.1 Pre-War Differences

Tables 1, 2 and 3 report sample means of available pre-war covariates. We follow Imbens and Wooldridge (2009) and report the means of all covariates in both groups and their normalized differences (see table notes for details). We do this separately for the same sub-samples that are used in the regression analysis. That is, we divide the sample according to gender, pre-war municipality type and the proximity of the pre-war residence municipality to the post-war border.

Table 1 reports average pre-war characteristics for men living in rural areas. We find very small differences in terms of age or imputed income. The most important difference is that the displaced men were less likely to be blue-collar workers and less likely to speak Swedish as their native tongue. They were also more likely to work in services. Average income in the municipality of residence is lower among the displaced if one compares the regions

close to the border, but similar if one uses data for the entire country. However, these differences are reasonably small in magnitude. Large differences can only be found in the share of Swedish speakers, a small minority group that is heavily concentrated on the western coast of Finland. Table 2 reports similar results for rural women. Again, displaced persons are less likely to be Swedish speakers or to be blue-collar workers.

Table 3 reports the same statistics for those who lived in urban areas in 1939. We report the sample means only for the entire country due to the small number of cities. In comparison to the rural population, the difference in the fraction of Swedish speakers is larger. Furthermore, the average income in the urban areas is lower in the ceded area, while imputed incomes are similar on both sides of the border. In terms of socioeconomic status, the most important difference is that displaced women are less likely to be blue-collar workers. A related difference can be observed in the industry structure. Both displaced men and women were less likely to be employed in manufacturing before the war. Among men, the displaced were also more likely to work in services.

To summarize, the pre-war covariate distributions among the displaced and non-displaced populations were reasonably similar. With the exception of a fraction of Swedish speakers and the average income in urban areas, the normalized differences are below 0.15.[10] Nevertheless, we note that some differences exist. Therefore we will primarily concentrate on the estimates that control for the pre-war differences and subject our results to several robustness checks in the analysis below.

5.2 Impact on Long-Term Income

Table 4 reports the results when we regress log annual taxable income measured in 1971 on the displacement status. Each coefficient stems from separate regressions which differ in the estimation sample used (columns) and the control variables included (rows).

[10] Imbens and Wooldridge (2009) suggest as a rule of thumb that controlling for covariates with linear regression tends to be sensitive to the specification if these differences exceed 0.25.

Panel A reports the main estimates. The top-left estimates are for men living in rural areas before the war. The first coefficient reveals that those who were forced to move during the war earned roughly 19 percent more in 1971 than those whose homes remained on the Finnish side of the post-war border. When we control for the pre-war characteristics, the point estimate decreases to 11 percent, but remains highly statistically significant.[11] Focusing only on those who were living within 100 kilometers of the new border before the war yields similar results, though the estimates are less precise. Dropping Swedish speakers from the data has virtually no impact on the estimates (not reported in the Table). Interestingly, however, none of the estimates for the urban population or for rural women is statistically significant.

In the lower part of Table 4, we examine the robustness of these results. Panel B reports estimates from alternative specifications. All of these specifications control for post-war covariates that are potentially affected by displacement. Hence, these estimates cannot necessarily be interpreted as causal effects. The first specification adds our earliest observation of the level of education (measured in 1950) to the control variables. The motivation is to address the concern that post-war income differences could be due to pre-war differences in education. While post-war education may have been affected by the displacement for some in the younger cohorts, most of the individuals in our sample had finished their formal schooling by 1939. In any case, including the 1950 level of education to the specification has little impact on the estimates.

The motivation for the next two specifications is to assess whether the results are sensitive to the way that income is measured. Here, the first concern is that taxable income may be a poor measure of true consumption

[11] The control variables are age (and its square), imputed earnings in 1939 (and its square), an indicator for speaking Swedish as one's mother tongue, six categories for socioeconomic status (entrepreneur, white-collar worker, blue-collar worker, assisting family member, out of the labor force), four categories for the sector of employment (agriculture, manufacturing, construction and services), latitude of the 1939 residence municipality and five categories of taxable income per capita (quintile groups) in the 1939 residence municipality.

possibilities or productivity. In particular, farmers could consume part of their own harvest, which would not be captured by taxable income. Below, we also show that the displaced farmers were more likely to move from agriculture to other sectors. Thus, the under-estimation of agricultural income would lead to an overestimate of the impact of the displacement on income. We assess this possibility by excluding those working in agriculture in 1970. The caveat to this approach is that displacement affected sorting into occupations, which may affect the estimates. However, in practice, this sub-sample yields very similar results to those obtained from the full sample.

In the next specification, we have changed the outcome variable from nominal to real income, taking into account regional price differences.[12] While nominal income is likely to reflect changes in productivity due to the displacement, it might not be informative about changes in purchasing power. The displaced were more likely to move from rural to urban areas and at least part of the gains in nominal income was probably lost due to the resulting higher rents or housing prices. As a consequence, part of the surplus created by migration might have been transferred to landlords in the areas attracting many displaced persons. In line with this hypothesis, the point estimates for the impact on real income among rural men are smaller than those for nominal income. Yet, they remain positive and, when we use the entire country as a control area, statistically significant. The estimates among rural women and the urban population are very similar to the estimates for nominal income and remain statistically insignificant.

The last row of panel B adds a full set of 1970 municipality of residence fixed effects to the specification. We note, again, that sorting across municipalities was probably affected by the displacement and thus these estimates do not have a causal interpretation. Nevertheless, it is interesting to note that adding fixed effects reduces the point estimates. While they remain

[12]We calculate real income by dividing the nominal income by the local price index. These price data were collected in the 1971 cost of living study that collected the 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 the local price level was mainly due to differences in (quality adjusted) rental prices. (Statistics Finland, 1972)

positive, they are no longer significantly different from zero.

Panel C reports regression-discontinuity (RD) estimates controlling for the distance to the post-war border and its interaction with displacement status. In most cases the results are qualitatively unaffected. The only qualitative difference between the OLS and RD results is that the estimate for urban men is significantly positive and the estimate for urban women significantly negative in the RD specification.

Another difference between the OLS and RD estimates is that the latter are less precise. When we focus on the 100 kilometer buffer zone, the 95% confidence interval for the baseline estimate for rural men is $[-0.05, 0.42]$. Furthermore, the impact of the distance to the border is very imprecisely estimated. On the western side of the border, even the sign of the distance to the border coefficient is sensitive to the choice of bandwidth. These observations suggest that our sample size is insufficient for conducting RD analysis using only individuals who lived close to the post-war border in 1939.

Precision is improved when we extend the estimation sample to include the entire country. The distance to the post-war border estimates are now statistically significant and suggest a positive association between the 1971 income and living further west in 1939. However, there is a large jump at the post-war border suggesting that being forced to migrate increased long-term income. The RD estimates are larger than the OLS estimates. If these differences followed from those living in the western parts of the country having more favorable unobserved characteristics, we should interpret the OLS estimates as a lower bound for the impact of displacement. However, the association between distance to the post-war border and later income could also follow from the eastern parts of the control area being more adversely affected by the shift of the border.

Panel D reports results from a placebo experiment where the border would have moved roughly 60 kilometers further west to the 1743 border as demanded by the Soviet Union in March 1944 (see Figure 2). This would have displaced another 250,000 persons in addition to the 430,000 persons who lived in the area that was actually ceded. We discard from the data the persons who were actually displaced and then estimate the effect of

displacement that did not happen—but could very well have happened—by defining persons who lived east of this proposed border as the treatment group. The effects of this placebo treatment on men turn out to be small and statistically insignificant. For women the point estimates are large, but statistically insignificant. We also experimented with a set of placebo borders moving from the current border westwards in 50 km steps. These estimates (not reported in the Table) are generally negative indicating that incomes are higher in the west as shown already in Figure 1 and suggesting that, if anything, our estimates are biased downwards and would therefore understate the true (positive) income effect of the displacement.

Our final robustness check concerns non-random sample attrition. The concern here is that since individuals had to be alive and be living in Finland at the time of the introduction of the social security numbers in 1964, sample attrition due to differences in mortality or emigration rates could explain the results. However, as we discuss in detail in the Appendix, the data do not support the hypothesis of non-random attrition.

5.3 Impact on Potential Channels

Taken together, the results discussed above suggest a positive impact of displacement on long-term income among rural men. For other groups, most of the estimates are not statistically significant. We now turn to the potential channels that may explain these effects by examining the effect of displacement on various labor market outcomes in 1950 and in 1970. As before, we examine the effects separately for men and women, for rural and urban populations, and for the whole country and for the group living close to the post-war border. To aid in their interpretation, we first report the control group means and then the effect of displacement. These effects are estimated by regressing each outcome on the displacement status and the pre-war characteristics listed in the footnote 11.

Table 5 reports the medium-term effects for men. All outcomes are measured in 1950, six years after the end of the war. We find no significant impacts on employment rates or on non-agricultural entrepreneurship. The

effect on imputed income is positive for sub-groups that originally lived in the rural areas, but negative for the sub-group that lived in urban areas in 1939. Both of these effects reflect changes in the urbanization rate and industry structure. The effects on both of these variables are large. Displacement increases the likelihood of moving from rural to urban areas by 1950 by about ten percentage points, i.e. approximately doubles the urbanization rate. For persons who already lived in urban areas, we find the opposite effect.[13] We also find substantial effects on the distribution of workers across industries. The displaced who lived in rural areas before the war are much less likely to work in agricultural occupations in 1950 and more likely to work in manufacturing and construction than their non-displaced comparison groups. For the displaced persons who lived in urban areas before the war, we find no significant changes in the industry structure.

In Table 6, we report medium-term effects for displaced women. The results are largely similar to those for men. We find a large increase in the fraction moving from rural to urban areas and a large decrease in the fraction employed in agriculture. Among women, the shift away from agricultural occupations is related to an increase in the fraction employed in services. Also employment rates increase more among the displaced than among the rest of the population.

In Tables 7 and 8, we examine the long-term effects of displacement using labor market outcomes in 1970. Again, we find little impact on employment or self-employment rates. The gap in imputed earnings has remained close to the level that was observed in 1950. By 1970, the urbanization rate has increased rapidly in the whole population and the share working in agricultural occupations simultaneously declined. However, the effect of displacement is still roughly as large as it was in 1950. The displaced men are 15 percentage points more likely to live in urban areas than those who lived elsewhere in the country in 1939. Interestingly, the comparison group living close to the new border is much more likely to move to cities than those living farther away from the border. Consequently, the difference between the displaced

[13]Note, however, that our classification counts only cities as urban areas. Therefore, moving to rural areas could be movement to suburban areas close to cities.

and this comparison group is not statistically significant, though the point estimate remains large at six percentage points. The movement away from agricultural occupations has continued. The displaced are eight percentage points less likely to work in agriculture than the comparison group. The effects on displaced women reported in Table 8 are quite similar. The only qualitative difference is that displacement appears to have increased employment rates for women, although the effect is statistically significant only when comparing the border regions.

In the bottom rows of Tables 7 and 8, we report the effect of displacement on later mobility. More precisely, we estimate the effect of displacement on the likelihood of changing the municipality of residence between 1950 and 1970. We note that during these years, there was a lot of internal migration in Finland. Between 30 and 40 percent of the control group move across municipalities during this period. However, war-time displacement significantly increases also later mobility. The displaced are about ten percentage points more likely to move between municipalities than the rest of the population. The effects are significant in all sub-groups, except when the displaced men are compared to the men who lived close to the border. In this case, mobility is exceptionally high also in the control group.

6 Discussion

In most historical accounts the Finnish post-war settlement policy is regarded as a success. The settlement policy had broad public support. Financial aid to the displaced people was generous and provided both land and monetary compensation for lost property. The settlers were integrated into the society reasonably well and no major conflicts or riots arose. This is remarkable given that Finland was not a rich country at the time. In addition, disputes about land had been one of the main reasons that led to a civil war that killed more than one percent of the population only twenty years earlier.[14]

[14]The relative importance of the factors leading to the civil war in 1918 are still debated. However, the status of crofters (tenants of very small farms) clearly contributed to the conflict. The crofters had long demanded a regulated process that would allow them to

Nevertheless, it seems unlikely that even a well-designed resettlement policy alone would explain the finding that forced migration increased the long-term earnings of the displaced. We next consider some possibilities that are consistent with this finding and discuss the implications for current policy. This discussion is inevitably more speculative than the evidence presented above.

In our view, the most likely explanation is that increased mobility—both in terms of geography and occupation—contributed to the effect on income. Of course, the displaced were geographically more mobile during the war. Less obviously, they were more likely to move from traditional to modern occupations and remained more mobile even after the resettlement. This occurred despite the fact that they were settled in farming communities and had an opportunity to continue farming. Increased mobility seems to have paid off in economic terms in the rapidly urbanizing and industrializing post-war Finland, at least for rural men. The estimates for urbanization among rural women are comparable to those of the men, but the effects on women's income are not statistically significant. While this lack of significance is partly due to the precision of the estimates, the results are consistent with later migrations being driven by the labor market considerations of husbands, thus making their wives "tied-movers". Explaining the income gains as a consequence of a faster transition from the traditional to the modern sector is also in line with the finding that those living in cities before the war did not experience income gains.

A key ingredient in our interpretation is that displacement lowered the costs of geographical and occupational mobility after the war. The farms provided by the resettlement policy were on average smaller than the farms that were left in the ceded area. Some farms were 'cold', established on poor uncultivated land and lacked basic farm buildings (Mead, 1951). Harsh conditions in the settlement farms probably decreased the opportunity cost of moving. Later migration might also have been affected by more subtle reasons such as the lack of location-specific social capital. Waris et al. (1952)

buy the land they were renting, but such demands were successfully opposed by land owners. (Jäntti et al., 2006)

report survey evidence according to which disputes between the displaced and the local population were not uncommon.[15] For instance, the displaced were less likely than the local population to consider compensation paid for lost property fair. Although the locals reported reasonably positive attitudes towards the displaced persons, they seem to have had some doubts about their trustworthiness. For example, very few respondents disapproved dating of displaced and non-displaced persons. Yet, 40 percent of the locals stated that if they were to marry now, they would prefer a local spouse. When asked to state two positive characteristics of the displaced and the locals, 15 percent of the locals mentioned honesty when describing the local population, while only 2 percent mentioned honesty when describing the displaced population. These results suggest that the displaced persons may have lost informal networks that had economic value. This may have decreased the economic benefits of staying in the traditional sector.[16]

Our results can also be interpreted in the context of regional policy. We find that moving into the modern sector, which often entailed geographic mobility, was financially rewarding for the displaced. Further work would be needed to determine the extent to which our findings generalize beyond the specific context we examine. Nevertheless, we note that the findings are consistent with the argument that policies promoting mobility are likely to help economic growth. Furthermore, our findings suggest that large economic gains from migrating would have been available also for those who were not displaced, but many chose not to exploit these opportunities unless forced to do so. This indicates that the cost of mobility is high and that it is higher for persons with stronger social and economic ties to the regions where they live.

Finally, we stress that we measure the impact of displacement on income, not on utility. Losing areas that their families had sometimes occupied for

[15]The research project "The Social Adjustment of Displaced Persons in Finland" was launched in 1948 with funding from the Rockefeller Foundation. It contained two general surveys of both the displaced persons and local populations supplemented with in-depth interviews in two rural municipalities and in one industrial town.

[16]See Banerjee and Newman (1998) for a formal model on the role of social networks on rural-urban migration.

centuries was deeply painful for many of the displaced persons. A clear indication of their preferences is that two thirds of the displaced persons returned to the ceded area when Finland reoccupied the territories in 1941.

7 Conclusions

Evacuation of the Eastern parts of Finland after World War II created a 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 later occupational and regional mobility and, eventually, on long-term income.

References

Banerjee, A. and A. Newman (1998): "Information, the Dual Economy, and Development," *Review of Economic Studies*, 65, 631–653.

Borjas, G. (1987): "Self-Selection and the Earnings of Immigrants," *American Economic Review*, 77, 531–553.

Chiquiar, D. and G. Hanson (2005): "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States," *Journal of Political Economy*, 113.

Gibson, J. and D. McKenzie (forthcoming): "The Microeconomic Determinants of Emigration and Return Migration of the Best and Brightest: Evidence from the Pacific," *Journal of Development Economics*.

Groen, J. and A. Polivka (2008): "The Effect of Hurricane Katrina on the Labor Market Outcomes of Evacuees," *American Economic Review*, 98, 43–48.

Grogger, J. and G. Hanson (2008): "Income Maximization and the Selection and Sorting of International Migrants," NBER Working Paper No. 13821.

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.

Imbens, G. and J. Wooldridge (2009): "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, 47, 5–86.

Jaeger, D., H. Bonin, T. Dohmen, A. Falk, D. Huffman, and U. Sunde (forthcoming): "Direct Evidence on Risk Attitudes and Migration," *Review of Economics and Statistics*.

Jäntti, M., J. Saari, and J. Vartiainen (2006): "Growth and equity in Finland," United Nations University, World Institute for Development Economics Research.

Kirby, D. (2006): *A Concise History of Finland*, Cambridge: Cambridge University Press.

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," *Economic Development and Cultural Change*, 57, 31–66.

Lee, D. and D. Card (2008): "Regression Discontinuity Inference with Specification Error," *Journal of Econometrics*, 142, 655–674.

Maddison, A. (2010): "Statistics on World Population, GDP and Per Capita GDP, 1-2008 AD," Dataset available at http://www.ggdc.net/maddison/.

McKenzie, D., J. Gibson, and S. Stillman (forthcoming): "How Important is Selection? Experimental vs Non-Experimental Measures of the Income Gains from Migration," *Journal of the European Economic Association*.

McKenzie, D. and H. Rapoport (forthcoming): "Self-Selection Patterns in Mexico-US Migration: The Role of Migration Networks," *Review of Economics and Statistics*.

Mead, W. (1951): "The Cold Farm in Finland: Resettlement of Finland's Displaced Farmers," *Geographical Review*, 529–543.

Miguel, E. and J. Hamory (2009): "Individual Ability and Selection into Migration in Kenya," MPRA Paper No. 19228.

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

Redding, S. and D. Sturm (2008): "The Costs of Remoteness: Evidence from German Division and Reunification," *American Economic Review*, 98, 1766–1797.

Rentola, K. (2001): "Stalin, Mannerheim ja Suomen rauhanehdot 1944," *Historiallinen aikakauskirja*, 99, 47–61.

Sacerdote, B. (2008): "When the Saints Come Marching in: Effects of Hurricanes Katrina and Rita on Student Evacuees," NBER Working Paper No. 14385.

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.

——— (1998): *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.

Vigdor, J. (2007): "The Katrina Effect: Was There a Bright Side to the Evacuation of Greater New Orleans?" *The BE Journal of Economic Analysis & Policy*, 7, 64.

Waris, H., V. Jyrkilä, K. Raitasuo, and J. Siipi (1952): *Siirtoväen sopeutuminen*, Otava.

World Commission on Dams (2000): *Dams and Development: A New Framework for Decision-Making*, Earthscan Publications.

A Appendix

A.1 Attrition

To examine whether sample attrition is related to displacement status we compare the number of individuals from each municipality in our 1970 sample to the population of these municipalities in 1939. This relationship is by definition

$$N_{j70} = o_j S_j P_{j39} \xi_j \tag{5}$$

where N_{j70} is the number of individuals observed in micro-data in 1970 who lived in municipality j in 1939, o_j is the sampling rate for this municipality, S_j is the survival rate, P_{j39} is the population in 1939 and ξ_j captures measurement error in P_{j39}. The sampling rate varies across municipalities because Statistics Finland picked every tenth box of census forms and therefore small municipalities that fit into one box were either fully included or not included at all. Measurement error is mainly due to the age restriction in the 1970 sample. We use the entire population of the municipality in 1939 as a proxy for the number of persons in the cohorts that are included in the sample.

We assume that the survival rate can be modeled as

$$\ln S_j = \ln \bar{S}_j + \beta D_j + X_j \gamma + v_j \tag{6}$$

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 a vector of pre-war characteristics and v_j is an error term. Taking logs of (5) and substituting with (6), we get an equation

$$\ln N_{j70} = \alpha + \beta D_j + \eta \ln P_{j39} + X_j \gamma + \epsilon_{j70} \tag{7}$$

where $\alpha = \ln \bar{o} + \ln \bar{S}_j$, and $\epsilon_{j70} = (\ln \bar{o} - \ln o_j) + v_j + \ln \xi_j$. Under the assumption that ϵ is uncorrelated with the covariates, the parameter β is informative on whether attrition among the displaced differs from the rest

of the population. Furthermore, the fact that (5) is an accounting relation also suggests a natural specification check: we should find that $\eta = 1$.

Table A1 reports the results from regressing the number of individuals observed in 1970 in our microdata from each 1939 municipality on the total pre-war population of this municipality and a dummy indicating that the municipality was ceded after the war. We also control for pre-war municipality characteristics (listed in the footnote below the table). We find no evidence on systematic differences in attrition rates between the displaced and non-displaced. We acknowledge, however, that the estimates are rather imprecise.

A.2 Constructed Variables

Imputed income

We construct an 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 were intra-industry income 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. To mitigate the latter problem, we also use an alternative imputed income measure based on tables listing mean taxable income 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) and assign each individual the mean income of their reference group.

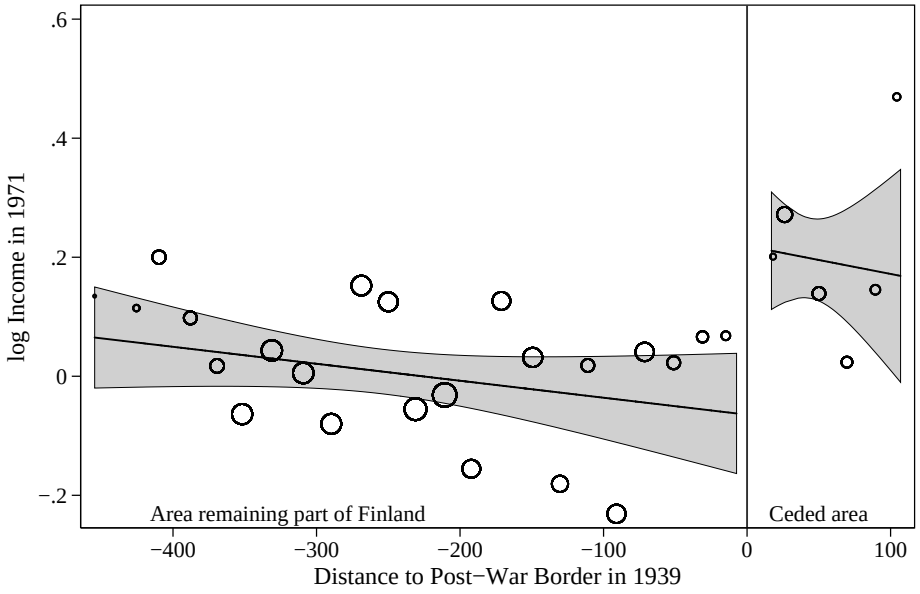
Lives in urban areas

Statistics Finland categorizes municipalities into cities, market towns and rural municipalities. Our definition of an 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 municipality of residence

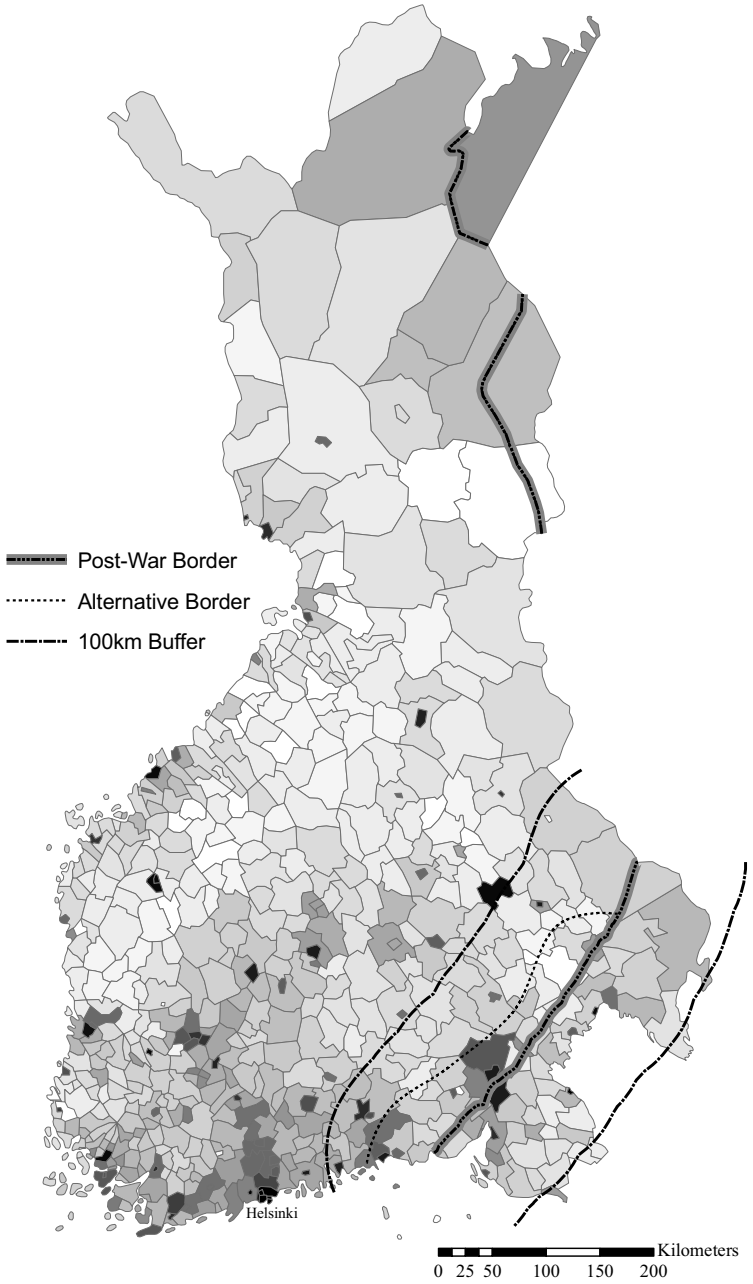
For 1939 and 1950, this variable is constructed by dividing the total taxable income (*veroäyri*) by the number of residents in the municipality.

Figure 1: Post-War Income and Pre-War Location



Note: Log taxable income in 1971 and the distance of the 1939 residence municipality to the post-war border (in kilometers). Sample: 6,495 men born between 1907 and 1925 (14–32 years old in 1939, 46–64 years old in 1971) and living in rural areas in 1939. Y-axis: 1971 log income (deviation from the non-displaced mean). X-axis: Shortest distance (in km) from the 1939 residence municipality centroid to the post-war border (negative values indicate areas remaining part of Finland). Circles correspond to mean log income by 20 kilometer bins. The size of the circle corresponds to the sample size. Lines are fitted values from the estimating $y_{i71} = \alpha + \beta_1 \{Z_i > 0\} + \beta_2 Z_i + \beta_3 Z_i \{Z_i > 0\} + \varepsilon_{ijt}$, where y_{i71} is log taxable income measured as deviation from the non-displaced's mean in 1971 and Z_i is the distance of the 1939 residence municipality to the post-war border. The 95% confidence intervals (shaded area) are based on standard errors that are clustered at the 1939 residence municipality level.

Figure 2: Ceded Areas and Taxable Income per capita, 1938



Note: Darker shades indicate higher taxable income per capita. Alternative border refers to the border the Soviet Union was preparing to offer in March 1944. Source: Rentola (2001)

Table 1: Pre-War Characteristics: Men Living in Rural Areas in 1939

	Sample means				Normalized Difference	
	All		100km Buffer			
	Disp.	Non-disp.	Disp.	Non-disp.	All	100km
Age	22.3	22.4	22.2	22.5	-0.01	-0.03
Native tongue Swedish	0.00	0.08	0.00	0.04	-0.29	-0.20
Imputed log income (def. 1)	2.25	2.26	2.25	2.28	-0.01	-0.03
Imputed log income (def. 2)	2.84	2.83	2.84	2.81	0.08	0.15
Taxable Income per capita in the Residence Mun.	28.8	27.6	28.9	33.3	0.05	-0.18
Socio-economic Status						
Entrepeneur	0.22	0.19	0.21	0.17	0.05	0.07
White Collar	0.07	0.06	0.07	0.06	0.04	0.02
Blue Collar	0.34	0.44	0.34	0.46	-0.14	-0.17
Family Member	0.22	0.18	0.23	0.16	0.07	0.13
Out of Labor Force	0.15	0.14	0.15	0.15	0.02	-0.01
Sector of Employment						
Agriculture	0.49	0.53	0.50	0.49	-0.06	0.02
Manufacturing	0.10	0.13	0.09	0.15	-0.07	-0.11
Construction	0.08	0.06	0.08	0.06	0.04	0.05
Services	0.15	0.10	0.15	0.12	0.10	0.07
Unknown	0.19	0.17	0.18	0.19	0.02	-0.02
Observations	786	7,268	708	960		

Note: Sample means of pre-war covariates. Normalized differences are calculated as $\Delta_X = \left(\overline{X}_1 - \overline{X}_0\right) / \left(\sqrt{S_1^2 + S_0^2}\right)$, where $\overline{X}_1$ and $\overline{X}_0$ (S_1^2 and S_0^2) are the sample means (variances) among the displaced and non-displaced persons.

Table 2: Pre-War Characteristics: Women Living in Rural Areas in 1939

	Sample means				Normalized Difference	
	All		100km Buffer			
	Disp.	Non-disp.	Disp.	Non-disp.	All	100km
Age	22.7	22.5	22.6	22.8	0.03	-0.02
Native tongue Swedish	0.00	0.07	0.00	0.04	-0.27	-0.19
Imputed log income (def. 1)	1.27	1.31	1.27	1.31	-0.05	-0.06
Imputed log income (def. 2)	2.90	2.90	2.90	2.91	0.01	-0.02
Taxable Income per capita in the Residence Mun.	28.9	28.4	28.9	35.4	0.02	-0.25
Socio-economic Status						
Entrepeneur	0.02	0.02	0.02	0.03	-0.01	-0.02
White Collar	0.06	0.06	0.06	0.07	-0.02	-0.04
Blue Collar	0.12	0.17	0.12	0.15	-0.11	-0.07
Family Member	0.21	0.19	0.21	0.16	0.04	0.09
Out of Labor Force	0.59	0.55	0.59	0.59	0.05	0.00
Sector of Employment						
Agriculture	0.25	0.26	0.25	0.21	-0.01	0.07
Manufacturing	0.05	0.06	0.05	0.07	-0.02	-0.05
Construction	0.00	0.00	0.00	0.00	-0.03	-0.03
Services	0.11	0.13	0.10	0.12	-0.04	-0.03
Unknown	0.59	0.56	0.60	0.61	0.05	-0.01
Observations	881	7,699	819	1,065		

Note: Sample means of pre-war covariates. Normalized differences are calculated as $\Delta_X = \left(\overline{X}_1 - \overline{X}_0\right) / \left(\sqrt{S_1^2 + S_0^2}\right)$, where $\overline{X}_1$ and $\overline{X}_0$ (S_1^2 and S_0^2) are the sample means (variances) among the displaced and non-displaced persons.

Table 3: Pre-War Characteristics: Urban Areas

	Sample means				Normalized Difference	
	Men		Women			
	Disp.	Non-disp.	Disp.	Non-disp.	Men	Women
Age	23.5	23.6	23.9	23.8	-0.02	0.02
Native tongue Swedish	0.02	0.17	0.01	0.14	-0.38	-0.36
Imputed log income (def. 1)	2.73	2.77	1.79	1.83	-0.03	-0.04
Imputed log income (def. 2)	3.04	3.01	3.22	3.16	0.05	0.10
Taxable Income per capita in the Residence Mun.	90.7	108.5	91.4	110.5	-0.43	-0.45
Socio-economic Status						
Entrepeneur	0.05	0.05	0.03	0.01	0.01	0.07
White Collar	0.20	0.18	0.20	0.20	0.05	0.00
Blue Collar	0.56	0.59	0.24	0.32	-0.04	-0.11
Family Member	0.01	0.00	0.01	0.01	0.07	0.01
Out of Labor Force	0.17	0.18	0.52	0.46	-0.02	0.08
Sector of Employment						
Agriculture	0.03	0.03	0.01	0.01	0.00	-0.03
Manufacturing	0.25	0.35	0.14	0.18	-0.15	-0.09
Construction	0.12	0.11	0.00	0.00	0.03	-0.01
Services	0.37	0.29	0.31	0.32	0.13	-0.01
Unknown	0.23	0.23	0.55	0.49	0.00	0.09
Observations	279	2,396	331	3,256		

Note: Sample means of pre-war covariates. Normalized differences are calculated as $\Delta_X = \left(\overline{X}_1 - \overline{X}_0\right) / \left(\sqrt{S_1^2 + S_0^2}\right)$, where $\overline{X}_1$ and $\overline{X}_0$ (S_1^2 and S_0^2) are the sample means (variances) among the displaced and non-displaced persons.

Table 4: Impact on 1971 log Income

	Lives in Rural Area in 1939				Lives in Urban Area in 1939	
	Men		Women			
	All	100 km	All	100 km	Men	Women
A: Baseline estimates						
No Covariates	0.193	0.166	0.022	-0.067	0.018	-0.012
	(0.042)	(0.076)	(0.046)	(0.069)	(0.059)	(0.057)
Conditional on	0.101	0.124	0.049	0.008	-0.013	-0.057
Pre-War Covariates	(0.040)	(0.057)	(0.043)	(0.065)	(0.026)	(0.044)
B: Alternative specifications						
Conditional on	0.096	0.106	0.063	0.021	0.011	-0.049
1950 Education	(0.038)	(0.057)	(0.042)	(0.061)	(0.032)	(0.042)
Excluding	0.098	0.104	0.013	-0.052	-0.043	-0.059
Agriculture in 1970	(0.044)	(0.068)	(0.045)	(0.067)	(0.027)	(0.048)
Inflating with	0.070	0.084	0.038	0.017	-0.034	-0.064
Local Prices	(0.040)	(0.060)	(0.042)	(0.063)	(0.027)	(0.043)
Conditional on	0.020	0.118	-0.012	-0.091	-0.046	-0.037
1970 Labor Market	(0.032)	(0.074)	(0.049)	(0.097)	(0.020)	(0.055)
C: Regression-discontinuity estimates						
No Covariates	0.283	0.119	-0.124	-0.221	0.273	-0.246
	(0.087)	(0.118)	(0.111)	(0.154)	(0.114)	(0.103)
Conditional on	0.198	0.109	-0.059	-0.107	0.064	-0.284
Pre-War Covariates	(0.078)	(0.118)	(0.094)	(0.135)	(0.079)	(0.094)
D: Placebo treatment						
Conditional on	-0.033	0.028	0.075	0.115	0.015	0.003
Pre-War Covariates	(0.051)	(0.064)	(0.051)	(0.067)	(0.044)	(0.093)

Note: OLS estimates for an indicator variable taking value one if the person was living in the ceded area on September 1st, 1939 and zero otherwise. Each estimate stems from a separate regression. Standard errors (in paretheses) are clustered at the 1939 municipality of residence level. Pre-War covariates: age (and its square), imputed earnings in 1939 (and its square), an indicator for speaking Swedish as ones mother tongue, six categories for socioeconomic status (entrepreneur, white-collar worker, blue-collar worker, assisting family member, out of labor force), four categories for the sector of employment (agriculture, manufacturing, construction and services), latitude of the 1939 residence municipality and five categories of taxable income per capita (quintile groups) in the 1939 residence municipality.

Table 5: Impact on 1950 outcomes (Men)

	Lives in Rural Area in 1939				Lives in Urban Area in 1939	
	Full Sample		100km Buffer			
Dependent Variable	Control Mean	Disp.	Control Mean	Disp.	Control Mean	Disp.
Employed	0.89	-0.02	0.88	0.01	0.97	0.01
		(0.01)		(0.02)		(0.00)
Entrepeneur (non-agricultural)	0.06	-0.02	0.05	-0.01	0.08	-0.02
		(0.01)		(0.01)		(0.01)
Imputed earnings	2.58	0.06	2.59	0.10	3.12	-0.04
		(0.02)		(0.03)		(0.01)
Lives in urban area	0.09	0.12	0.10	0.11	0.87	-0.21
		(0.02)		(0.02)		(0.04)
Works in agriculture	0.52	-0.10	0.49	-0.13	0.04	0.00
		(0.02)		(0.03)		(0.01)
... manufacturing	0.18	0.04	0.19	0.07	0.40	-0.02
		(0.01)		(0.02)		(0.04)
... construction	0.10	0.04	0.09	0.03	0.13	0.01
		(0.01)		(0.02)		(0.01)
... services	0.17	0.02	0.19	0.03	0.39	0.02
		(0.02)		(0.02)		(0.03)

Note: Means of the control group and OLS estimates for displacement status. Standard errors (in paretheses) are clustered at the 1939 municipality of residence level. The specification also includes age (and its square), imputed earnings in 1939 (and its square), an indicator for speaking Swedish as ones mother tongue, six categories for socioeconomic status (entrepreneur, white-collar worker, blue-collar worker, assisting family member, out of labor force), four categories for the sector of employment (agriculture, manufacturing, construction and services), latitude of the 1939 residence municipality and five categories of taxable income per capita (quintile groups) in the 1939 residence municipality.

Table 6: Impact on 1950 outcomes (Women)

	Lives in Rural Area in 1939				Lives in Urban Area in 1939	
	Full Sample		100km Buffer			
Dependent Variable	Control Mean	Disp.	Control Mean	Disp.	Control Mean	Disp.
Employed	0.63	0.05	0.68	0.05	0.95	-0.00
		(0.02)		(0.02)		(0.01)
Entrepeneur (non-agricultural)	0.06	0.02	0.06	0.02	0.08	-0.02
		(0.01)		(0.01)		(0.02)
Imputed earnings	1.71	0.12	1.78	0.11	2.38	-0.04
		(0.03)		(0.04)		(0.02)
Lives in urban area	0.10	0.12	0.12	0.10	0.83	-0.19
		(0.02)		(0.02)		(0.04)
Works in agriculture	0.50	-0.12	0.45	-0.13	0.05	0.00
		(0.02)		(0.03)		(0.01)
... manufacturing	0.20	0.02	0.24	0.02	0.38	-0.07
		(0.02)		(0.02)		(0.03)
... construction	0.06	0.01	0.06	0.02	0.08	0.02
		(0.01)		(0.01)		(0.02)
... services	0.22	0.09	0.23	0.08	0.46	0.05
		(0.02)		(0.03)		(0.04)

Note: Means of the control group and OLS estimates for displacement status. Standard errors (in paretheses) are clustered at the 1939 municipality of residence level. The specification also includes age (and its square), imputed earnings in 1939 (and its square), an indicator for speaking Swedish as ones mother tongue, six categories for socioeconomic status (entrepreneur, white-collar worker, blue-collar worker, assisting family member, out of labor force), four categories for the sector of employment (agriculture, manufacturing, construction and services), latitude of the 1939 residence municipality and five categories of taxable income per capita (quintile groups) in the 1939 residence municipality.

Table 7: Impact on 1970 outcomes (Men)

	Lives in Rural Area in 1939				Lives in Urban Area in 1939	
	Full Sample		100km Buffer			
Dependent Variable	Control Mean	Disp.	Control Mean	Disp.	Control Mean	Disp.
Employed	0.83	-0.02	0.81	-0.00	0.83	-0.03
		(0.01)		(0.02)		(0.02)
Entrepeneur (non-agricultural)	0.05	-0.01	0.04	-0.00	0.08	0.00
		(0.01)		(0.01)		(0.01)
Imputed earnings	2.61	0.06	2.63	0.08	3.00	-0.04
		(0.02)		(0.03)		(0.03)
Lives in urban area	0.24	0.16	0.30	0.07	0.87	-0.14
		(0.03)		(0.07)		(0.03)
Works in agriculture	0.28	-0.08	0.26	-0.09	0.05	0.02
		(0.01)		(0.02)		(0.01)
... manufacturing	0.15	0.02	0.14	0.04	0.27	0.04
		(0.02)		(0.02)		(0.02)
... construction	0.11	0.03	0.12	0.01	0.09	-0.01
		(0.01)		(0.02)		(0.01)
... services	0.16	0.02	0.17	0.05	0.32	-0.04
		(0.02)		(0.02)		(0.02)
Migrates between 1950 and 1970	0.35	0.10	0.41	0.05	0.34	0.14
		(0.03)		(0.05)		(0.02)

Note: Means of the control group and OLS estimates for displacement status. Standard errors (in paretheses) are clustered at the 1939 municipality of residence level. The specification also includes age (and its square), imputed earnings in 1939 (and its square), an indicator for speaking Swedish as ones mother tongue, six categories for socioeconomic status (entrepreneur, white-collar worker, blue-collar worker, assisting family member, out of labor force), four categories for the sector of employment (agriculture, manufacturing, construction and services), latitude of the 1939 residence municipality and five categories of taxable income per capita (quintile groups) in the 1939 residence municipality.

Table 8: Impact on 1970 outcomes (Women)

	Lives in Rural Area in 1939				Lives in Urban Area in 1939	
	Full Sample		100km Buffer			
Dependent Variable	Control Mean	Disp.	Control Mean	Disp.	Control Mean	Disp.
Employed	0.62	0.01	0.62	0.03	0.76	-0.06
		(0.01)		(0.02)		(0.01)
Entrepeneur (non-agricultural)	0.04	0.00	0.04	0.00	0.06	0.00
		(0.01)		(0.01)		(0.01)
Imputed earnings	1.76	0.09	1.79	0.08	2.21	-0.09
		(0.02)		(0.03)		(0.02)
Lives in urban area	0.25	0.14	0.32	0.04	0.85	-0.12
		(0.03)		(0.06)		(0.01)
Works in agriculture	0.27	-0.07	0.24	-0.07	0.05	0.02
		(0.01)		(0.02)		(0.01)
... manufacturing	0.13	0.00	0.15	0.01	0.21	-0.00
		(0.01)		(0.01)		(0.01)
... construction	0.04	0.00	0.03	0.02	0.03	-0.00
		(0.01)		(0.01)		(0.00)
... services	0.24	0.04	0.27	0.04	0.44	-0.04
		(0.02)		(0.02)		(0.02)
Migrates between 1950 and 1970	0.32	0.10	0.35	0.08	0.29	0.11
		(0.02)		(0.04)		(0.04)

Note: Means of the control group and OLS estimates for displacement status. Standard errors (in paretheses) are clustered at the 1939 municipality of residence level. The specification also includes age (and its square), imputed earnings in 1939 (and its square), an indicator for speaking Swedish as ones mother tongue, six categories for socioeconomic status (entrepreneur, white-collar worker, blue-collar worker, assisting family member, out of labor force), four categories for the sector of employment (agriculture, manufacturing, construction and services), latitude of the 1939 residence municipality and five categories of taxable income per capita (quintile groups) in the 1939 residence municipality.

Table A1: Attrition

	y = log #observations, 1970		
	(1)	(2)	(3)
Ceded municipality	0.069	0.043	0.039
	(0.127)	(0.122)	(0.122)
log Population, 1939	1.078	1.015	0.992
	(0.047)	(0.046)	(0.047)

Note: OLS estimates for the likelihood of remaining in the sample. Sample: 528 municipalities. The coefficients correspond to β (first row) and η (second row) of equation (7), see the Appendix for discussion. Specifications: (1) no additional covariates, (2) controlling for log mean age and the share of Swedish-speaking population, (3) controlling for log mean age, the share of Swedish-speaking population, log taxable income per capita in 1938 and a city dummy.