Habit Formation and the Misallocation of Labor: Evidence from Forced Migrations*

Matti Sarvimäki Roope Uusitalo Markus Jäntti

April 22, 2020

Abstract

We examine the long-term effects of resettling 11% of the Finnish population during World War II. Entire rural communities were moved to locations that resembled the origin areas and displaced farmers were given farms similar to those they had lost. Despite this policy of reconstructing the pre-war situation, we find that forced migration increased the likelihood of leaving agriculture, which in turn led to a large increase in long-term income among the displaced rural population. By contrast, being displaced decreased the income of the resettled urban population. We examine the extent to which these effects can be explained by the impact of forced migration on farm quality, education, networks, learning and discrimination, but find only limited support for the relevance of these mechanisms. Instead, we argue that a Roy model augmented with habit formation for residential location provides the most compelling rationalization for these results.

*We thank Abhijit Banerjee, Sascha Becker, David Card, Kristiina Huttunen, David McKenzie, Kaivan Munshi, Pauli Murto, Tuomas Pekkarinen, Mark Rosenzweig, Daniel Sturm, Marko Terviö, Olof Åslund and various seminar participants for helpful comments. Matti Mitrunen provided superb 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 Palkansaajasäätiö and the Jenny and Antti Wihuri Foundation. *Sarvimäki* (corresponding author): Aalto University School of Business, VATT and Helsinki Graduate School of Economics, matti.sarvimaki@aalto.fi. *Uusitalo*: University of Helsinki, VATT and Helsinki Graduate School of Economics, roope.uusitalo@helsinki.fi. *Jäntti*: Stockholm University, markus.jantti@sofi.su.se. This paper supersedes "The Long-Term Effects of Forced Migration", which first appeared as Chapter 3 of Sarvimäki's doctoral dissertation.

1 Introduction

In a typical country, a quarter of the labor force works in agriculture, where their productivity is only half of the productivity of comparable workers in other sectors (Gollin et al., 2014). At face value, this observation suggests that a reallocation of workers from agriculture to the modern sector would substantially increase aggregate productivity. Yet, the agricultural productivity gap also poses us with a puzzle: if people could earn much more in the modern sector, why do they stay in agriculture?

This paper examines transitions from agriculture to nonagriculture in the mid-20th century Finland—a poor, predominantly agrarian society at the time. Our research design is based on a large-scale population resettlement following the cession of Finland's eastern parts to the Soviet Union during World War II. In total, 11% of the population was forced to migrate and resettled into the remaining parts of Finland. For those working in agriculture—roughly one half of the population—the government attempted to reconstruct the pre-war conditions as closely as possible. Displaced farmers were given land and assistance to establish new farms in areas that had similar soil and climate as the origin regions. Former neighbors were resettled close to each other in order to preserve social networks. Once the resettlement was completed in 1948, the displaced farmers were not subject to any special policies. In particular, they received no further subsidies and, like everyone else, were free to sell and buy land and to move across locations and sectors.

We start our analysis by estimating the impact of forced migration on long-term income and mobility. The top panel of Figure 1 presents our first result using a sample of men who were working in agriculture just before the war in 1939. The horizontal axis shows the distance from their pre-war municipality of residence to the post-war border and the vertical axis presents their average annual income in 1971. The figure shows that a quarter-century after being forced to migrate, displaced farmers earned more than other men who worked in agriculture before the war.

The post-war difference between displaced and non-displaced farmers suggests that forced migration increased long-term income. This interpretation is supported by the fact that the entire population living in the ceded area was evacuated and resettled in an orderly manner. Thus, the post-war differences do not arise from self-selection into migration or survival bias. Furthermore, there are little differences in the pre-war observable characteristics of the displaced and non-displaced farmers. Combining estimates from alternative approaches to get plausible bounds, we find that being displaced increased long-term income by 16–30% among men working in agriculture before the war.

We next examine potential channels behind the positive effect on income. The bottom panel of Figure 1 shows that displaced farmers were more likely to move from agriculture to other sectors between 1939 and 1970. Estimates from alternative specifications suggest that forced migration

increased the likelihood of leaving agriculture by 12–17 percentage points from a baseline of 28%. Importantly, this effect reflects voluntary transitions, because the displaced farmers were given new farms in the resettlement areas.

We also find that forced migration increased the likelihood of moving to a city and to complete secondary education among the displaced farmers, and that the impacts on income and mobility closely mirror each other when we extend the analysis to other groups. Specifically, being displaced decreased income and increased the likelihood of moving to rural locations among the urban population. In addition, the average income of displaced persons was similar in the 1970s to that of non-displaced persons working in the same industries and living in the same locations after the war. Taken together, our results suggest that the positive impact of forced migration on the income of farmers can be attributed to an increased likelihood of leaving agriculture, and that the returns to leaving agriculture were large in the mid-20th century Finland.

If farmers could substantially increase their income by moving to the modern sector, why did most of them decide to stay on their farms? Much of the previous work suggesting answers to this question has focused on the riskiness of urban labor markets (Harris and Todaro, 1970; Bryan et al., 2014), local prices and amenities (Rosen, 1979; Roback, 1982), and sectoral differences in human capital and returns to skills (Caselli and Coleman, 2001; Lucas, 2004; Lagakos and Waugh, 2013; Young, 2013). However, these models are unlikely to explain our results, because the displaced and non-displaced persons did not significantly differ from each other along the dimensions they examine.

We use two complementary approaches to shed light on the likely mechanisms at play. First, we interpret our results through the lens of a Roy model with heterogeneous comparative advantages, farm qualities and migration costs. This model provides structure for our discussion by highlighting that post-war mobility and income could be driven both by changes in returns to migration as well as by changes in migration costs. It also provides some additional predictions that may help in distinguishing between alternative mechanisms. Second, we draw from earlier work based on interviews and surveys. A particularly powerful source is an early study on the adaptation of the displaced population based on two large surveys and in-depth interviews conducted in 1949 and 1951 (Waris et al., 1952). This work provides insights into how the displaced and non-displaced persons thought about the resettlement policy and its aftermath a few years after its completion. Based on their findings, we also suggest a variant of the Roy model that we believe to provide the most compelling rationalization for our results.

We start examining the potential mechanisms by discussing the extent to which the displacement affected returns to migration through a reduction of income available from agriculture. This channel was likely to be particularly severe for the owners of large farms, because the new farms had at most 15 hectares of agricultural land. Importantly, however, less than a tenth of the displaced

farmers owned farms larger than 15 hectares. Furthermore, we find no evidence on the pre-war farm size explaining post-war mobility and income. On the other hand, using pre-war regional yield data, we estimate that the resettlement area had roughly 2% lower average yields than the ceded area. We also discuss the possibility that the displaced farmers could have received below-average quality land in the resettlement areas. Taken together, the available evidence suggests that while some farmers may have experienced significant deterioration in their farm quality—and were thus pushed out from agriculture—the average change is likely to have been modest. Furthermore, it is hard to rationalize the finding that forced migration increased the long-term income of displaced farmers through a decline in farm quality alone. Hence, we argue that this channel is unlikely to explain the full pattern of our results.

Next, we examine whether our results can be explained by the forced migration affecting human capital investments (Becker et al., 2019), networks (Banerjee and Newman, 1998; Munshi and Rosenzweig, 2016) or discrimination. Our data and the details of the resettlement policy allow us to conduct empirical tests for each of these channels, but we fail to find evidence supporting them. Furthermore, our reading of the qualitative and survey-based evidence is that while some of these channels may have played a role, they are likely to leave an important part of the story untold.

We end by considering a hypothesis that an important part of migration costs may arise from people growing attached to a place. This hypothesis is inspired by qualitative and survey-based evidence, where displaced persons typically describe having lost their *homes* rather than just jobs and productive assets. For example, Waris et al. (1952) found that the displaced persons tended to express a strong desire to return to their former homes. Importantly, their revealed preferences are in line with their survey responses. The first evacuation took place during the Winter War (1939–40) and the first version of the resettlement policy was executed during what later became known as the Interim Peace (1940–41). In 1941, Finland joined Germany's attack on the Soviet Union and reoccupied the ceded areas. Despite much destruction in the reoccupied areas and the genuine opportunity to remain on their new farms, the vast majority of the displaced farmers returned. This was a costly and risky decision, given that their old farms had in many cases been destroyed and that the outcome of the war was anything but certain. Indeed, their investments in repairing their old farms were lost in 1944, when the same areas were again ceded to the Soviet Union and the return migrants were evacuated and resettled for the second, and final, time.

We rationalize the survey responses and return migration behavior by augmenting the Roy model with habit formation for residential location. The key ingredient of this model is the assumption that people derive utility both from income and from their residential location—and that utility from a location increases with the time the person has already lived there. We call the latter property habit formation in the spirit of Pollak (1970) and follow Becker and Murphy (1988) by modeling it as an accumulation of "location capital" that directly affects contemporaneous utility.

This accumulation process starts already in childhood and hence affects the choice of location in adulthood. Specifically, a person who has grown up on a farm may choose to remain in agriculture in order to enjoy her location capital even if she could earn more somewhere else. However, if she is forced to move, she loses the location capital tied to her old home and chooses the location providing her with the highest income after the displacement. While we do not claim that other mechanisms are irrelevant in explaining our empirical results, we argue that an explanation including habit formation is substantially more compelling than one based purely on other mechanisms. We also discuss how this formulation of the Roy model yields more nuanced welfare implications than a model without habit formation.

This paper is part of the recent literature evaluating the long-term effects of forced migrations. The work closest to us has examined population displacements created by territorial changes of Germany and Poland after WWII (Bauer et al., 2013; Becker et al., 2019), the internment of Japanese Americans during WWII (Arellano-Bover, 2018), a volcanic eruption in Iceland in 1973 (Nakamura et al., 2019), the demolition of public housing in Chicago in the 1990s (Chyn, 2018), and Hurricane Katrina in 2005 (Deryugina et al., 2018).[1] Despite vast differences in the contexts these papers study, all find positive effects on long-term income among agricultural workers and/or individuals who were relatively young or not yet born at the time of displacement.

In comparison to the other forced migration episodes examined thus far, the Finnish experience is unique in combining three features. First, the resettlement policy was designed to keep rural communities together and to give displaced farmers farms that were comparable to the ones they had lost. Thus, we study decisions to voluntarily leave agriculture in a situation where the pre- and post-migration circumstances are largely comparable, apart from the loss of the original farm. Second, we are able to examine alternative mechanisms behind the overall effect using variation created by the details of the evacuation and resettlement policies and high-quality contemporary survey-based research. The displaced farmers were also given the opportunity to reveal their location preferences during the period when Finland temporarily reconquered the ceded areas. Third, we conduct our analysis using longitudinal data that follows a large number of individuals over several decades and is unlikely to suffer from non-random attrition or recall bias. Together, these aspects give rise to plausible identification of the impact of forced migration and allow us to paint a more nuanced picture of the underlying mechanisms than have been feasible in other contexts.

More broadly, our findings add to the large literature examining the possibility that misallo-

[1] The broader literature on the impacts of forced migration is reviewed in Ruiz and Vargas-Silva (2013) and Becker and Ferrara (2019). Other quantitative work examining the post-WWII population displacement in Finland include Waris et al. (1952) (which we discuss in detail in Section 5.4); Saarela and Finnäs (2009) and Haukka et al. (2017), who focus on mortality; Sarvimäki (2011), who examines the impact on the industrial structure of the receiving areas; and Lynch et al. (2019), who examine the associations between intermarriage, fertily and socioeconomic outcomes within the displaced population.

cation of labor across sectors and locations constitutes a major obstacle to development. This hypothesis goes back to at least Lewis (1955) and remains an active area of research.[2] Our results, together with the work cited above, suggest that migration costs are an important factor affecting the allocation of labor. Large migration costs, particularly for leaving one's place of birth, also show up in quantitative structural models of migration (e.g. Kennan and Walker, 2011; Diamond, 2016; Bryan and Morten, 2018; Lagakos et al., 2018). Thus, policies reducing these costs could have large effects. However, the effectiveness and welfare implications of alternative policies crucially depend on the reasons behind migration costs.

We contribute to this broader literature by discussing the potential importance of habit formation as an impediment to mobility. Earlier work closest to us in this regard has examined habit formation towards locally abundant food (Atkin, 2013, 2016). This form of habit formation is unlikely to be important in our context because displaced farmers were resettled into locations that had similar soil as their origin areas. More importantly, while alternative sources of habit formation provide partly similar insights, there are also significant differences. The key similarity is that people may be stopped from pursuing their comparative advantage if they have accumulated location or consumption capital before choosing their sector of employment. Unlike food or other consumption goods, however, living in a certain location is fundamentally nontradable. Hence, attachment to a place may be less affected by technological and institutional changes than migration costs arising from other forms of habit formation.[3]

We proceed as following. The next two sections introduce the historical episode we study and our data. We report our main results in Section 4 and discuss possible interpretations in Section 5. The final section concludes.

2 The Resettlement

2.1 Historical Context

At the beginning of World War II, Finland was a poor 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. In 1938, Finland's GDP per capita was roughly $4,000 (in 2011 USD, see Bolt et al., 2018) and more than half of the population was working in agriculture,

[2] In addition to the papers cited above, examples include Gollin et al. (2002), Caselli (2005), Munshi and Wilson (2011), Adamopoulos and Restuccia (2014), and Fernando (2016). Hopenhayn (2014) and Restuccia and Rogerson (2017) review the broader literature on misallocation.

[3] For example, Atkin (2013) examines the welfare implications of a reduction in trade costs in the presence of habit formation for food varieties. In this context, regional price differences create migration costs that vanish if prices converge. By contrast, trade costs do not affect migration costs arising from habit formation towards a specific location.

typically owning small farms and working as hired labor in forest work during the winter. Finland modernized and grew rapidly after World War II. In 1970, GDP per capita was about $14,000 and less than one-fifth of the population worked in the primary sector.

The Soviet Union attacked Finland in November 1939 after negotiations they had initiated on moving the Finnish-Soviet border collapsed. The civilian population living in the conflict areas was evacuated and transported to designated evacuation areas in the middle and western parts of the country, where the local population was obliged to provide them with shelter. In the peace treaty ending the hostilities in March 1940, Finland ceded roughly a tenth of its territory to the Soviet Union. Part of this area had remained under Finnish control during the war and the civilian population living in these areas were evacuated as part of the peace treaty.

In July 1940, the Finnish Parliament enacted an Emergency Settlement Act (*Pika-asutuslaki*) guiding the resettlement policy. However, the 1940 resettlement policy turned out to have limited long-term effects, because Finland joined Germany in its attack on the Soviet Union in June 1941 and reoccupied almost all of the ceded areas. As we discuss in more detail in Section 5.4, 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 almost three years of trench warfare, the Soviet Union launched a massive attack in June 1944. The armistice signed in September, and later ratified in the Paris Peace Treaty, restored the 1940 border with some additional areas ceded to the Soviet Union. The entire population living in the ceded area was again evacuated and resettled. The border has been unchanged and undisputed ever since.

Figure 2 shows the pre-war and the post-war borders and the 1945 resettlement plan discussed in detail below. It seems reasonable to consider the 1944 border as good as randomly assigned from the viewpoint of the population living in Eastern Finland in 1939. 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 since 1809, belonging first to the Russian Empire as part of the autonomous Grand Duchy of Finland and, from 1917 onwards, to independent Finland.

In the peace negotiations between Finland and the Soviet Union, historical borders were used as reference point. 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.[4] The post-WWII border closely follows the border set in the treaty of Nystad in 1721. Rentola (2001) discusses archive material indicating that when the Soviet Union offered peace talks in March 1944, it was preparing to negotiate based on the 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, together with the need to reallocate Soviet troops to the Baltic front, had

[4]See e.g. www.tacitus.nu/historical-atlas/scandinavia/finland.htm.

improved Finland's position in the negotiations and thus moderated the Soviet demands. Below we will also show that average pre-war characteristics were similar before the war on both sides of the post-war border.

2.2 The Resettlement Policy

Resettling the 430,000 displaced persons was a major challenge. 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 industrial production capacity was destroyed in the war and further cuts in capacity were caused by the war reparations that amounted to roughly a sixth of the government budget between 1945 and 1949 (Mitrunen, 2019).

Despite the grave economic situation, the Parliament approved a series of laws in 1940 and 1945 that offered compensation for the property lost due to the displacement. The rate of compensation varied from full reimbursement for small losses to compensation of only ten percent for very large ones. Those who had owned or leased agricultural land in the ceded areas were given agricultural and forest land (Pihkala, 1952). Those who had lost other kinds of property received their compensation primarily in the form of inflation-indexed government bonds for which a liquid secondary market quickly emerged.

The resettlement was financed by levying a massive tax on wealth. Land for the settlers was first taken from the state, the local authorities (municipalities) and the church. However, roughly two-thirds of the cultivated fields, one half of the land that could be cleared for cultivation and a third of forest land were seized from private owners using an explicit progressive expropriation schedule.[5]

The aim of the resettlement policy was to match the pre-war conditions as closely as possible. In order to preserve social connections, farmers from each ceded village were settled together to a designated target area. Furthermore, the soil quality and average temperatures of the source and destination areas were matched as closely as possible. As illustrated by Figure 2, 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 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. While

[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 four 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 forms of wealth. Pihkala (1952) discusses the land acquisition policy in detail and argues that landowners did not suffer more than other property owners.

those who had not worked in agriculture were not explicitly allocated, the settlement plan appears to have influenced also their migration, probably due to family ties and employment opportunities with former employers. In June 1949, 53 percent of the displaced persons lived in their designated placement areas (Waris et al., 1952).

The resettlement was completed in 1948, after which no further policies targeted to the displaced population were introduced and the displaced and non-displaced population had equal legal status. In particular, everyone could sell and buy land and migrate anywhere in the country.

3 Data

Statistics Finland constructed our data by linking a 10% sample of the 1950 population census to the 1970 census and the 1971 tax records. The information for pre-war municipality of residence, occupational status and industry codes comes from the 1950 census, which included retrospective questions referring to September 1st, 1939—two months before the war began. We augment these individual-level data with municipality-level information on the pre-war income distribution and industry structure. We discuss the details of the data and variable definitions in the Appendix.

Table 1 reports the average pre-war characteristics included in our data for individuals born between 1907 and 1925. We focus on these 78,549 individuals—of whom 7,805 were displaced—because they remain of working-age throughout the period we study.[6] Overall, the displaced and non-displaced populations have quite similar pre-war characteristics. The largest differences are in the share of people speaking Swedish as their mother tongue (a relatively prosperous group heavily concentrated on the southern and western coasts of Finland) and in the share of population who were members of the Orthodox church (a less prosperous group concentrated in the eastern parts of the country). Furthermore, the displaced rural population was less likely to work as blue-collar workers and in manufacturing, and tended to live in somewhat poorer municipalities in 1939.

The earliest information on individual-level income comes from the 1971 tax register. These data provide an accurate measure of annual earned income. Tax records are likely to provide a comparable measures of income across agricultural and non-agricultural households, because The Finnish tax authorities treated agricultural income similarly to wages and the extent of home production was modest in the 1970s. Indeed, taxable earned income predicts consumption in a very similar way for farmers and non-farmers in the 1971 Household Budget Survey (Appendix Figure A1 and Appendix Table A1).

Other outcome variables come from the 1950 and 1970 censuses. We use industry codes to construct an indicator for working outside of agriculture, and municipality codes and Statistics Finland's pre-war definition of cities for an indicator for living in an urban area. We also con-

[6] They were 14–32 years old when the war started in 1939 and 46–64 years old in 1971.

struct an imputed income measure for 1950 using mean taxable income in 38 industry-occupation-socioeconomic status groups for 1950 as reported in Statistics Finland (1953, Table 2). Our main measure for education is an indicator for holding at least a secondary degree in 1970.

Our data also include information on the education and income of the children of individuals present in our main sample. We focus on children born after 1948, i.e. those who did not experience the evacuation or the implementation of the resettlement policy themselves. A limitation of these data is that we observe only one parent for 42% of the children. As we discuss in more detail in Section 4.2 and in the Appendix, this gives rise to somewhat complex measurement error in parents' displacement status.

4 Impact of Forced Migration

This section reports our main results. We start with a discussion of our empirical strategies and then report the estimates on the impact of forced migration. We also report estimates for income conditional on industry and location, and discuss what our results imply on returns to leaving agriculture during this period.

4.1 Empirical Strategies

We evaluate the impact of forced migration by comparing the outcomes of displaced persons to control groups of persons who were not displaced. As we discuss below, each of these comparisons may yield biased estimates. However, alternative approaches are likely to suffer from biases of opposite signs and thus provide plausible bounds for the impact of being forced to migrate.

In practice, we estimate variants of the regression equation

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

where y_{it} is the outcome of interest for individual i at time t, D_i is an indicator for the person living in the ceded area just before the war, X_{0i} is a vector of observed pre-war characteristics, and ε_{it} captures unobserved factors. We implement the various comparisons by estimating (1) for different subsamples and by varying the content of X_{0i}.

We recognize that the resettlement was likely to affect the entire population of post-war Finland. Hence, our aim to is estimate a causal relationship in the sense of a thought experiment in which one would manipulate the displacement status of a single individual, while 11% of the population were still forced to migrate.

Baseline Estimates and Oster Bounds Our baseline estimates come from comparisons between all displaced and non-displaced individuals. A limitation of this approach is that the displaced and non-displaced populations differ somewhat in their pre-war characteristics (Table 1) and may thus also differ in their unobservable characteristics. However, the observed differences are relatively small and our data allow us to condition on a rich set of pre-war observables. As a baseline, we thus report estimates with and without controlling for pre-war differences. We then use the difference between the unconditional and conditional estimates to bound the likely remaining omitted-variables bias (Altonji et al., 2005; Oster, 2019). Specifically, we report bounds under the assumption that selection-on-unobservables is as important as selection-on-observables and that the hypothetical maximum R^2 from a regression including all relevant background characteristics is $1.3\tilde{R}$, where $\tilde{R}$ is the R^2 from the regression including the control variables observed in our data; see Oster (2019) for discussion.

Spatial Regression Discontinuity Design Our second comparison is between individuals, who lived just east of the post-war border (and were thus displaced) and people who lived slightly more to the west (and were thus not displaced).[7] This spatial regression-discontinuity design builds on the plausibility of locally random assignment into forced migration (see Section 2). However, its limitation is that those living in the control areas may have been affected by the shift of the border more than those living further away. For example, 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 the Finnish municipalities close to the new border suffered from similar adverse effects, the spatial RD estimates would be biased upwards. Thus, we interpret these estimates as upper bounds on the treatment effect.

Within-Resettlement-Area Comparisons Our third comparison is between displaced persons and the local population of their resettlement areas.[8] The main advantage of these within resettlement area comparisons is that the destination areas were far away from the post-war border, but were designed to match the origin areas by soil quality and average temperature. That is, the reset-

[7] We implement the spatial RD comparisons using standard local linear estimators. That is, we add pre-war distance to the post-war border and its interaction with the displacement status to X_{0i}, restrict the estimation sample to persons who lived close to the post-war border before the war (using the Imbens and Kalyanaraman (2012) algorithm to choose the optimal bandwidth) and weight the observations close to the border more than those further away using a triangle-shaped kernel.

[8] We implement this comparison by including resettlement area fixed-effects in X_{0i} and dropping the non-displaced persons living outside of the resettlement area from the sample. These fixed-effects are constructed using the 1939 residence municipality information and, for the displaced persons, refer to the areas where the displaced persons would have been living in after the war if they had followed their resettlement plan (regardless of where they actually lived after the war). The displaced persons were not able to choose their resettlement areas and thus these regressions do not suffer from the "bad control" problem.

tlement areas were designed to provide as similar an environment as possible to what the displaced farmers would have had if they had not been forced to migrate. However, the caveat is that the destination areas tended to be slightly richer and more industrialized before the war (Appendix Tables A2–A5). Furthermore, the resettlement shock itself may have pushed rural municipalities to industrialize faster and thus increased local wages (Sarvimäki, 2011). Thus, we interpret estimates from these comparisons as lower bounds of the treatment effect.

4.2 Long-Term Effects

Tables 2–5 present estimates for the differences between displaced and non-displaced persons or their children. Each entry comes from a separate regression that differ in the population examined (rows) and specification (columns). In order to assess the magnitudes of the estimates, we also report the mean outcomes among the non-displaced persons. We cluster standard errors at the level of the 1939 residence municipality.

Income The first row of Table 2 reports results for men working in agriculture before the war. In 1971, displaced farmers earned €2,080 more annually than non-displaced farmers. In comparison to the €10,500 average earnings among non-displaced farmers, this difference corresponds to 20% higher income. Controlling for the observable pre-war characteristics reduces the point estimate very slightly to €2,060. Assuming that selection-on-unobservables is as important as selection-on-observables, the difference in the point estimates—together with an increase in the R^2 from 0.005 to 0.123—suggests a lower bound of €2,050 or 19%. The spatial RD estimates show that farmers' income jump by €3,120 or 30% in comparison to comparable non-displaced farmers at the post-war border. On the other hand, displaced farmers had about €1,670 or 16% higher long-term income than non-displaced farmers with similar pre-war characteristics living in the resettlement areas already before the war. As we discussed in the previous subsection, we interpret the within-resettlement-area comparisons as lower bounds and the spatial RD estimates as upper bounds. Thus, we conclude that forced migration increased the long-term income of the displaced male farmers by 16–30%.

The remainder of Table 2 shows the same estimates for other groups. The baseline estimates and the Oster bound for men living in rural areas but working outside of agriculture before the war correspond to a 7–11% increase in income. For this group the spatial RD and within resettlement area comparisons yield smaller and statistically insignificant estimates. We also find strong positive effects for rural women. The baseline estimates and Oster bounds correspond to 30–37% higher income among women working in agriculture before the war and 23–34% higher income for rural women working outside of agriculture before the war. The spatial RD and within-resettlement-area

estimates for rural women are comparable to the baseline estimates. By comparison, the impact of forced migration on the urban population is very different. While the unconditional differences are not statistically significant, estimates from regressions controlling for pre-war characteristics are significant and suggest 19% (men) or 15% (women) decreases in long-term income.[9]

Appendix Table A6 reports similar estimates for another measure of income, where we have taken into account local prices. The results are very similar, though slightly smaller than those reported in Table 2. Furthermore, Appendix Table A7 shows that the effects are larger for individuals who were displaced at a younger age. We next show that these patterns are closely mirrored also for industry, urban status and education.

Industry, urbanization and education Table 3 reports estimates for the key outcomes measured in the 1970 census. For brevity, we report only baseline and within-resettlement-area estimates conditioning on pre-war characteristics. Appendix Figure A2 presents the results from all specifications used for income above.

The results for the likelihood of working outside of agriculture are similar to those for income. In 1970, displaced male farmers were 15 percentage points or 53% more likely to work in a non-agricultural job—most prominently in manufacturing and construction (Appendix Table A8)—than comparable non-displaced farmers. The estimates for rural men who worked in non-agriculture before the war are much smaller and statistically insignificant. For rural women, we find strong positive effects for holding a non-agricultural job in 1970. Similar to the results for income, there is a negative effect on working outside of agriculture for the urban population. Furthermore, the effects are again larger for those who were displaced at a younger age (Appendix Table A7).

For farmers, increased working outside of agriculture is matched by a comparable, or larger, decrease in the likelihood of working in agriculture. Hence, the point estimates for being employed in 1970 are negative for both male and female farmers, although only the within-resettlement-area estimate for men is statistically significant. We find no employment effect for rural men working outside of agriculture already before the war, while the results suggest that being displaced increased long-term employment of non-agricultural rural women by four percentage points. Again, the estimates suggest that forced migration had a negative impact on the urban population.

The remainder of Table 3 shows similar estimates for the likelihood of living in a city and for education. The displaced rural population was substantially more likely to move to urban areas than the comparable non-displaced population, while the opposite is true for the urban population. Finally, the displaced rural population were more likely to hold a secondary degree in 1970 than

[9] We do not report spatial RD due to very few cities being located close to the post-war border. Furthermore, as the urban population was not resettled into certain locations, the within-resettlement-area research design is not suitable for them.

the non-displaced rural population, while we find no statistically significant differences for those living in urban areas before the war.

Intergenerational effects Table 4 examines even longer-term effects by reporting estimates for the children of the individuals included in our main sample. We use a structure similar to that in Table 3, but the dependent variable is now children's average income when they are 30–40 years old (columns 1–3) or an indicator for having completed a secondary degree by 2011 (columns 4–6). The treatment status is based on the father (panel A) or mother (panel B), regardless of the status of the other parent.

The estimates are qualitatively similar to those for the first generation. Among the non-displaced population, children of farmers have lower income and educational attainment than children of urban parents, while the children of rural parents working outside of agriculture before the war fall in between. The same pattern is present also for education. More importantly, children of displaced farmers have higher income than children of comparable non-displaced farmers, while the opposite is true for children of urban parents. Furthermore, the estimates suggest that forced migration increased educational attainment of the children of rural parents, while we find no impact for the urban population.

In terms of magnitudes, the intergenerational effects are substantially smaller than those for the first generation. This finding is consistent with the impact of forced migration fading away over generations. However, it could also follow from measurement error, because some of the non-displaced persons in our data are likely to have a displaced spouse who we do not observe (see Section 3). A full evaluation of the intergenerational effects would also benefit from an investigation of this population displacement on the marriage market. We leave these analyses for later work in the hope that more comprehensive data will become available in the future.

4.3 Medium-Term Effects

A limitation of the long-term effects discussed above is that they could lead to misleading conclusions about lifetime outcomes. For example, Lucas (1997) proposes a rationalization for rural-urban wage gaps based on the assumption that leaving agriculture reduces short-term income, but leads to faster human capital accumulation. As a consequence, incomes of migrants eventually overtake incomes of farmers. However, examining only the long-term outcomes would miss the initial investment phase and thus lead to an overstatement of the impact of forced migration on lifetime income.

A challenge for examining income dynamics in our context is that we observe individual-level income only from 1971 onwards. However, the 1950 census includes information on occupation,

industry and socio-economic status that we can use to construct a measure of imputed income (see Section 3). Using this measure as an outcome variable, the first columns of Table 5 shows that displaced male farmers had jobs associated with 15% higher income than comparable non-displaced farmers already in 1950. Furthermore, the estimates for other groups are also quite similar to our estimates for long-term income. While we cannot rule out the possibility that the displaced farmers could have had below-average earnings in these jobs, these results suggest that the displacement had a positive (negative) effect on the income of the rural (urban) population already by 1950.

The rest of Table 5 presents similar analysis for sector, employment, urban status and education in 1950. Again, the results are very similar to those for 1970 outcomes. In fact, the impact of working outside of agriculture is larger in 1950 than in 1970. This pattern arises from the non-displaced rural population partially catching up with the displaced persons over time. For example, 20% of non-displaced male farmers had moved from agriculture to the modern sector by 1950, while the number for displaced farmers was 40%. By 1970, the share had increased to 28% among non-displaced farmers, while it was 43% among displaced farmers. Thus forced migration appears to have both increased the share of population moving to the modern sector as well as pushed them to make the transition earlier than the non-displaced population. On the other hand, we do not find a positive effect on education by 1950, suggesting that the investments in human capital due to being displaced primarily took place sometime between 1950 and 1970.

4.4 Conditioning on Post-War Sector and Location

The results discussed thus far are consistent with the hypothesis that increased sectoral mobility, typically accompanied by urbanization and investments in education, led to higher earnings among the displaced farmers. Table 6 presents complementary evidence supporting this hypothesis by comparing the annual income of displaced to non-displaced persons who worked in the same industries and lived in the same locations in 1970. For reference, columns 1 and 5 report estimates controlling only for pre-war characteristics. We then gradually condition on working outside of agriculture (columns 2 and 6), education in 1970 (columns 3 and 7), and fixed effects for 1970 residence municipality and 2-digit industry (columns 4 and 8). Among the rural population and urban women, the point estimates for displacement status approach zero and become statistically insignificant as we add further post-war control variables. However, the estimates for urban men remain negative and significant in all specifications.

We emphasize that the estimates reported in Table 6 do not have a causal interpretation, because we are now conditioning on post-war outcomes that were themselves affected by forced migration. Nevertheless, the estimates can be interpreted as informative descriptive statistics showing that

the long-term income of displaced rural persons did not differ from the income of non-displaced persons who worked in the same industries and locations after the war.

4.5 Returns to Leaving Agriculture

Taken together, our results suggest that leaving agriculture had high returns in the mid-20th century Finland. Table 7 attempts to further quantify these returns using data on pre-war farmers. For reference, column 1 shows that men who still worked in agriculture earned roughly €10,000 in 1970, while women who had remained in agriculture earned only €700. The OLS estimates show that men who had left agriculture by 1970 had about €6,000 or 57% higher annual earnings than observationally identical farmers, who had remained in agriculture. For women, the earnings difference is between €8,200, corresponding to more than a tenfold increase in income. These estimates would measure the true returns to leaving agriculture if selection into the modern sector was as good as random (once we condition on observable characteristics). This identifying assumption seems unlikely to hold, because people are likely to self-select across sectors based on their unobservable characteristics.

In order to complement the OLS estimates, we report 2SLS estimates, where we use displacement status as an instrument for working outside of agriculture in 1970. These estimates would measure the returns to leaving agriculture if the impact of forced migration on long-term income was mediated *entirely* through the transition to the modern sector. Clearly, other possible mechanisms exist. For example, as we discuss in more detail in the next Section, being displaced may have affected human capital investments (Becker et al., 2019) or economically valuable social networks. Thus, we emphasize that the IV approach is based on stronger, and less plausible, identifying assumptions than the results on the overall impact of forced migration. Nevertheless, they provide a potentially informative summary of the impacts of forced migration on income and sectoral mobility.

The IV estimates paint a very similar picture as the OLS estimates. Since IV approaches are informative only about the subpopulation of "compliers" (see the next section), we first report estimates of what the compliers' would have earned if they had stayed in agriculture. In comparison to this baseline, the 2SLS estimates suggest that leaving agriculture increased the income of men by 84%. Again, the estimates for women are similar in levels, but much larger in comparison to their baseline income in agriculture.

5 Interpretation

The results reported thus far suggest that returns to leaving agriculture were substantial in the mid-20th century Finland. This leaves us with the question of why most farmers chose to forgo these opportunities and why forced migration pushed many of them into the modern sector. In this section, we address this question through the lens of a simple Roy model and examine which of its variants are the most consistent with our data. We also suggest an extension incorporating habit formation for residential location, which we believe to provide the most compelling rationalization for our results. Throughout, we contrast our results to those from an early study by Waris et al. (1952), who collected survey data and conducted in-depth interviews among the displaced and non-displaced population.[10]

5.1 An Illustrative Roy Model

In order to organize thoughts, we consider a simple Roy model with heterogeneous comparative advantage and moving costs. Our aim is to present the simplest possible framework for structuring discussion and thus we keep the model as bare-bones as possible. More elaborate models starting from similar building blocks include Lagakos and Waugh (2013), Young (2013), Bryan and Morten (2018), Lagakos et al. (2018) and Nakamura et al. (2019). We discuss our own extention in Section 5.4.

Consider an economy consisting of two sectors, agriculture denoted by a and non-agriculture denoted by n. Individuals, denoted by i, inelastically supply one unit of labor and maximize utility by choosing their sector of employment. They differ in their migration cost $C(i)$, and in their industry-specific efficiency units of labor $z_a(i)$ and $z_n(i)$. An individual working in agriculture receives income $A(i) z_a(i)$, where $A(i)$ summarizes the quality of his farm. Those working in non-agriculture receive income $z_n(i)$. That is, we assume that farm quality and farmer's productivity are complements, and that non-agricultural labor markets are competitive. Furthermore, we normalize non-agricultural wages per efficiency unit to one. Given these assumptions, a person starting in in

[10]The research project "The Adaptation of Displaced People: A Study on the Social Adaptation of Finnish Karelian Displaced People" was led by Heikki Waris, an eminent professor of social policy at the University of Helsinki. It was launched in 1948 with funding from the Rockefeller Foundation. The research group conducted two surveys in 1949 and 1951. The final survey data include 1,982 displaced and 1,150 non-displaced persons living around in the resettlement areas (see Appendix Figure A4 for the geographical distribution of the survey). The sample was constructed by first stratifying municipalities into groups based on the population shares of displaced persons and then using quota-sampling to ensure representativeness in terms of gender and age within each location. In addition to the baseline survey, the research group conducted in-depth interviews in two rural municipalities and in one industrial town in 1949. The results were published (in Finnish) in Waris et al. (1952).

agriculture will move to non-agriculture if

$$\underbrace{z_n(i) - A(i) z_a(i)}_{\text{Returns to leaving agriculture}} > \underbrace{C(i)}_{\text{Cost of migration}} \tag{2}$$

While simple, this model illustrates the potential complexity in who selects into making the transition to the modern sector. Those with stronger comparative advantage in non-agriculture (larger $z_n(i)/z_a(i)$) or lower quality farms, are more likely to leave agriculture. However, given sufficiently large migration costs, they are willing to forgo large monetary returns to migration. The resulting selection pattern depends on the joint distribution of comparative advantages, farm qualities and migration costs. All of these factors are typically unobserved and thus this model has little direct empirical content (Heckman and Honore, 1990).

For our analysis, the value of organizing thoughts through equation (2) is twofold. First, it provides structure for our discussion by listing factors through which the displacement may affect transitions from agriculture to the modern sector. Specifically, it highlights that increased mobility following the displacement is consistent with both changes in returns to leaving agriculture and changes in the cost of migration. Second, as we discuss below, it provides us with some additional predictions that may allow us to distinguish between alternative mechanisms.

5.2 Returns to Leaving Agriculture

We start by asking whether we can rationalize our empirical findings solely by the displacement increased returns to leaving agriculture, while having no impact on migration costs. Such effect would arise if the resettlement either reduced income available from agriculture or increased earnings potential in the modern sector. Here, we discuss two potential mechanisms that could lead to such effects: a reduction in the quality of farms and a direct impact of the displacement on education.

Quality of the New Farms The most direct way the resettlement may have affected returns to leaving agriculture is through quantity and/or quality of agricultural land.[11] In terms of quantity, the changes were mechanical as the size of the new farms was limited to 6–15 hectares of cultivable land. Hence, those who had derived their primary income from a farm smaller than 6 hectares of land were given more land than the one they had lost. Those who had owned more than 15 hectares experienced a reduction in their farm size.

[11] However, it is unlikely that the type of imperfect skill-transferablity examined by Bazzi et al. (2016) in the context of Indonesian resettlement program would be relevant in our case, because the resettlement policy we examine was designed to allocate farmers to areas with similar soil quality and average temperature.

We use municipality-level information from the 1930 Agricultural Census to assess the likely importance of this compression of the farm size distribution. These data show that among the farmers who were likely to get a new farm as part of the resettlement policy, less than a tenth had more than 15 hectares, while a third had 3–5 hectares of agricultural land in 1930 (Appendix Table A9). The rest fell in between and thus should have received a new farm of equal size as the one left in the ceded area. Hence, the resettlement seems to have affected primarily the distribution of farm size rather than the average size of the farms. Furthermore, we find no evidence that displaced farmers coming from municipalities that had had more large farms were more likely to leave agriculture than those from municipalities with smaller farms (Appendix Table A10). However, as we discuss in more detail in the Appendix, this analysis yields quite imprecise estimates. While the results strongly suggest that changes in farm size are unlikely to be the main mechanism behind our results, we are not able to rule out economically meaningful variation in the impact of displacement across this dimension.

Of course, the resettlement may have also affected opportunities in agriculture through land quality. This could occur in two ways. First, the average land quality in the resettlement areas may have been lower than in the ceded areas. We investigate this possibility using regional-level information on yields per hectare of various agricultural crops as reported in the 1930 Agricultural Census. These data suggest that the ceded areas had around 2% higher yields than the resettlement area (Appendix Table A11). Second, displaced farmers may have been given below-average quality land within the resettlement areas. Importantly, however, the resettlement was implemented through a highly regulated process, where the displaced farmers had strong representation and thus local land owners faced severe constraints on choosing which plots of land to give up for expropriation.[12] In the surveys and interviews conducted by Waris et al. (1952), displaced farmers express many complaints about their new farms. However, the criticism was almost exclusively directed towards the size of the new farms and the overall differences in land quality between the source and resettlement areas rather than receiving lower quality land within the resettlement area.

In short, the resettlement policy was unlikely to entirely achieve the aim of replacing the lost farms with fully comparable new ones. On the other hand, the average deterioration of farm quality was likely to be relatively modest for most of the displaced farmers. Some were even likely to end

[12]The Department of Land Settlement at the Ministry of Agriculture was in charge of the resettlement policy. It was led by one of the most influential politicians of the post-war Finland, Johannes Virolainen, who himself was a son of a displaced farmer and became known for defending the interests of the displaced population. The expropriation of land was entrusted to 147 Land Redemption Boards, each consisting of a surveyor engineer acting as a chairman, two expert members (a graduate in agricultural sciences and a forester), a lay member representing the local land-owners, and a lay member representing the displaced farmers. The distribution of the land among the displaced farmers was conducted by another 147 Settlement Boards, consisting of a graduate in agricultural sciences as a chairman, one representative of the local land owners and two representatives of the displaced farmers. In addition, eight Supervisory Bureaus and eight Courts of Appeals were set up to ensure the fairness of the process. (Pihkala, 1952)

up with a better farm than what they had lost.

Nevertheless, it is worthwhile to examine how deterioration of farm quality would play out in the context of the Roy model discussed above. The first prediction is straightforward: when $A(i)$ decreases, condition (2) holds for a larger number of farmers and more farmers move to the modern sector. However, reconciling the positive impact on income with worsening opportunities in agriculture is complicated. In the model, income increases for those farmers who could have earned more in the modern sector than in agriculture already before the war and now leave agriculture due to getting a sufficiently bad farm. However, the impact on other groups is negative. Displaced farmers who remain in agriculture despite having to settle for a lower quality farm will earn less. The same is true for farmers who would have maximized their income by staying in their old farms, but now move to the modern sector due to worsening opportunities in agriculture. Thus, changes in farm quality would increase average income only if the increase in income among the first type fo farmers were sufficiently large to more than offset the income losses among the other groups. This is logically possible, but would require a very specific joint distribution of skills, migration costs and farm qualities.[13]

We also note that an interpretation based solely on deterioration of farm quality would need to address the question of why the displaced farmers did not acquire more or better land. Improving one's farm in this period was clearly feasible for the displaced farmers, who were entitled to subsidized loans from the State Settlement Fund. Furthermore, the quality of land is relatively easy to assess and rural Finland has an abundance of forest land, which can be cleared into fields. Indeed, the government made a significant investment in land clearing by establishing a joint-stock company, *Pellonraivaus Oy,* to ensure access to modern equipment for this purpose. Thus, availability of land, credit constraints or asymmetric information about the quality of the land are unlikely to have prevented displaced farmers from buying more land. For these reasons, we conclude that while deterioration of farm quality was likely to push some farmers to leave agriculture, it is unlikely to fully explain our results.

Human Capital Another way the resettlement could have affected returns to migration is through a direct impact on human capital. This channel would be in line with Becker et al. (2019), who show that the offspring of individuals forced to move from areas Poland ceded to the Soviet Union at the end of World War II are substantially more educated than the offspring of the non-displaced

[13]Empirically, we found that displaced and non-displaced farmers had similar average income conditional on their sector and location (Table 6). However, as we discussed in Section 4.4, these estimates do not have a causal interpretation, because they condition on factors that are themselves affected by the resettlement. This "bad control" problem is easy to see in the context of the Roy model. Assuming that migration costs are independent of displacement status and that the average post-war farm quality is lower among the displaced farmers, it follows that the displaced farmers who remain in agriculture must have stronger average comparative advantage in agriculture than non-displaced farmers who stay in agriculture.

population living in the same locations. They interpret this difference as evidence for forced migration shifting preferences towards investing in portable assets, such as education, and present compelling complementary evidence supporting this interpretation.

In the context of our Roy model, a direct effect on education would improve skills. If returns to formal education were higher outside of agriculture, additional education would increase $z_n(i)$ more than $z_a(i)$, thus increasing returns to leaving agriculture and pushing displaced farmers towards the modern sector. As a consequence, their income would have increased through both the migration and the human capital channel. However, it is important to note that the causal chain could also run in the opposite direction. That is, if the displacement affected the likelihood of leaving agriculture through other mechanisms, higher returns to education in the modern sector would create incentives to acquire more education.[14]

Our data appear to be more consistent with transitions to the modern sector driving education than the other way around. We find a positive impact of displacement on education for those living in rural areas before the war and no effect on the urban population (Table 3 and Section 4.2). Furthermore, consistent with migration preceding educational investments, the effect of forced migration on moving to non-agriculture and cities among the rural population are clearly present already in 1950, while the impact on education appears to occur sometime between 1950 and 1970 (Tables 3 and 5). We stress that these observations do not rule out the possibility that forced migration affected preferences towards education. However, it seems unlikely that a direct impact on education is a major factor explaining our results.

5.3 Cost of Migration I: Networks, Culture, Discrimination and Learning

The analysis above suggests that changes in returns to migration alone are unlikely to explain our main results. Thus, we believe that an important part of the story lies on the other side of condition (2), i.e. forced migration reducing migration costs. We next discuss how expansion of dispersed networks, destruction of local networks, cultural differences, discrimination, and learning could lead to such effects. We leave our final candidate, attachment to a place (habit formation), for the last subsection.

Expansion of Dispersed Networks One way the displacement may have reduced migration costs is that it may have created valuable social networks. In particular, the initial evacuations could have created geographically dispersed networks that could have facilitated the flow of information about job and business opportunities. In the evacuation phase, the displaced population

[14] See also Nakamura et al. (2019) for related discussion in the case of Iceland and for a formal model where location and education choices are made simultaneously.

of each ceded municipality was transported to a designated evacuation area and the local population was obliged to provide them with shelter. As a consequence, most displaced persons were hosted by a local family during the winters of 1940-41 and 1944-45. During the summer of 1945, the displaced farmers received their new farms from locations that were, on average, roughly 150 kilometers away from their 1944–45 evacuation areas. According to Waris et al. (1952, p. 240), at least some of the displaced persons maintained contacts with their former host families also after moving to the resettlement areas.

The key challenge in explaining our results with evacuation networks is that the evacuation areas were rural and the families hosting the displaced population were largely farmers (who had space to accommodate the evacuees). Thus, these networks were not particularly well-suited for conveying information about non-agricultural job opportunities. On the other hand, some rural municipalities became local manufacturing centers in the post-war period (Sarvimäki, 2011; Mitrunen, 2019) and being evacuated to such a municipality could thus have been valuable. Furthermore, some of the locals living in the evacuation areas in the 1940s migrated to cities later on and could thus expand the network available for the displaced population.

We examine the role of the evacuation networks by comparing displaced persons exposed to different kinds of evacuation areas. This approach builds on the assumption that if the evacuation networks facilitated information flows, they were more valuable to displaced persons who had been evacuated into more prosperous or faster-growing locations. Estimates reported in Panel A of Table 8 show that displaced persons evacuated into more economically viable areas—as measured by the 1971 average income of individuals living in these locations already in 1939—do not earn more than those evacuated into other places. The estimates using data on all displaced persons suggest that a one-euro increase in the 1971 earnings of locals living in the 1940 and 1944 evacuation areas, respectively, predicts 0.00 (95% confidence interval -0.13–0.12 euros) and and -0.03 (CI -0.16–0.10) euros higher 1971 income among the displaced. Breaking down the displaced population by gender and pre-war status yields both positive and negative point estimates of comparable magnitude that are all statistically insignificant.[15] As a robustness check, we also report similar estimates using pre-war taxable income per capita (panel B) as an alternative measure of evacuation area quality. Again, we find precisely estimated zeros. The only exception is the estimate for rural men working outside of agriculture in 1939, for whom the estimates suggest that a standard deviation increase in the pre-war per capita income of the 1944 evacuation area would decrease 1971 income by 1,210 euros. However, as we report 28 estimates in Table 8, giving weight to one statistically significant estimate is unlikely to be appropriate. Thus, we interpret these results as suggesting that while the evacuation areas may have influenced the lives of the

[15] We do not report estimates for the urban population here, because there were only three cities in the ceded areas and thus standard errors clustered at the 1939 residence municipality level are unlikely to yield valid inference.

displaced population, they did not play an important economic role.

Destruction of Local Networks In addition to creating new social networks, forced migration may have destroyed old ones. In particular, it may have disrupted close-knit local networks that allow informal credit and insurance arrangements to persist (Banerjee and Newman, 1998; Karlan et al., 2009; Munshi and Rosenzweig, 2016). Losing access to such informal arrangements would reduce the opportunity cost of migration and could thus account for our results. Furthermore, the displacement could have affected access to farmers' co-operatives and thus pushed displaced farmers away from agriculture.

The importance of local networks was not lost on Finnish policy makers, who made every effort to resettle displaced villagers close to each other. However, the extent to which this principle could be implemented in practice varied across locations. As Figure 2 illustrates, even neighboring municipalities ended up being resettled into areas that differed vastly in size. This variation is driven by differences in the presence of large farms and government-owned land—which determined the amount of land that could be distributed to displaced farmers—and thus comparable displaced farmers were resettled to areas of different sizes. Specifically, our hypothesis is that being resettled into a larger resettlement area led to longer geographical distances between members of pre-war local networks and thus weakened these networks. If local networks were an important force holding back migration, displaced farmers resettled further away from their old network members would be more likely to move to the modern sector and thus to earn more than those resettled into more compact areas.

Columns (1) and (5) of Table 9 report results from regressing annual income in 1971 (panel A) and an indicator for working outside of agriculture in 1970 (panel B) on the size of the resettlement area and observable pre-war characteristics. We focus on displaced farmers because those working outside of agriculture were not directly affected by the resettlement plan. The treatment variable is the size of the resettlement area scaled with the size of the origin municipality (interquartile range 1.4), where the scaling is due to accounting for pre-war differences in population density. The estimates suggest that ending up into one unit larger resettlement area increased income of men by €130 (95% confidence interval -€810–€1,070) and decreased income of women by €110 (CI -€370–€120). The corresponding estimates for the likelihood of working outside of agriculture are a decline of 0.42 percentage points for men (CI -3.8–2.9 percentage points) and 0.77 percentage points for women (CI -4.0–2.4 percentage points).

We stress that this result does not necessarily imply that local networks were irrelevant. Indeed, they could be so valuable that the displaced persons maintained them despite the increased distance between the members of the network. Nevertheless, we do not find support for the hypothesis that the destruction of local networks explains why forced migration affected income and mobility.

This conclusion is also in line with Waris et al. (1952, p. 220-230), who found that displaced persons were welcomed by local farmers' co-operatives and other local clubs and societies. As we discuss in more detail next, the displaced persons seem to have integrated well also to other kinds of local networks.

Cultural Differences and Discrimination Our third candidate for why many displaced farmers decided to leave agriculture is that they may have felt out of place in their resettlement areas. Finland has a rich variety of local dialects and customs, and cultural differences between displaced and local populations may have been relatively large, in particular in the resettlement areas further away from the ceded areas. On the other hand, the displaced population could have faced discrimination, particularly if the locals held a grudge for having had their land expropriated.

A large fraction of Waris et al. (1952) is devoted to this question. Their conclusion is that while many respondents recalled tensions during the evacuation phase (when the local population suddenly had to share their homes with the evacuees), the displaced population seem to have quickly integrated into their resettlement areas' social life. Three-quarters of the displaced persons reported having visited at least one local during the past month and the same share of locals reported visiting at least one displaced family. About half of the displaced persons participating in the in-depth interviews included a local in their list of five best friends. Another sign of integration is the high rate of intermarriage and the fact that the displaced population actively entered local politics as part of the established parties rather than forming their own parties. However, later qualitative work has argued that Waris et al. (1952) paints an overly harmonious picture of the interactions between the displaced and local populations. In particular, more recent work emphasizes prejudices towards displaced persons who were members of the Orthodox Church (Alasuutari and Alasuutari, 2009; Kananen, 2018; Tepora, 2018).

Table 9 adds to this evidence by examining whether displaced farmers resettled into a culturally more different or more hostile location were more likely to leave agriculture and to have higher income. We use geographical distance from source area as a proxy for cultural distance and the share of the redistributed land coming from private landowners (instead of government-owned land) as a proxy for the hostility of the locals. All point estimates are small and the only statistically significant one suggests that women resettled to areas where a larger share of redistributed land came from private landowners were less likely to leave agriculture.[16] In a separate analysis, we show

[16]Specifically, the estimates suggest that being resettled 100 kilometers further to the west increased annual income by €60 (CI -€600–€710) for men and by €50 (CI -€190–€290) for women and decreased the likelihood of leaving agriculture by 0.1 percentage points (CI -2.8–2.6 percentage points) for men and by 1.1 percentage points (CI -3.6–1.4 percentage points) for women. Similarly, a ten percentage points increase in the share of private land increased annual income in 1971 by €260 (CI -€120–€640) for men and by €38 (CI -€100–€180) for women, while increasing the likelihood of leaving agriculture by 0.4 percentage points (CI -1.3–2.2 percentage points) for men and decreasing it by 1.9 percentage points (CI -3.2–0.4 percentage points) for women.

that the impacts of being displaced were very similar to the members of the (majority) Lutheran and (minority) Orthodox churches despite the fact that the latter group was likely to experience substantially more discrimination than the former (Appendix Table A12). Thus we conclude that while the displaced population was likely to face at least some cultural differences and discrimination, these factors are unlikely to drive our main results.

Learning and inertia Forced migration could have reduced migration costs also through learning. Specifically, the shock of being displaced could have helped people to overcome inertia or allowed them to update their beliefs about their ability to settle in a new environment. However, the survey evidence does not support this hypothesis. Indeed, a recurring theme in Waris et al. (1952) is that displaced persons were tired of moving. These survey responses, and the return migration discussed next, do not support the hypothesis that forced migration would have made the displaced persons more open or willing to migrate to a new location.

5.4 Cost of Migration II: Habit Formation

We end by examining what the displaced farmers replied when asked about their future migration intentions and why they were planning to move (or to stay). Again, we draw from Waris et al. (1952), who included direct questions on this topic in their surveys and interviews. They conclude that most of the displaced persons expected to settle down in their current locations and summarize their results as follows:

> *"The explanations for why it was time to settle down varied widely, [but our] overall conclusion is that the displaced Karelians started to feel part of their new communities. The only reservation that came up again and again was: "but if only one could move back to Karelia...!". The lost area, and everything related to it, gave rise to overwhelming emotions. Just saying the word, Karelia, seemed to put everything that belonged to the past, and that was now lost, into a bright, admiring light. In comparison to that everything else was gray, dull, inferior."* (p. 314)

A typical sentiment appears to be captured by a displaced farmer saying: "*Since I cannot go back to my old land, it does not matter where I live. Besides, the locals here are nice people.*" (p. 180). The perception that the displaced persons held a high opinion of the ceded areas is supported also from another direction: when locals were asked to name an annoying trait among the displaced persons, the most frequently mentioned ones were categorized as "exaggeration, praising the past too much" (p. 212).

Return Migration Of course, the respondents may have been tempted to exaggerate their desire to move back to Karelia in low-stake surveys and interviews. However, an informative feature of the Finnish resettlement experience is that the displaced persons were given an opportunity to reveal their preferences. As we discussed in Section 2, Finland reconquested the lost areas during the summer of 1941 and held them for three years. In December 1941, the ownership rights in the ceded area were formally restored and displaced persons could apply for a permit to return. If they were granted the permit, they could give up the farm they had received as part of the resettlement policy and move back to what was left of their old farms.

Importantly, there was a genuine opportunity to stay in the resettlement areas and, in many ways, staying would have been a sensible choice. Much of the housing stock of the reconquested area was destroyed and conditions were often harsh. For example, roughly 22,000 farm buildings needed to be rebuilt (Pihkala, 1952) and some return migrants endured food shortages (Waris et al., 1952). Despite the high cost and risk, return migration proved extremely popular, particularly among the farmers. Indeed, 97% of the farms allocated as part of the 1940 Emergency Settlement Act (*Pika-asutuslaki*) were returned in exchange for the old farms (Pihkala, 1952). In total, 70% of the displaced population returned, even though not everyone was granted a permit to return due to housing shortage, proximity to the frontline or for being considered politically unreliable. Remarkably, almost half of the population returned even to the municipalities where more than 85% of the housing stock had been destroyed (Appendix Figure A5).

In 1944, the return migrants were again evacuated and resettled. About six years later, Waris et al. (1952) asked their interviewees to describe their return migration experience and how they felt about it afterward. Among those who had returned, 84% argued that they did not regret doing so. Instead, they described at length the joy of returning home. Waris et al. (1952, p. 170) summarize the sentiment of the displaced persons by quoting one saying: *"It was better to spend those years [in Karelia]. Maybe it was a financial loss, but that is not the most important thing*".

These observations suggest that impact of being forced to migrate were not limited to farm quality, human capital, networks, cultural distance, discrimination or learning. In fact, these factors play a relatively minor role in qualitative research and popular writings about the displaced Karelians (see e.g. Alasuutari and Alasuutari, 2009). Instead, a major theme in this work is the displaced persons' intense longing for their *homes* that goes far beyond economic opportunities available in the ceded area.

A Roy Model with Habit Formation In order to rationalize the survey answers and return migration behavior described above, we augment the Roy model sketched in Section 5.1 with habit formation for living in a location. As before, we assume that individuals differ from each other in terms of their productivities across sectors and in their migration costs. Now, we also explicitly

define the source for migration costs. Specifically, we assume that in addition to consumption, people derive utility from their residential location and that this utility increases with the time the person has already lived in the location. Following Becker and Murphy (1988), we call the latter property habit formation and model it as "location capital" that affects contemporaneous utility and is accumulated by "consuming" more of a location, i.e. living there. We assume that people are heterogeneous in their taste for location capital and that they maximize their lifetime utility by choosing a sequence of locations while taking into account their lifetime budget constraint, the accumulation process of location capital and their initial location capital accumulated during childhood.[17]

Figure 3 illustrates the key tradeoffs in this model. We consider individuals in early adulthood, who have grown up on a farm and are now choosing whether to remain or to move to the modern sector. For simplicity, we assume that one can leave agriculture only by moving to another location. The choice is determined by the relative strengths of comparative advantage (vertical axes) and taste for location capital (horizontal axis). The solid line plots the indifference curve between moving and staying. It slopes upwards, because the extra income required for making a person indifferent between moving and staying increases with the extent to which he values staying at home. Thus, those who have a combination of sufficiently high returns to leaving agriculture and/or sufficiently low taste for location capital move, while the rest stay on their farms.

Suppose that at some later stage, some people are forced to migrate. They now lose the location capital accumulated before the displacement and hence their indifference curves become flat. Thus they will choose the location based on available income alone. Those who had chosen to leave agriculture already before the displacement remain in the modern sector. Similarly, those who maximize their income in agriculture will remain in agriculture. However, farmers who previously chose to stay on their farms due to taste for location capital now move to the modern sector. As a consequence, forced migration increases their income.

This variant of the Roy model also provides a simple rationalization for the return migration behavior discussed above. That is, the displaced persons started to accumulate a new stock of consumption capital in their new locations—in the words of Waris et al. (1952), they "started to feel part of their new communities"—but this process takes time and the location capital tied to their old homes may depreciate slowly. Thus, there can be a long period during which they would be willing to give up part of their income in order to return to their previous homes.

In our view, the most important insight from this version of the Roy model is that the welfare implications of the resettlement may have an important intergenerational element. As all Roy

[17]For illustration, we present one version of such model in the Appendix, but the basic intuition should be general to any model including the following ingredients: (a) people differ in their comparative advantages across locations, (b) people derive utility from consumption and location, (c) utility from living in a location increases with location capital, and (d) people choose their locations only after having first accumulated some initial location capital.

models, our variant suggests that while forced migration increased average income of the first generation displaced persons, it is likely to have reduced their welfare. After all, if the sector and location they ended up in after the war provided them with higher welfare than their home farms, they could have moved there even if they were not displaced. Importantly, however, in a model including habit formation, forced migration increases the welfare of their children. The reason is that these children, particularly those born after the displacement, will now accumulate their initial location capital in economically more viable places. Thus, they will be able to reap the benefits of better labor market opportunities without having to pay the price of leaving home.

6 Conclusions

This paper examines the long-term effects of resettling 11% of the Finnish population during World War II. Our key finding is that forced migration increased the likelihood of leaving agriculture and long-term income among the rural population. We examine the extent to which these effects arise from forced migration affecting farm quality, education, networks, learning, cultural differences and discrimination, and find limited or no evidence supporting these mechanisms. Thus, we argue that the most compelling explanation for our results is that people are willing to forgo even large monetary gains to stay at home. When this opportunity is removed from their choice set, they become more responsive to economic incentives and hence more mobile. We formalize this argument with a Roy model extended with habit formation for residential location.

At some level, the conclusion that people value their homes is obvious. Stories of someone choosing a humble life in her home town instead of a lucrative career elsewhere are familiar to most readers. Yet, while anecdotes are abundant, systematic evidence remains scarce—most likely because any empirical investigation of habit formation faces the challenge that people who differ in their personal histories are likely to also differ along other (unobservable) dimensions. Thus, it is hard to distinguish between alternative mechanisms using observational data, but experimental manipulation of personal histories is typically infeasible. Forced migrations can create research designs for studying habit formation, but they typically come with the limitation of affecting people in multiple ways. In this regard, Finland's attempt to reconstruct the pre-war situation for displaced farmers as closely as possible provides a rare opportunity for gaining insight into this question.

Our results are consistent with the hypothesis that habit formation for residential location is a quantitatively important factor affecting the allocation of labor across sectors and locations. However, we emphasize that it alone is unlikely to explain the persistence of rural-urban income differences. We do not advocate a view that other mechanisms are irrelevant, but merely suggest that habit formation matters and that models including it provide additional insight. In particular, such models illustrate that even if labor is misallocated in terms of lost production, this need not entail

lost welfare. In the language of Becker and Murphy (1988), farmers may be "addicted" to their farms, but it may be "rational addiction" in the sense that given their initial location capital, they maximize their lifetime utility by staying in the low productivity traditional sector.

Such preferences are unlikely to be limited to farmers and may thus explain more broadly why people choose to stay in declining areas or industries. Indeed, we see no reason to think that mid-20th century Finns would have been particularly attached to their homes. Furthermore, economic development may ease other forms of migration costs and thus increase the relative role of habit formation. For example, better availability of formal insurance reduces the importance of risk and informal insurance in migration decisions. Similarly, improvements in information technology are likely to make dispersed networks less important for learning about job and business opportunities. By contrast, migration costs due to habit formation are not affected by such institutional or technological changes. If anything, people may become more likely to decide that they can afford to stay at home, even if it is costly, when they become more prosperous.

Finally, we note that analyzing location choices through the lens of a habit formation model reveals a potential intergenerational conflict, because children do not choose where to accumulate their initial location capital. Thus, if people could choose their locations at birth, they might choose differently than what they end up choosing later in life. In the context of post-war Finland, it is possible that while the displacement was a tragedy for those who had to leave their homes, it may have liberated their offspring to enjoy the benefits of growing up in locations that provide better economic opportunities. In other contexts, creating (preferably softer) incentives for people to move to higher productivity areas could also give rise to similar intergenerational effects.

References

Adamopoulos, T. and D. Restuccia (2014). The Size Distribution of Farms and International Productivity Differences. *American Economic Review 104*(6), 1667–97.

Alasuutari, P. and M. Alasuutari (2009). Narration and Ritual Formation of Diasporic Identity: The Case of Second Generation Karelian Evacuees. *Identities: Global Studies in Culture and Power 16*(3), 321–341.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy 113*(1), 151–184.

Arellano-Bover, J. (2018). Displacement, Diversity, and Mobility: Career Impacts of Japanese American Internment. Mimeo, Stanford University.

Atkin, D. (2013). Trade, Tastes, and Nutrition in India. *American Economic Review 103*(5), 1629–1663.

Atkin, D. (2016). The Caloric Costs of Culture: Evidence from Indian Migrants. *American Economic Review 106*(4), 1144–81.

Banerjee, A. and A. Newman (1998). Information, the Dual Economy, and Development. *Review of Economic Studies 65*(4), 631–653.

Bauer, T. K., S. Braun, and M. Kvasnicka (2013). The Economic Integration of Forced Migrants: Evidence for Post-War Germany. *Economic Journal 123*(571), 998–1024.

Bazzi, S., A. Gaduh, A. D. Rothenberg, and M. Wong (2016). Skill Transferability, Migration, and Development: Evidence from Population Resettlement in Indonesia. *American Economic Review 106*(9), 2658–2698.

Becker, G. S. and K. M. Murphy (1988). A Theory of Rational Addiction. *Journal of Political Economy 96*(4), 675–700.

Becker, S. O. and A. Ferrara (2019). Consequences of Forced Migration: Recent Findings. *Labour Economics, forthcoming*.

Becker, S. O., I. Grosfeld, P. Grosjean, N. Voigtländer, and E. Zhuravskaya (2019). Forced Migration and Human Capital: Evidence from Post-WWII Population Transfers. *forthcoming, American Economic Review*.

Bolt, J., R. Inklaar, H. de Jong, and J. L. van Zanden (2018). Rebasing Maddison: New Income Comparisons and the Shape of Long-run Economic Development. GGDC Research Memorandum 174.

Bryan, B., S. Chowdhury, and A. M. Mobarak (2014). Under-Investment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh. *Econometrica 82*(5), 1671–1748.

Bryan, G. and M. Morten (2018). The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia. *Journal of Political Economy, forthcoming*.

Caselli, F. (2005). Accounting for Cross-Country Income Differences. *Handbook of Economic Growth 1*, 679–741.

Caselli, F. and W. Coleman (2001). The US Structural Transformation and Regional Convergence: A Reinterpretation. *Journal of Political Economy 109*(3), 584–616.

Chyn, E. (2018). Moved to Opportunity: The Long-Run Effects of Public Housing Demolition on Children. *American Economic Review 108*(10), 3028–56.

Deryugina, T., L. Kawano, and S. Levitt (2018). The Economic Impact of Hurricane Katrina on Its Victims: Evidence from Individual Tax Returns. *American Economic Journal: Applied Economics 10*(2), 202–33.

Diamond, R. (2016). The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000. *American Economic Review 106*(3), 479–524.

Fernando, A. N. (2016). Shackled to the Soil: The Long-Term Effects of Inherited Land on Labor Mobility and Consumption. Mimeo.

Gollin, D., D. Lagakos, and M. E. Waugh (2014). The Agricultural Productivity Gap. *Quarterly Journal of Economics 129*(2), 939–993.

Gollin, D., S. Parente, and R. Rogerson (2002). The Role of Agriculture in Development. *The American Economic Review 92*(2), 160–164.

Harris, J. and M. Todaro (1970). Migration, Unemployment and Development: a Two-Sector Analysis. *American Economic Review 60*(1), 126–142.

Haukka, J., J. Suvisaari, M. Sarvimäki, and P. Martikainen (2017). The Impact of Forced Migration on Mortality: A Cohort Study of 242,075 Finns from 1939–2010. *Epidemiology 28*(4), 587–593.

Heckman, J. J. and B. E. Honore (1990). The Empirical Content of the Roy Model. *Econometrica 58*(5), 1121–1149.

Hopenhayn, H. A. (2014). Firms, Misallocation, and Aggregate Productivity: A Review. *Annual Review of Economics 6*(1), 735–770.

Imbens, G. and K. Kalyanaraman (2012). Optimal Bandwidth Choice for the Regression Discontinuity Estimator. *Review of Economic Studies 79*(3), 933–959.

Imbens, G. W. and D. B. Rubin (1997). Estimating Outcome Distributions for Compliers in Instrumental Variables Models. *Review of Economic Studies 64*(4), 555.

Kananen, H. (2018). Ortodoksisen siirtoväen kontrolloitu sopeutuminen. In P. Kanervo, T. Kivistö, and O. Kleemola (Eds.), *Karjalani, Karjalani, maani ja maailmani*, pp. 105–1022. Turku: Siirtolaisuusinstituutti & Oy Sigillum Ab.

Karlan, D., M. Mobius, T. Rosenblat, and A. Szeidl (2009). Trust and Social Collateral. *Quarterly Journal of Economics 124*(3), 1307–1361.

Kennan, J. and J. R. Walker (2011). The Effect of Expected Income on Individual Migration Decisions. *Econometrica 79*(1), 211–251.

Lagakos, D., A. M. Mobarak, and M. E. Waugh (2018). The Welfare Effects of Encouraging Rural-Urban migration. NBER WP No. 24193.

Lagakos, D. and M. E. Waugh (2013). Selection, Agriculture, and Cross-Country Productivity Differences. *American Economic Review 103*(2), 948–80.

Lewis, W. A. (1955). *Theory of Economic Growth*. Routledge.

Lucas, R. (1997). Internal Migration in Developing Countries. In M. Rosenzweig and O. Stark (Eds.), *Handbook of Population and Family Economics*, Volume 1, pp. 721–798. Amsterdam: Elsevier.

Lucas, R. (2004). Life Earnings and Rural-Urban Migration. *Journal of Political Economy 112*(1), 29.

Lynch, R., V. Lummaa, K. Panchanathan, K. Middleton, A. Rotkirch, M. Danielsbacka, D. O'Brien, and J. Loehr (2019). Integration Involves a Trade-Off between Fertility and Status for World War II Evacuees. *Nature Human Behaviour*.

Mitrunen, M. (2019). War Reparations, Structural Change, and Intergenerational Mobility. Mimeo, Stockholm University.

Munshi, K. and M. Rosenzweig (2016). Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap. *American Economic Review 106*(1), 46–98.

Munshi, K. and N. Wilson (2011). Identity and Mobility: Historical Fractionalization, Parochial Institutions, and Occupational Choice in the American Midwest. Working paper.

Nakamura, E., J. Sigurdsson, and J. Steinsson (2019). The Gift of Moving: Intergenerational Consequences of a Mobility Shock. Mimeo, UC Berkeley and Stockholm University.

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics 37*(2), 187–204.

Pihkala, K. U. (1952). The Land Settlement Program of Finland. *Land Economics 28*(2), 147–159.

Pollak, R. A. (1970). Habit Formation and Dynamic Demand Functions. *The Journal of Political Economy*, 745–763.

Redding, S. and D. Sturm (2008). The Costs of Remoteness: Evidence from German Division and Reunification. *American Economic Review 98*(5), 1766–1797.

Rentola, K. (2001). Stalin, Mannerheim ja Suomen rauhanehdot 1944. *Historiallinen aikakauskirja 99*(1), 47–61.

Restuccia, D. and R. Rogerson (2017). The Causes and Costs of Misallocation. *Journal of Economic Perspectives 31*(3), 151–74.

Roback, J. (1982). Wages, Rents, and the Quality of Life. *Journal of Political Economy 90*(6), 1257–1278.

Rosen, S. (1979). Wage-Based Indexes of Urban Quality of Life. In P. Mieszkowski and M. Straszheim (Eds.), *Current Issues in Urban Economics*. Johns Hopkins University Press.

Ruiz, I. and C. Vargas-Silva (2013). The Economics of Forced Migration. *The Journal of Development Studies 49*(6), 772–784.

Saarela, J. and F. Finnäs (2009). Forced Migration and Mortality in the Very Long Term: Did Perestroika Affect Death Rates also in Finland? *Demography 46*(3), 575–587.

Sarvimäki, M. (2011). Agglomeration in the Periphery. Mimeo, Aalto University.

Statistics Finland (1942). *IV Varallisuustilastoa: Tulo ja omaisuusverotilastoa vuodelta 1938*. Helsinki: Valtioneuvoston kirjapaino.

Statistics Finland (1953). *IV Varallisuustilastoa: Tulo ja omaisuusverotilastoa vuodelta 1950*. Helsinki: Valtioneuvoston kirjapaino.

Statistics Finland (1979). *Väestön elinkeino: Väestö elinkeinon mukaan kunnittain vuosina 1880-1975*. Helsinki: Statistics Finland.

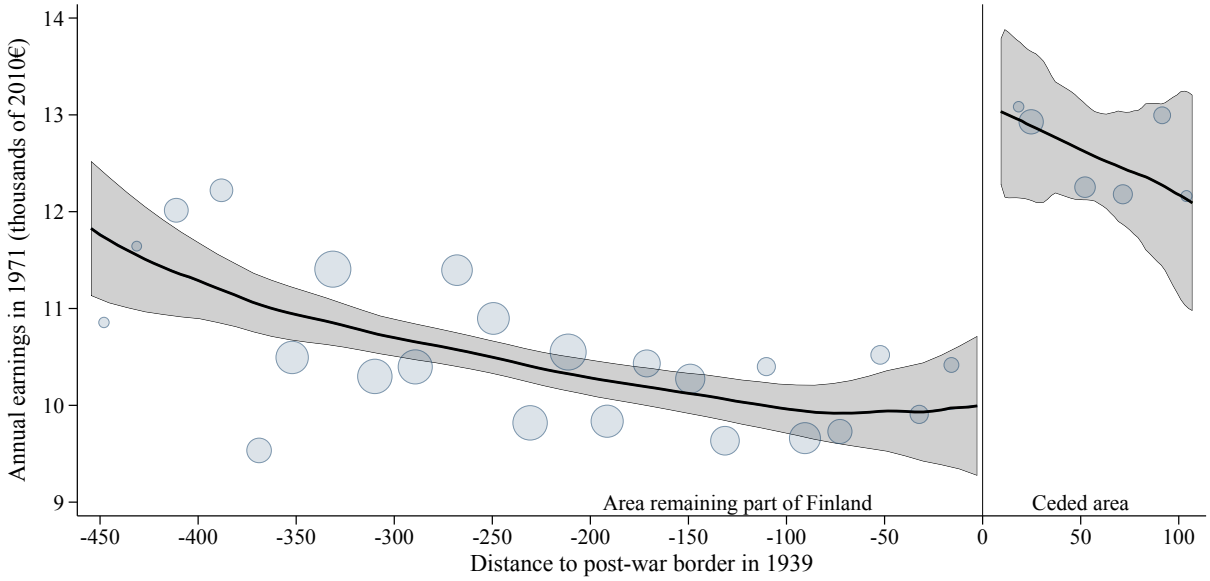
Tepora, T. (2018). Itsekkäät evakot ja ahneet isännät – talvisodan evakuoinnin kokemushistoriaa. In P. Kanervo, T. Kivistö, and O. Kleemola (Eds.), *Karjalani, Karjalani, maani ja maailmani*, pp. 52–72. Turku: Siirtolaisuusinstituutti & Oy Sigillum Ab.

Waris, H., V. Jyrkilä, K. Raitasuo, and J. Siipi (1952). *Siirtoväen sopeutuminen: Tutkimus Suomen karjalaisen siirtoväen sosiaalisesta sopeutumisesta*. Helsinki: Otava.

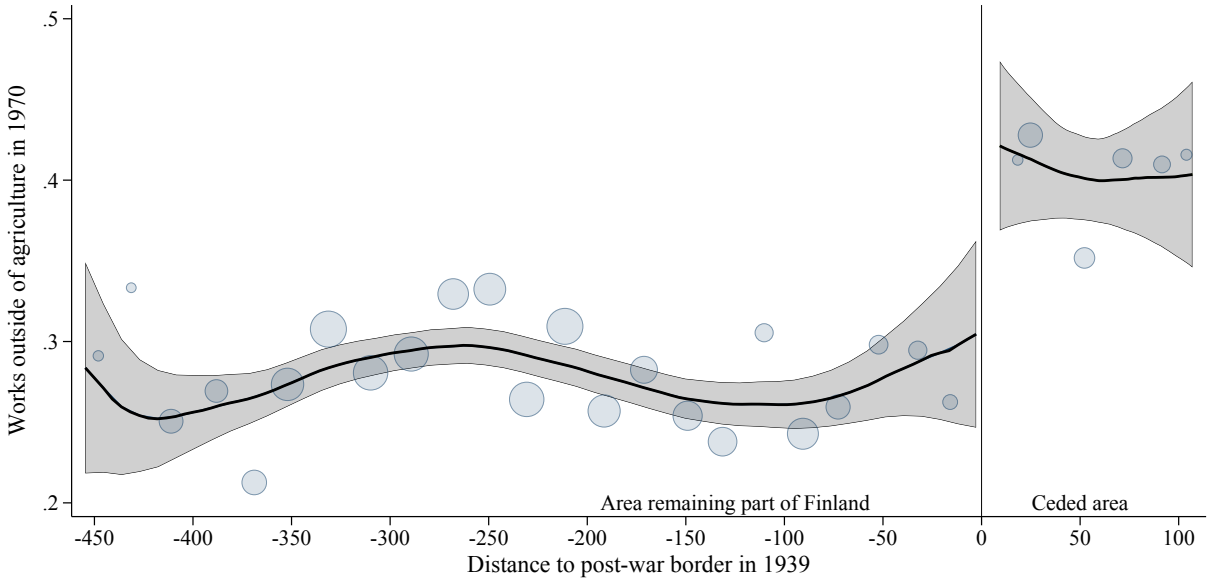
Young, A. (2013). Inequality, the Urban-Rural Gap and Migration. *Quarterly Journal of Economics 128*(4), 1727–1785.

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

(a) Annual Income in 1971 (including zeros)

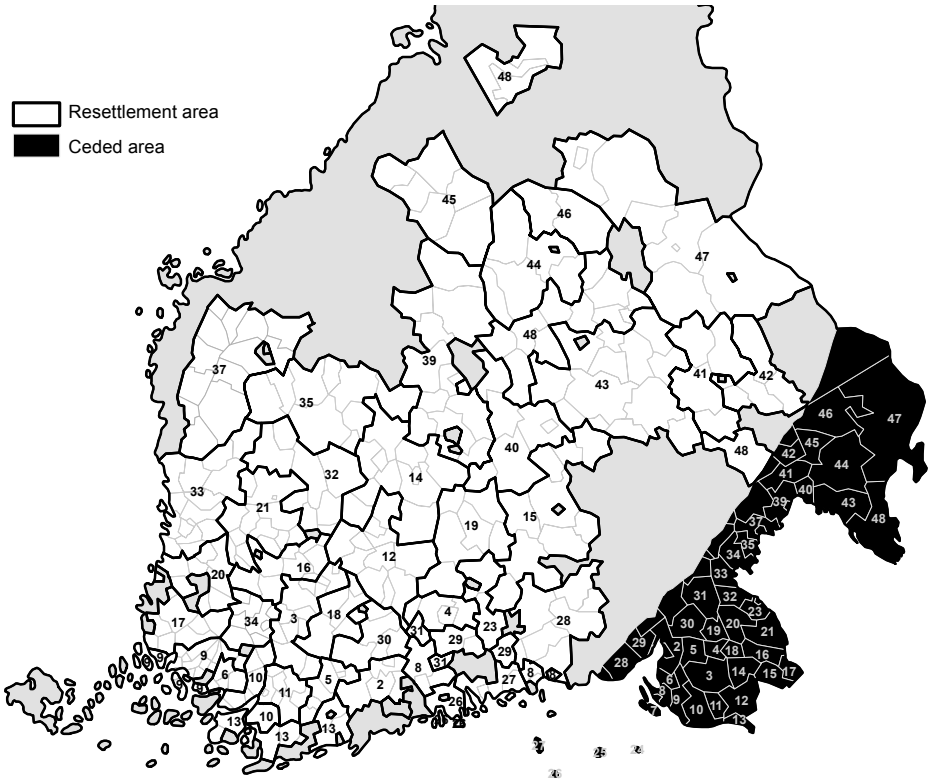


(b) Non-Agricultural Employment in 1970



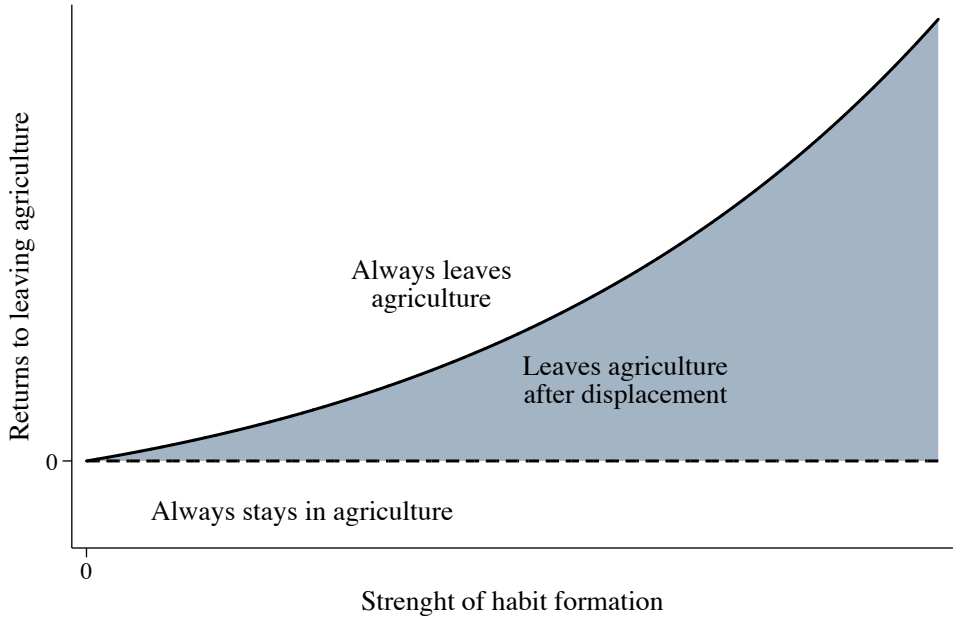
Note: The top panel plots annual income in 1971 in thousands of 2010€ (including zeros) by distance of the 1939 residence municipality to the post-war border (in kilometers). Positive numbers on the x-axis refer to areas that were ceded to the Soviet Union in 1940/45. The sample used in this analysis consists of 14,317 men born between 1907 and 1925, who worked in agriculture in 1939. They were 14–32 years old in 1939 and 46–64 years old in 1971. The bottom panel presents similar analysis for an indicator for the person working outside of agriculture in 1970. The lines represent local linear estimates using the edge kernel and the optimal bandwidth of Imbens and Kalyanaraman (2012). The dots correspond to the sample means by 20km bins. On average, each dot represents 477 individuals.

Figure 2: The 1945 Resettlement Plan



Note: This map represents the ceded area and the resettlement areas. The numbers refer to the ceded municipalities in the ceded area and to their corresponding resettlement areas in the remaining parts of Finland.

Figure 3: Migration Decisions in an Illustrative Model of Habit Formation



Note: See Section 5.4 for discussion.

Table 1: Pre-War Characteristics

	Men						Women					
	Agricultural		Other rural		Urban		Agricultural		Other rural		Urban	
	Non-disp. (1)	Dis-placed (2)	Non-disp. (3)	Dis-placed (4)	Non-disp. (5)	Dis-placed (6)	Non-disp. (7)	Dis-placed (8)	Non-disp. (9)	Dis-placed (10)	Non-disp. (11)	Non-disp. (12)
A: Demographics												
Age	22.8	22.8	22.9	22.8	24.0	23.4	24.2	24.7	22.7	22.5	24.0	24.1
Swedish-speaker	0.07	0.00	0.08	0.00	0.16	0.01	0.07	0.00	0.07	0.00	0.13	0.01
Migrated prior to 1939	0.17	0.15	0.38	0.36	0.61	0.55	0.28	0.24	0.39	0.35	0.70	0.62
Orthodox	0.00	0.12	0.00	0.11	0.01	0.04	0.00	0.13	0.00	0.12	0.01	0.04
B: Socioeconomic status												
Entrepeneur	0.30	0.42	0.08	0.06	0.05	0.05	0.04	0.03	0.02	0.02	0.02	0.03
White-collar	0.02	0.02	0.10	0.13	0.20	0.25	0.00	0.00	0.10	0.10	0.23	0.23
Blue-collar	0.36	0.16	0.56	0.53	0.59	0.52	0.23	0.13	0.16	0.13	0.31	0.25
Assisting family member	0.32	0.41	0.26	0.28	0.00	0.01	0.73	0.84	0.72	0.75	0.00	0.01
Out of labor force	0.00	0.00	0.34	0.34	0.20	0.23	0.00	0.00	0.74	0.76	0.47	0.52
C: Sector of employment												
Manufacturing	0.00	0.00	0.28	0.20	0.33	0.25	0.00	0.00	0.08	0.06	0.18	0.14
Construction	0.00	0.00	0.14	0.14	0.11	0.11	0.00	0.00	0.00	0.00	0.00	0.00
Services	0.00	0.00	0.24	0.32	0.33	0.38	0.00	0.00	0.18	0.18	0.33	0.33
Other	0.00	0.00	0.00	0.00	0.16	0.18	0.00	0.00	0.00	0.00	0.44	0.49
D: Characteristics of the municipality of residence												
Average taxable income	1.41	1.38	1.95	1.65	6.75	5.61	1.45	1.38	1.81	1.58	6.80	5.62
Agricultural LFS	0.83	0.81	0.74	0.76	0.14	0.01	0.83	0.80	0.76	0.78	0.13	0.01
Latitude	69.4	67.7	68.9	67.7	67.9	67.5	69.3	67.7	69.0	67.8	67.9	67.5
Observations	12,940	1,377	11,142	1,258	8,079	889	7,366	831	19,633	2,259	11,584	1,191

Note: Panels A–C report individual-level information referring to 1939 as measured by the retrospective questions in the 1950 census. Pre-war migration is defined as living outside of municipality of birth in 1939. Panel D reports municipality-level information from Statistics Finland (1942, 1979) that is linked to the individual-level data using information on the 1939 residence municipality.

Table 2: Impact of Forced Migration on Annual Income in 1971

	Cont. Means	Baseline		Oster's Bound	Spatial RD		Resettlement Area FEs	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				A: Men by 1939 status				
Agri-cultural	10.5	2.08 (0.37)	2.06 (0.36)	2.05 (0.36)	3.12 (0.86)	3.07 (0.73)	1.26 (0.33)	1.67 (0.47)
Other rural	16.2	1.14 (0.59)	1.83 (0.70)	2.18 (0.57)	1.01 (2.54)	0.39 (2.80)	0.81 (0.78)	2.06 (1.15)
Urban .	23.7	-2.48 (1.95)	-4.52 (1.32)	-5.23 (0.86)	.	.	.	.
				B: Women by 1939 status				
Agri-cultural	1.9	0.57 (0.16)	0.66 (0.19)	0.71 (0.23)	0.49 (0.48)	0.37 (0.44)	0.45 (0.15)	0.64 (0.25)
Other rural	4.8	1.10 (0.19)	1.46 (0.22)	1.65 (0.24)	1.51 (0.69)	2.11 (0.58)	0.84 (0.20)	1.53 (0.28)
Urban .	8.8	-0.65 (0.73)	-1.29 (0.64)	-1.52 (0.34)	.	.	.	.
Controlling for:								
Pre-war char.		no	yes	.	no	yes	no	yes
Resettlement area		no	no	.	no	no	yes	yes

Note: Point estimates and standard errors (in parentheses) from regressing annual income in 1971 in thousands of 2010€ (including zeros) on an indicator for displacement status and, in some specifications, observable characteristics measured before the war (year of birth dummies, indicators for speaking Swedish as one's mother tongue, belonging to the Orthodox church, living outside of one's municipality of birth, sector of employment, socioeconomic status, quintile dummies for residence municipality's taxable income per capita, labor force share in agriculture and latitude). For Oster's bounds, we set the hypothetical R^2 as $1.3\tilde{R}$, where $\tilde{R}$ is the R^2 from the regression including the control variables observed in our data; see Oster (2019) for discussion. RD estimates come from local linear regression allowing different slopes for the distance to the post-war border by displacement status, using triangle-shaped kernel around the border and optimal bandwidths of Imbens and Kalyanaraman (2012). Estimates reported in the last columns control for resettlement area fixed-effects constructed using the 1939 residence municipality information for locals and the 1945 Resettlement Plan for displaced population (regardless of where they actually lived after the war). These estimates exclude non-displaced persons living outside of the resettlement area from the sample.

Table 3: Impact of forced migration on industry, employment, urbanization and education in 1970

	Works outside of agriculture			Employed			Lives in a city			Holds a secondary degree		
	Cont. Means	Estimates		Cont. Means	Estimates		Cont. Means	Estimates		Cont. Means	Estimates	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A: Men by 1939 status												
Agricultural	28.2	15.0	17.0	72.2	-1.4	-4.1	11.5	13.4	15.1	8.9	3.2	2.0
		(2.1)	(2.2)		(1.5)	(2.1)		(1.5)	(2.2)		(1.0)	(1.2)
Other rural	63.4	2.5	1.7	75.9	0.0	-2.7	27.7	16.9	20.2	17.5	5.9	7.9
		(2.0)	(2.4)		(1.6)	(2.0)		(2.1)	(2.6)		(1.4)	(1.4)
Urban	76.1	-4.5	.	78.5	-4.7	.	84.2	-15.7	.	31.9	0.9	.
		(0.8)			(1.0)			(1.6)			(1.6)	
B: Women by 1939 status												
Agricultural	14.7	5.1	5.2	40.8	-1.8	-2.9	10.9	10.6	11.1	9.4	0.5	-1.0
		(1.9)	(2.4)		(2.6)	(3.1)		(2.3)	(3.0)		(1.7)	(2.0)
Other rural	34.0	9.5	8.2	44.7	4.2	3.9	23.7	18.9	23.7	13.9	2.8	4.3
		(1.4)	(1.8)		(1.3)	(1.6)		(2.2)	(2.3)		(1.1)	(1.1)
Urban	54.7	-5.1	.	55.8	-4.4	.	83.3	-14.9	.	22.1	-0.2	.
		(1.4)			(1.2)			(2.8)			(1.5)	
Controlling for:												
Pre-war char.		yes	yes		yes	yes		yes	yes		yes	yes
Resettlement area		no	yes		no	yes		no	yes		no	yes

Note: Point estimates and standard errors (in parentheses) from regressing indicators for working outside of agriculture (columns 1–3), being employed (columns 4–6), living in a city (columns 7–9), or holding a secondary degree (columns 10–12) in 1970 on an indicator for displacement status and observable characteristics measured before the war. Estimates reported in the last subcolumn of each supercolumn also control for resettlement area fixed effects and exclude non-displaced persons living outside of the resettlement area. See the note to Table 2 for details and Appendix Figure A2 for alternative specifications.

Table 4: Intergenerational Impact of Forced Migration

	Income			Secondary degree		
	Cont. Means	Parent Displaced		Cont. Means	Parent Displaced	
	(1)	(2)	(3)	(4)	(5)	(6)
			A: Father's 1939 status			
Agricultural	21.6	1.16	0.78	73.9	4.64	4.88
		(0.46)	(0.53)		(1.63)	(2.20)
Other rural	23.4	0.69	0.78	76.2	1.31	6.32
		(0.52)	(0.61)		(1.64)	(2.31)
Urban	26.4	-2.13	.	77.7	-0.44	.
		(0.34)			(1.78)	
			B: Mother's 1939 status			
Agricultural	21.2	0.18	0.03	71.2	2.25	3.59
		(0.79)	(0.97)		(2.80)	(3.50)
Other rural	22.2	0.81	0.78	72.5	3.83	4.85
		(0.39)	(0.52)		(1.53)	(2.03)
Urban	25.9	-2.92	.	75.7	1.20	.
		(0.48)			(1.39)	
Controlling for:						
Pre-war characteristics		yes	yes		yes	yes
Resettlement Area FE		no	yes		no	yes

Note: Point estimates and standard errors (in parentheses) from regressing children's average annual income at age 30–40 in thousands of 2010€ (columns 1–3) or an indicator for the child holding a secondary degree in 2011 (columns 4–6) on an indicator for parent's displacement status and characteristics measured before the war. Standard errors are two-way clustered at parent and parent's 1939 municipality level. Estimates reported in the last subcolumn of each supercolumn also control for resettlement area fixed-effects.

Table 5: Impact of forced migration on imputed income, industry, employment, urbanization and education in 1950

	Imputed income			Works outside of agriculture			Employed			Lives in a city			Holds a secondary degree		
	Cont. Means	Estimates		Cont. Means	Estimates		Cont. Means	Estimates		Cont. Means	Estimates		Cont. Means	Estimates	
	(1)	(2)	(3)	(3)	(4)	(5)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
A: Men by 1939 status															
Agricultural	4.9	0.72 (0.09)	0.69 (0.12)	19.9	20.5 (2.1)	19.6 (2.4)	79.9	4.4 (1.7)	3.5 (2.3)	4.3	7.8 (1.4)	7.4 (1.3)	0.9	-0.3 (0.3)	-0.5 (0.4)
Other rural	7.0	0.25 (0.15)	0.29 (0.15)	74.4	6.2 (1.8)	7.0 (2.1)	90.0	0.4 (1.1)	0.6 (1.4)	15.3	16.6 (2.0)	16.0 (2.5)	8.3	2.4 (1.1)	2.1 (1.4)
Urban	7.9	-0.19 (0.07)	.	90.7	-2.5 (1.2)	.	94.1	-1.2 (0.6)	.	85.4	-24.7 (2.4)	.	22.8	-6.1 (2.2)	.
B: Women by 1939 status															
Agricultural	5.5	0.56 (0.21)	0.68 (0.21)	15.2	13.0 (2.1)	13.2 (2.8)	31.4	11.8 (2.0)	12.9 (2.6)	3.5	3.8 (1.4)	2.1 (1.6)	1.3	0.1 (0.5)	0.5 (0.7)
Other rural	6.9	0.39 (0.10)	0.45 (0.14)	57.1	16.4 (2.0)	18.4 (2.7)	69.9	10.8 (1.7)	9.9 (2.2)	12.9	17.6 (1.6)	19.0 (1.8)	8.2	2.6 (0.9)	3.8 (1.3)
Urban	7.7	-0.08 (0.08)	.	89.8	-2.5 (0.7)	.	92.1	-2.4 (0.6)	.	82.8	-20.8 (3.4)	.	23.3	-3.3 (2.7)	.
Controlling for:															
Pre-war char.		yes	yes		yes	yes		yes	yes		yes	yes		yes	yes
Resettlement area		no	yes		no	yes		no	yes		no	yes		no	yes

Note: Point estimates and standard errors (in parentheses) from regressing indicators for working outside of agriculture (columns 1–3), working in agriculture (columns 4–7), being employed (columns 7–9), living in a city (columns 10–12), or holding a secondary degree (columns 13–15) in 1970 on an indicator for displacement status and observable characteristics measured before the war. Estimates reported in the last subcolumn of each supercolumn also control for resettlement area fixed effects and exclude non-displaced persons living outside of the resettlement area. See the note to Table 2 for details and Appendix Figure A3 for alternative specifications.

Table 6: Annual Income in 1971 Conditional on Post-War Sector and Location

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Agricultural	2.06	0.93	0.67	0.34	0.66	0.25	0.29	0.02
	(0.36)	(0.32)	(0.31)	(0.32)	(0.19)	(0.13)	(0.12)	(0.21)
Other rural population	1.83	1.53	0.73	-0.46	1.46	0.50	0.46	0.10
	(0.70)	(0.63)	(0.52)	(0.52)	(0.22)	(0.16)	(0.14)	(0.18)
Urban population	-4.52	-3.80	-2.87	-2.47	-1.29	-0.73	-0.50	-0.36
	(1.32)	(1.34)	(1.20)	(0.53)	(0.64)	(0.50)	(0.40)	(0.19)
Controlling for:								
Works outside of agriculture in 1970	no	yes	yes	yes	no	yes	yes	yes
Education in 1970	no	no	yes	yes	no	no	yes	yes
2-digit industry and municipality in 1970	no	no	no	yes	no	no	no	yes

Note: Estimates from regressing annual income in 1971 on an indicator for displacement status and background characteristics (see the notes to Table 2). In columns 2–4 and 6–8, we gradually control for an indicator for working outside of agriculture in 1970, indicators for level of education in 1970 and indicators for 2-digit industry and municipality of residence in 1970. Standard errors (in parentheses) are clustered at the 1939 municipality of residence level. Each estimate stems from a separate regression. Pre-war characteristics: see note to Table 2.

Table 7: Returns to Leaving Agriculture

	OLS			2SLS			
	Average income in agriculture	Returns to nonagri-culture		Compliers' agricultural income		Returns to nonagri-culture	
Status in 1939	(1)	(2)	(3)	(5)	(6)	(7)	(8)
Men	9.9	6.3	6.0	12.1	11.1	10.8	9.3
		(0.3)	(0.3)	(1.8)	(1.5)	(2.6)	(2.1)
Women	0.7	8.3	8.2	0.8	-0.1	12.6	12.7
		(0.2)	(0.2)	(1.0)	(1.3)	(2.4)	(2.9)
Controlling for pre-war characteristics		no	yes	no	yes	no	yes

Note: Column 1 shows the annual earnings in 1971 in 2010€ for those working in agriculture. Columns 2–3 report OLS estimates for an indicator variable taking value one if the person works outside of agriculture in 1970 and zero otherwise. Standard errors (in parentheses) are clustered at the 1939 residence municipality level. Columns 5–6 report estimates for the average income of compliers if they would remain in agriculture (Imbens and Rubin, 1997). All regressions control for resettlement area fixed-effects. Pre-war characteristics: see the notes to Table 2.

Table 8: Evacuation Area Quality and Long-Term Income

		Men			Women		
	All	All	Agri-cultural	Other rural	All	Agri-cultural	Other rural
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A: Average income of "locals" in 1971							
1940 evacuation area	0.00	-0.10	0.16	-0.38	0.08	-0.09	0.08
	(0.06)	(0.11)	(0.16)	(0.30)	(0.05)	(0.05)	(0.09)
1944 evacuation area	-0.03	-0.06	-0.17	0.17	-0.04	-0.10	0.03
	(0.07)	(0.13)	(0.09)	(0.25)	(0.06)	(0.05)	(0.06)
B: Taxable income per capita in 1939 (standardized)							
1940 evacuation area	0.01	-0.07	0.20	-0.62	0.12	0.04	-0.24
	(0.20)	(0.26)	(0.29)	(0.68)	(0.18)	(0.15)	(0.22)
1944 evacuation area	0.09	0.30	-0.50	1.21	-0.12	-0.03	-0.04
	(0.16)	(0.27)	(0.26)	(0.46)	(0.10)	(0.13)	(0.13)
Observations	7,506	3,382	1,337	1,156	4,124	831	2,122

Note: Estimates for β from a regression $y_i = \alpha + EA_i\beta + X_{io}\delta + \varepsilon_{ijt}$, where y_{ij} is annual income in 1971, EA_i is the average 1971 income of individuals living in her evacuation area already in 1939 (panel A) or standardized taxable income per capita in 1939 (panel B), and X_{i0} is a vector of observable pre-war characteristics (see the notes to Table 2). Each column reports estimates from a separate regression. Standard errors (in parentheses) are clustered at the 1939 municipality of residence. The number of observations deviates slightly from those reported in Table 1, because we have not been able to find information on the evacuation area of a few municipalities. We do not report estimates separately for the urban population due to the small number of ceded urban municipalities.

Table 9: Resettlement Area Characteristics and Long-Term Outcomes

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Annual income in 1971								
Size relative	0.08		.	0.14	-0.09		.	-0.08
to origin area	(0.30)			(0.30)	(0.09)			(0.10)
Distance to	.	0.06	.	-0.12	.	0.05	.	0.02
origin area		(0.33)		(0.38)		(0.12)		(0.14)
Exproriated	.		2.60	3.34	.		0.38	0.10
private land			(1.94)	(2.18)			(0.73)	(0.84)
B: Works outside of agriculture in 1970								
Size relative	-0.26		.	-0.20	-0.55		.	-1.29
to origin area	(1.06)			(1.14)	(1.16)			(1.12)
Distance to	.	-0.10	.	-0.55	.	-1.13	.	0.01
origin area		(1.36)		(1.47)		(1.27)		(1.24)
Exproriated	.		4.37	5.34	.		-18.6	-22.2
private land			(8.80)	(9.75)			(7.3)	(6.9)
Observations	1,376				831			

Note: Estimates for β from a regression $y_{ij} = \alpha + \beta RA_j + X_{io}\delta + \varepsilon_{ijt}$, where y_{ij} is the either annual income in 1971 in thousands of 2010€, including zeros (panel A) or an indicator for working outside of agriculture in 1970 (panel B), RA_j is a measure of resettlement area characteristics, and X_{i0} is a vector of observable pre-war characteristics. All regressions are run using data including only displaced farmers. Standard errors (in parentheses) are clustered at the 1939 municipality of residence level. Size is measured as the total area of the resettlement area divided by the area of the origin municipality (standard deviation (SD) 1.53, interquartile range (IQR) 1.42), distance to origin area in 100 of kilometers (SD 0.85, IQR 1.52), and expropriated land as the share of all redistributed agricultural land expropriated from private landowners (SD 0.18, IQR 0.26).

Appendix

Data details

Census data The starting point of our data is the 1950 population census. The original census forms were sorted by municipality, within municipalities in alphabetical order and then filed in folders. In 1995, Statistics Finland drew a sample from the full 1950 census by picking every tenth folder, see Statistics Finland (1998) for details. Almost all of the information on the census forms was manually inserted into a database. The resulting sample contains information on 411,629 persons from 392 of municipalities (out of a total of 547 municipalities) corresponding to 10.3% of the full population.

The 1950 census data is linked to other individual-level data sources using social security numbers. The social security numbers were introduced in September 1964 and thus had to be collected from the Population Register using names, dates of birth, gender and place of birth. The match rate is very high. Social security numbers were found for 82.5% of the individuals included in the 1950 census sample. Furthermore, information from the 1970 census could be found for 73.1% of the original sample. In comparison, taking into account mortality and emigration, at most 74.5% of the population present in the 1950 census was also present in the 1970 census (Statistics Finland, 1998).

Income data Our information on individual-level income comes from the 1971 tax register. A key concern in using tax data in an analysis like ours is that taxable income might not be a comparable measure of true income for individuals working in and outside of agriculture. However, by 1971, agricultural profits were treated as taxable earned income and taxed according to the same rates as wage earnings. While production for own consumption was not taxed, agriculture had become increasingly specialized and, for example, Pihkala (1982) estimates that 90% of agricultural products were sold on the market and hence taxed. Much of the remaining 10% consisted of feeder crops used on the farm as intermediate inputs.

Figure A1 provides another check for the comparability of our income measure by plotting consumption expenditure against gross-income for farming and non-farming households using data from the 1971 Household Budget Survey (Statistics Finland, 1976). These data contain information on 1,186 households, of which we categorize 372 as farmers and 814 non-farmers based on the household reference person's main occupation. The consumption information was collected by the households during four weeks and includes the purchase value of items produced by the households (e.g. vegetables grown in their own garden). However, gifts and transfers to other households are not included as consumption. Gross income consists of all earnings and capital incomes as well as all public income transfers. The income information stems from the tax records of each household member and thus corresponds closely to the

income measure used in our analysis. The expenditures have been annualized by Statistics Finland.

Panel A of Figure A1 represents the full data, while panel B focuses only on the inner 98% of each marginal distribution.[1] We find no indication of the tax records underestimating the consumption possibilities of farmers. In fact, the only statistically significant difference—at very high levels of income—suggest the opposite. However, these differences are driven by outliers. Once we drop the outliers, the confidence regions of the two groups overlap throughout (Panel B, Figure A1). Table A1 confirms this result by reporting regression coefficients using up to a fourth-order polynomial in income interacted with an indicator for the household's reference person working in farming.

Urban status We defined municipalities use Statistics Finland's definitions of cities used in the 1950 census (*kuntatyyppi*). We augment this definition to include three municipalities—Espoo, Vantaa and Kauniainen—that are part of the Helsinki metropolitan area.

Second-generation We also have information on the children of persons included in the 1950 census sample. Statistics Finland has identified these children using their dataset on parent-child links. For each child, we observe education taken from the register of degrees, income from tax registers and basic demographic variables from the census and administrative registers.

A limitation of these data is that we observe only one parent for 42% of the children. This feature of the data is likely driven by the fact that for both parents to be observed, they have to be included in the 1950 census sample. Given the sampling scheme described above, we observe both (known) parents if they lived in the same household. Furthermore, we are more likely to observe them if they were living in the same municipality in 1950.[2] As a consequence, children for whom we observe both parents are likely to differ from children for whom we observe only one parent. This selection process is hard to characterize and may have been influenced by the resettlement. For these reasons, we have opted for an analysis sample in which the first generation is defined using information on only one parent. While this approach facilitates the interpretation of the estimates, it also means that some children categorized into the control group do have one displaced parent.

Additional empirical results

Descriptive statistics Tables A2–A5 show a detailed examination of the pre-war differences between displaced and non-displaced rural population. For comparison, columns 1–2 report the same numbers we reported in Table 1, i.e. sample averages of the background vari-

[1] That is, panel B uses "shaved" data, where we have excluded 42 observations with income or expenditure that is lower than the 1st or higher then the 99th percentile of each distribution.

[2] Moreover, this likelihood is affected by the size of the municipality. Specifically, if all census forms of a municipality would fit into one folder, we would observe everyone living in the municipality. As the size of the municipality grows, i.e. the census forms fill more folders, the expected share of individuals ending into our sample approaches 10%.

ables by future displacement status. The remaining columns show estimates and standard errors from regressions of the form

$$x_{0i} = \alpha + \beta D_i + \mathbf{Z_{0i}}\gamma + \varepsilon_{it} \tag{A1}$$

where x_{0i} is a pre-war characteristic of individual i, D_i is an indicator for future displacement status (i.e. living in the ceded area before the war) and $\mathbf{Z_{0i}}$ is a vector of other pre-war characteristics we may want to condition on. Columns 3–4 report the baseline differences, columns 5–6 control for the distance of the 1939 residence municipality to the post-war border (and its interaction with future displacement status), and columns 7–8 condition on resettlement area fixed-effects.

Farm size Table A9 presents farm size distributions for the ceded, partly ceded and other municipalities as reported in the 1930 and 1940 Agricultural Censuses. In 1930, there were 30,415 farms in the ceded area of which 10,530 were smaller than three hectares and 19,885 larger three hectares. The majority of the small farms were likely owned or leased by people whose primary job was not in agriculture, but who complemented their income with part-time farming.

Part-time agriculture was taken into account in the resettlement policy, which distributed 13,362 "Part-time holdings" (2–6 hectares) and 19,622 "Agricultural holdings" (6–15 hectares) to the displaced population (Pihkala, 1952, Table V). We categorize farms smaller than three hectares in 1930 as likely part-time farms (to be replaced by "part-time holdings") and those larger than likely full-time farms (to be replaced by "agricultural holdings"), because this provided the best match between the number of farms distributed in the 1940s and the number of farms present in the ceded area in 1930. We recognize that a limitation of this approximation is that the share of large farms in the ceded area may have changed between 1930 and 1939. The next Agricultural Census was conducted in the fall of 1941, when the first part of the resettlement policy had been completed and the first return migrants had moved back to their old farms in the area Finland had taken back in the summer of 1941 (see Sections 2.1 and 6.4 of the main paper for details). Nevertheless, we note that the farm size distribution remained roughly constant at the national level between 1930 and 1941.

Table A10 examines whether farmers who lost more land due to the resettlement were more likely to leave agriculture. Ideally, we would have examined effect heterogeneity by the size of origin farm at farmer level, but this information is not included in the 1950 census. Thus, we measure the role of reduction in farm size using municipality-level information on the share of farms larger than 15 hectares in 1930 in the municipality where each person lived in 1939. We calculate these shares using data only on farms that were larger than three hectares in order focus on full-time farms (see above).

The results show that rural men coming from municipalities that had more large farms tended to earn more in 1971 and to work more outside of agriculture in 1970 than those coming from places where farms were smaller. The point estimate for interaction between pre-war farm size and displacement status suggests that male farmers coming from 90th percentile of the pre-war farm size distribution (15% of farms larger than 15 hectares in 1930) had 320 euros higher income in 1971 and were 1.4 percentage points more likely to work outside of agriculture in 1970 than those coming from the 10th percentile of the pre-war

farm size distribution (2% large farms). However, the estimates are imprecise and statistically insignificant and far from statistical significance. Thus, we are not able to provide strong evidence in favor or against the reduction of farm size pushing farmers towards the modern sector. Importantly, however, the main effect of being displaced—corresponding to all farms of the source area being below 15 hectares before the war—remain large. In fact, some of the estimates are larger than our main estimates reported in Tables 2 and 3 of the main paper due to the strong association between the outcome variables and pre-war farm size (which is not controlled for in our main specification).

Regional yields Table A11 reports average yields using regional-level information of various crops as reported in the 1930 Agricultural Census. The average yields for the most popular crop (oats) were about 8% higher in the Viipuri region (which included most of the ceded area) in comparison to the average yield in all of Finland. On the other hand, the yields for the second most popular crop (rye) were 5% lower than the national average. In order to summarize the yield information, we first calculate a weighted average for each region using the agricultural land shares for each crop in the Viipuri area as weights. According to this index, yields in the Viipuri region were 3% higher than the national average. However, part of the national average reflects the conditions in northern Finland, where few displaced persons were resettled. For benchmark, we thus calculate a weighted average of the regional yields using the number of farms distributed to the displaced farmers in each region as weights. The results suggest that yields in the resettlement area were 1% higher than the national average. According to this proxy, the ceded areas would thus have had around 2% higher yields than the resettlement areas.

An illustrative Roy model with habit formation

We assume that agents' contemporaneous utility is an additive function of location capital and consumption

$$u_{jt}\left(c_t, l_{jt}\right) = c_t + l_{jt}^{\alpha} \tag{A2}$$

where t and j index time and location, c is consumption, l is the time the person has lived in the location ("location capital"), and $\alpha \geq 0$ is a parameter governing the strength of habit formation.

Individuals live in one location during childhood and then work for T periods. During their working life, they choose a sequence of locations, I_{jt}, to maximize lifetime utility

$$\max_{\{I_{jt}\}} U = \sum_{t=1}^{T} u_j\left(c_t, l_{jt}\right) \tag{A3}$$

subject to a budget constraint and accumulation of location capital. The budget constraint is

$$\sum_{t=1}^{T} c_t \leq \sum_{t=1}^{T} z_j \tag{A4}$$

where z_j is the income the agent would earn each period in location j. These incomes are constant over time and drawn at birth from a joint distribution of sectoral incomes $G(z_s)$. Locational capital is accumulated as

$$l_{jt} = l_{j,t-1} + I_{jt}, \tag{A5}$$

where I_{jt} is an indicator function taking the value one if the agent lives in location j in period t and zero otherwise.

We define period one as the stage when the person starts to make her own decisions and assume that she enters this stage with initial location capital, l_{j0}, accumulated during her childhood and thus reflecting the decisions of her parents. In order to keep the model as simple as possible, we treat locations and jobs as isomorphic and abstract away from discounting, depreciation of location capital, price and wage dynamics, local amenities, differences in regional prices and other migration costs.

Given these assumptions, utility is maximized by spreading consumption evenly over the life-cycle. Furthermore, if the agent migrates, she does so immediately at $t = 1$ in order to start accumulating location capital in the new location as soon as possible. For the same reason, it is never optimal to migrate twice. Thus the maximum utility the agent can derive from choosing location j for the remaining of her life is

$$V(z_j, l_{j0}, \alpha, T) = T z_j + \sum_{t=1}^{T} (l_{j0} + t)^{\alpha} \tag{A6}$$

Consider now the choice of whether to stay at home location h or to move to some other location. We denote the location providing her the highest income with m. If she migrates, she will choose m as she has no reason to move to a lower-paying location for which she does not have any location capital. She migrates to m if $V(z_m, 0, \alpha, T) > V(z_h, l_{h0}, \alpha, T)$ or

$$z_m - z_h \geq \frac{\sum_{t=1}^{T} (l_{h0} + t)^{\alpha} - \sum_{t=1}^{T} t^{\alpha}}{T} \tag{A7}$$

Condition (A7) illustrates that even in a highly stylized model, individuals may choose between migrating or staying for many reasons. Some stay because their skills have the highest return in their home location ($z_m - z_h < 0$). Others could increase their income by migrating, but would lose too much utility by giving up their initial location capital. This trade-off gives rise to the income difference required for migration that is larger for individuals who have lived longer in the same place (and thus have higher l_{h0}) and for those who have stronger location preferences (higher α). Furthermore, the minimum income difference required for migration decreases with the length of the future career, T.

References

Pihkala, E. (1982). Maa-ja metsätalouden uusi asema. In J. Ahvenainen, P. Erkki, and R. Viljo (Eds.), *Suomen taloushistoria 2: Teollistuva Suomi*. Helsinki: Tammi.

Pihkala, K. U. (1952). The Land Settlement Program of Finland. *Land Economics 28*(2), 147–159.

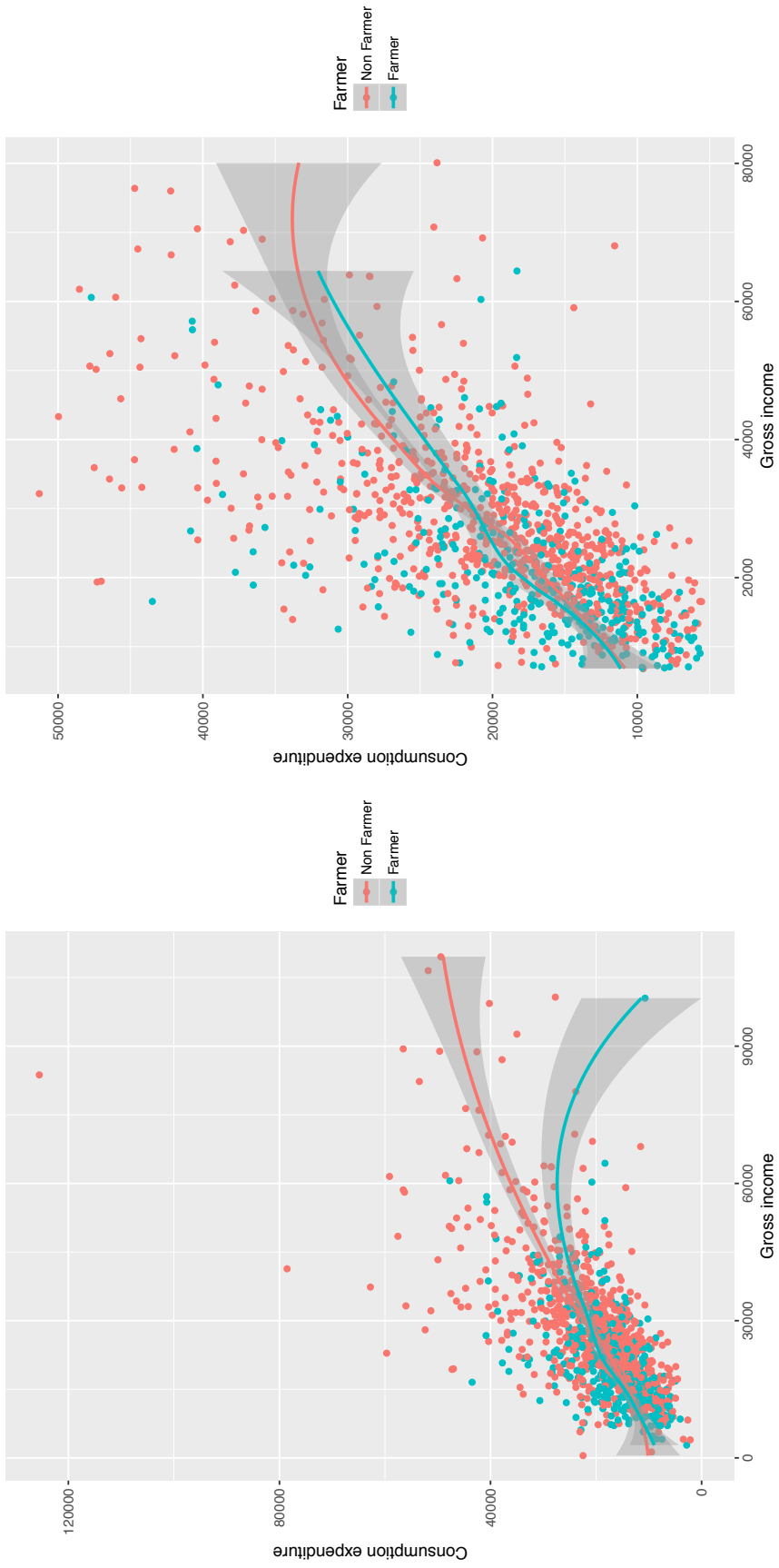
Statistics Finland (1972). Vuoden 1971 kuntien kalleustutkimus. Studies 19, Statistics Finland.

Statistics Finland (1976). Kotitaloustiedustelu 1971, osa 1, kotitalouksien kulutusmenot. Tilastollisia tiedonantoja 55. Helsinki, Finland.

Statistics Finland (1998). The Finnish Longitudinal Census Data File. Statistics Finland.

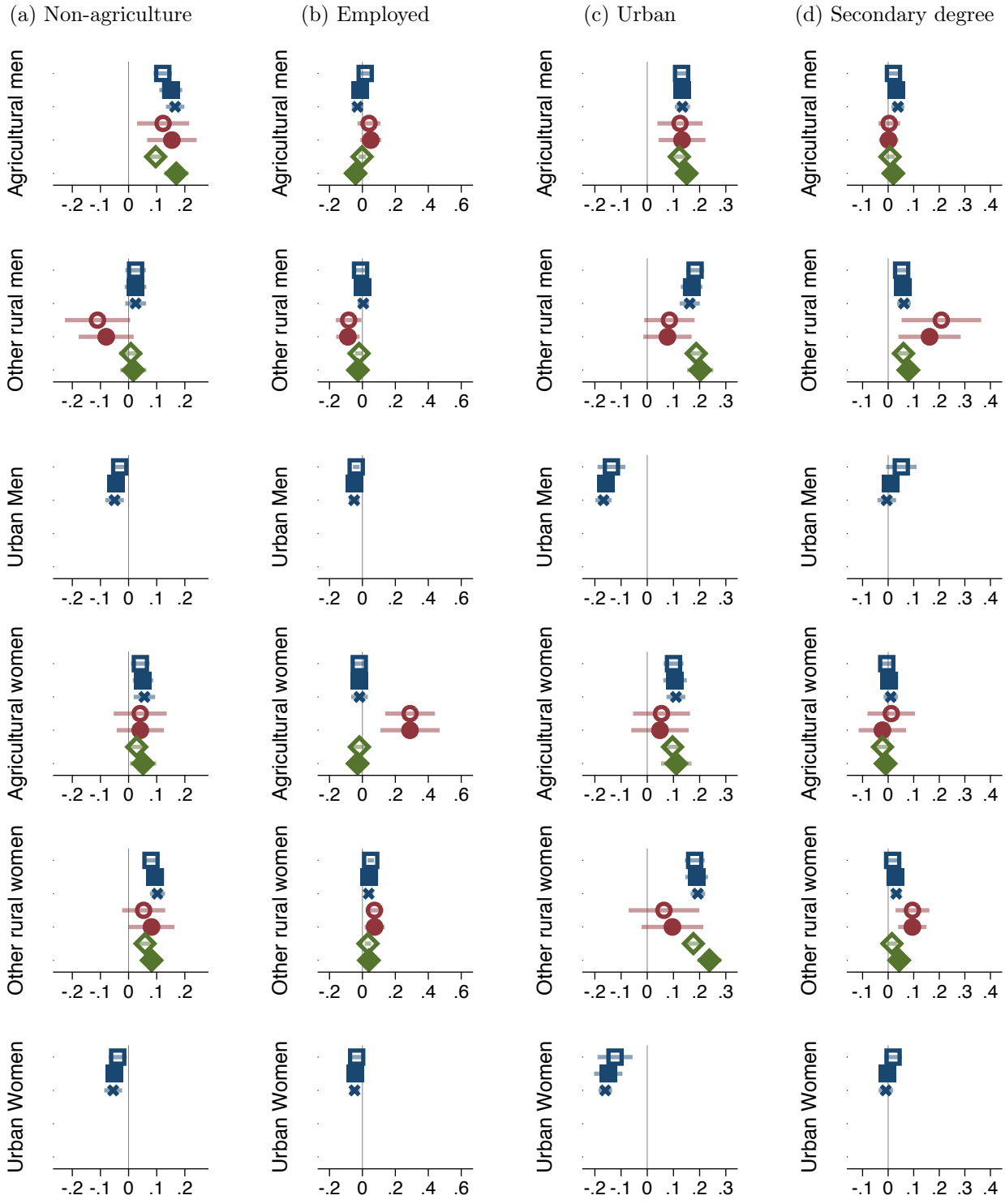
Waris, H., V. Jyrkilä, K. Raitasuo, and J. Siipi (1952). *Siirtoväen sopeutuminen: Tutkimus Suomen karjalaisen siirtoväen sosiaalisesta sopeutumisesta*. Helsinki: Otava.

Figure A1: Consumption expenditure and gross income among farm and non-farm households in 1971



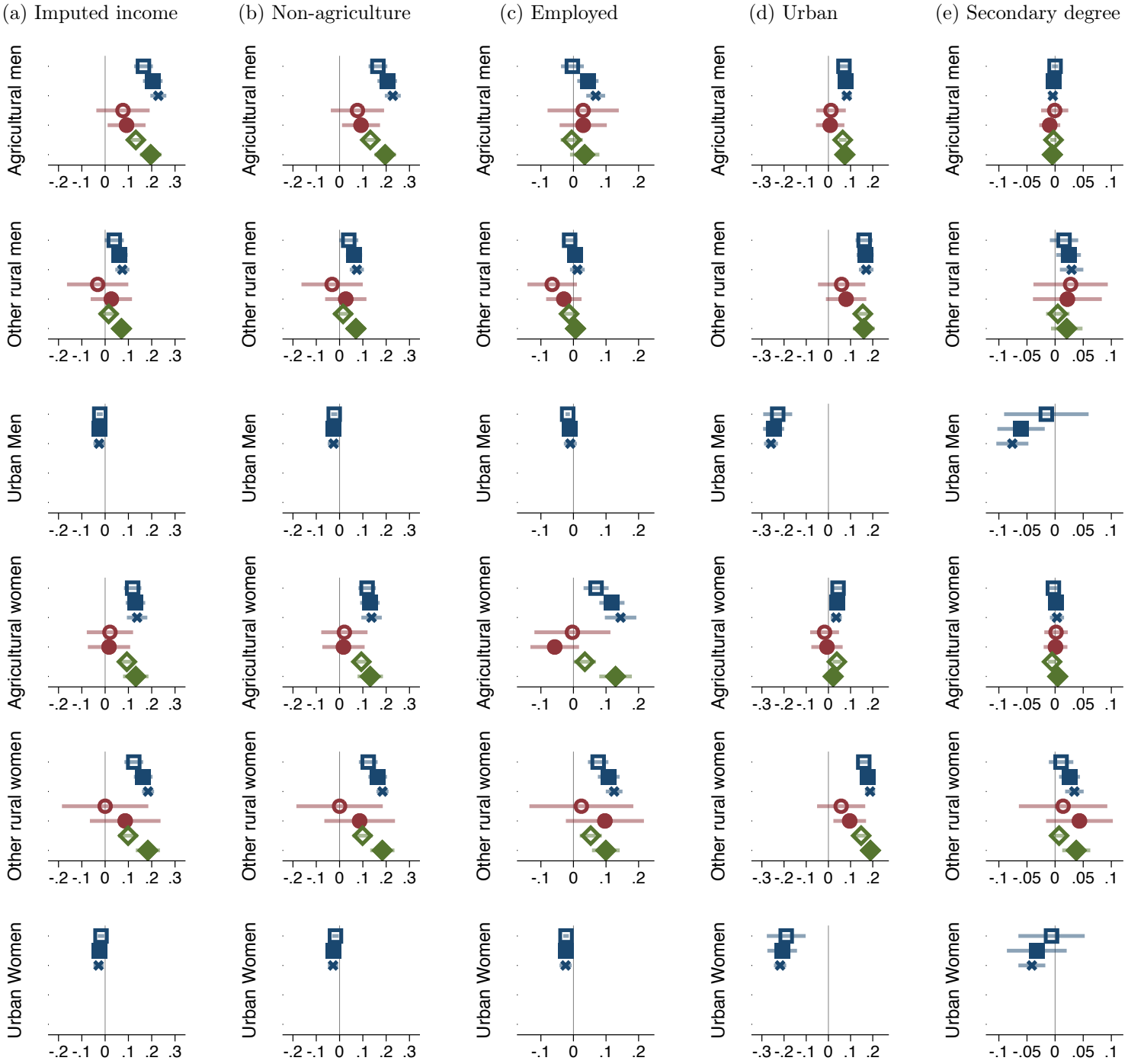
Note: Each point corresponds to a household. The lines represent loess-lines within the group (farmer, non-farmer). The shaded areas show the corresponding 95% confidence regions.

Figure A2: Impact of forced migration on sector, employment, urbanization and education in 1970



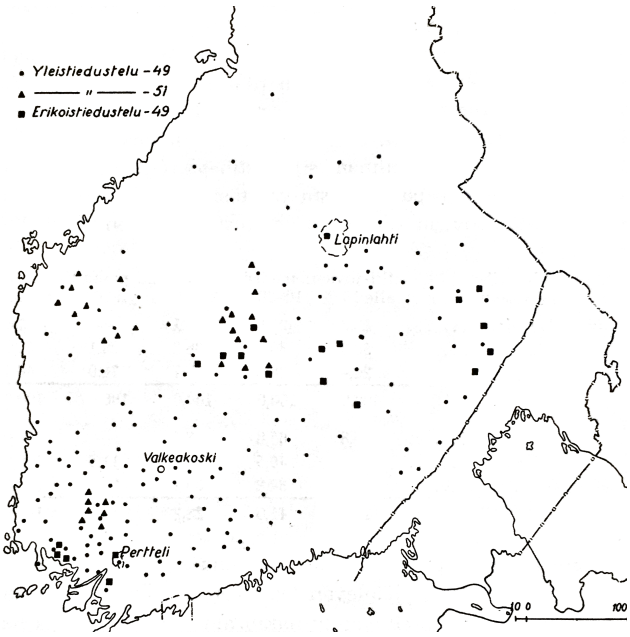
Note: Point estimates and 95% confidence intervals for the impact of forced migration on 1970 outcomes using specifications corresponding to those in Table 2. Solid markers refer to estimates controlling for pre-war observable characteristics. Squares show estimates from the baseline specification, crosses are the Oster Bounds, circles are the spatial RD specification and diamonds for specifications controlling for resettlement area fixed-effects.

Figure A3: Impact of forced migration on sector, employment, urbanization and education in 1950



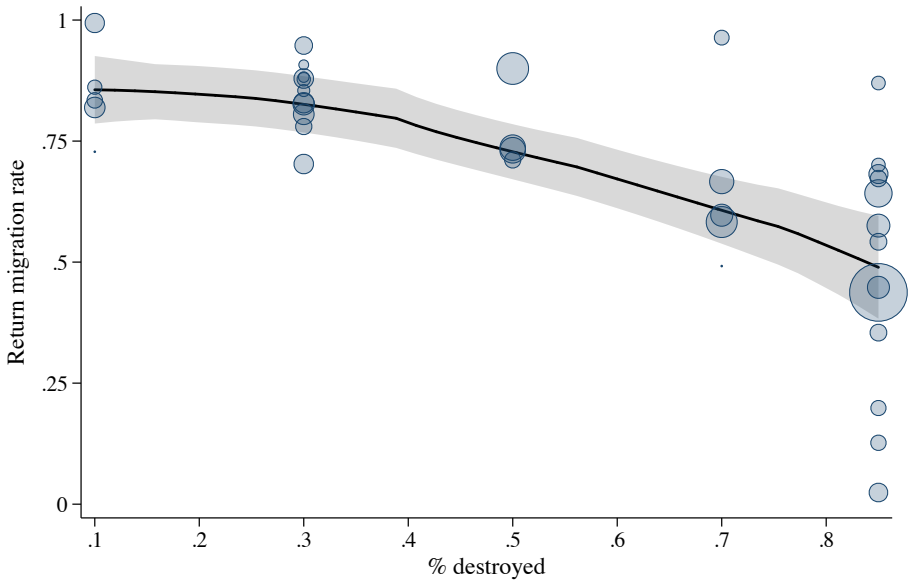
Note: Point estimates and 95% confidence intervals for the impact of forced migration on 1970 outcomes using specifications corresponding to those in Table 2. Solid markers refer to estimates controlling for pre-war observable characteristics. Squares show estimates from the baseline specification, crosses are the Oster Bounds, circles are the spatial RD specification and diamonds for specifications controlling for resettlement area fixed-effects.

Figure A4: Survey Locations of Waris et al. (1952)



Source: Waris et al. (1952), Figure 17.

Figure A5: Return Migration and the Destruction of the Housing Stock



Y-axis: Share of the pre-war population who had returned by January 1st, 1944. Drafted men are included in the denominator, but not in the numerator. X-axis: The share of existing housing stock destroyed by December 31st, 1941. Source: Waris et al. (1952, Appendix Tables 7 and 9)

Table A1: Consumption expenditure and gross income

	(1)	(2)	(3)	(4)
Constant	8.0518	4.3229	8.5764	11.6125
	(0.5795)	(1.1665)	(2.2106)	(3.9673)
Farmer	0.8223	2.4083	−2.3713	−10.1319
	(0.9718)	(1.9341)	(3.6753)	(6.9078)
Farmer×gross income	−0.0291	−0.0922	0.4298	1.7449
	(0.0382)	(0.1421)	(0.4134)	(1.0932)
Farmer×gross income^2/10		0.0020	−0.1610	−0.8899
		(0.0234)	(0.1403)	(0.5929)
Farmer×gross income^3/100			0.0147	0.1736
			(0.0143)	(0.1300)
Farmer×gross income^4/1000				−0.0117
				(0.0098)
Gross income	0.4385	0.6917	0.2398	−0.2068
	(0.0189)	(0.0714)	(0.2120)	(0.5290)
Gross income^2/10		−0.0361	0.1007	0.3138
		(0.0098)	(0.0612)	(0.2392)
Gross income^3/100			−0.0119	−0.0518
			(0.0053)	(0.0436)
Gross income^4/1000				0.0025
				(0.0027)
Observations	1,144	1,144	1,144	1,144
Adjusted R^2	0.3913	0.3987	0.4004	0.4003

Note: Coefficients and standard errors (in parantheses) from regressing consumption expenditure on gross income (both measured in 1000s of *markka*) using data from the 1971 Household Budget Survey.

Table A2: Pre-War Characteristics of the Rural Population: Agricultural men

	Means		Differences					
			Baseline		Spatial RD		Within Resettlement Area	
	Non-disp. (1)	Disp. (2)	Diff. (3)	se. (4)	Diff (5)	se. (6)	Diff (7)	se. (8)
A: Demographics								
Age	22.8	22.8	-0.01	(0.14)	0.24	(0.50)	-0.15	(0.19)
Swedish-speaker	0.07	0.00	-0.07	(0.01)	0.01	(0.01)	-0.02	(0.01)
Migrated prior to 1939	0.17	0.15	-0.02	(0.02)	0.02	(0.13)	-0.08	(0.03)
Orthodox	0.00	0.12	0.12	(0.06)	-0.11	(0.06)	0.08	(0.03)
B: Socioeconomic status								
Entrepeneur	0.30	0.42	0.12	(0.02)	0.14	(0.07)	0.14	(0.02)
White-collar	0.02	0.02	-0.01	(0.00)	-0.01	(0.02)	0.00	(0.01)
Blue-collar	0.36	0.16	-0.20	(0.02)	-0.04	(0.06)	-0.25	(0.02)
Assisting family member	0.32	0.41	0.09	(0.02)	-0.04	(0.07)	0.11	(0.02)
C: Characteristics of the municipality of residence								
Average taxable income	1.41	1.38	-0.02	(0.12)	-0.91	(1.11)	-0.56	(0.13)
Agricultural LFS	0.83	0.81	-0.02	(0.03)	-0.01	(0.15)	0.02	(0.03)
Latitude	69.4	67.7	-1.63	(0.15)	-0.47	(0.82)	0.00	(0.00)

Note: Pre-war differences between displaced and non-displaced rural population. Columns 1–2 report averages, columns 3–4 baseline differences, columns 5–6 control for the distance of the 1939 residence municipality to the post-war border (and its interaction with future displacement status), and columns 7–8 condition on resettlement area fixed-effects.

Table A3: Pre-War Characteristics of the Rural Population: Other rural men

	Means		Differences					
			Baseline		Spatial RD		Within Resettlement Area	
	Disp. (1)	Non-disp. (2)	Diff. (3)	se. (4)	Diff (5)	se. (6)	Diff (7)	se. (8)
A: Demographics								
Age	22.9	22.8	-0.15	(0.16)	-0.96	(0.53)	-0.22	(0.19)
Swedish-speaker	0.08	0.00	-0.08	(0.02)	0.01	(0.01)	-0.07	(0.03)
Migrated prior to 1939	0.38	0.36	-0.02	(0.03)	-0.28	(0.15)	-0.06	(0.03)
Orthodox	0.00	0.11	0.10	(0.04)	-0.05	(0.06)	0.08	(0.02)
C: Socioeconomic status								
Entrepeneur	0.08	0.06	-0.01	(0.01)	0.02	(0.02)	-0.01	(0.01)
White-collar	0.10	0.13	0.03	(0.01)	0.09	(0.04)	0.03	(0.01)
Blue-collar	0.56	0.53	-0.04	(0.02)	-0.29	(0.09)	-0.09	(0.03)
Assisting family member	0.26	0.28	0.02	(0.02)	0.10	(0.07)	0.07	(0.02)
D: Sector of employment								
Manufacturing	0.28	0.20	-0.08	(0.02)	-0.30	(0.09)	-0.13	(0.02)
Construction	0.14	0.14	-0.01	(0.01)	0.01	(0.03)	-0.02	(0.01)
Services	0.24	0.32	0.08	(0.02)	0.22	(0.06)	0.09	(0.02)
E: Characteristics of the municipality of residence								
Average taxable income	1.95	1.65	-0.30	(0.18)	-1.33	(0.58)	-0.88	(0.17)
Agricultural LFS	0.74	0.76	0.03	(0.03)	0.03	(0.12)	0.09	(0.03)
Latitude	68.9	67.7	-1.18	(0.17)	-0.38	(0.72)	0.00	(0.00)

Note: Pre-war differences between displaced and non-displaced rural population. Columns 1–2 report averages, columns 3–4 baseline differences, columns 5–6 control for the distance of the 1939 residence municipality to the post-war border (and its interaction with future displacement status), and columns 7–8 condition on resettlement area fixed-effects.

Table A4: Pre-War Characteristics of the Rural Population: Agricultural women

	Means		Differences					
			Baseline		Spatial RD		Within Resettlement Area	
	Non-disp. (1)	Disp. (2)	Diff. (3)	se. (4)	Diff (5)	se. (6)	Diff (7)	se. (8)
A: Demographics								
Age	24.2	24.7	0.53	(0.19)	-0.36	(0.48)	0.38	(0.22)
Swedish-speaker	0.07	0.00	-0.07	(0.02)	0.01	(0.01)	-0.01	(0.01)
Migrated prior to 1939	0.28	0.24	-0.04	(0.03)	-0.02	(0.11)	-0.10	(0.03)
Orthodox	0.00	0.13	0.12	(0.06)	-0.01	(0.06)	0.08	(0.03)
B: Socioeconomic status								
Entrepeneur	0.04	0.03	-0.01	(0.01)	0.01	(0.02)	0.00	(0.01)
White-collar	0.00	0.00	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
Blue-collar	0.23	0.13	-0.10	(0.02)	0.08	(0.05)	-0.19	(0.03)
Assisting family member	0.73	0.84	0.11	(0.02)	-0.09	(0.07)	0.19	(0.03)
C: Characteristics of the municipality of residence								
Average taxable income	1.45	1.38	-0.07	(0.11)	-1.04	(0.85)	-0.57	(0.13)
Agricultural LFS	0.83	0.80	0.01	(0.01)	-0.06	(0.04)	-0.01	(0.01)
Latitude	69.3	67.7	-1.51	(0.15)	0.37	(0.48)	0.00	(0.00)

Note: Pre-war differences between displaced and non-displaced rural population. Columns 1–2 report averages, columns 3–4 baseline differences, columns 5–6 control for the distance of the 1939 residence municipality to the post-war border (and its interaction with future displacement status), and columns 7–8 condition on resettlement area fixed-effects.

Table A5: Pre-War Characteristics of the Rural Population: Other rural women

	Means		Differences					
			Baseline		Spatial RD		Within Resettlement Area	
	Disp. (1)	Non-disp. (2)	Diff. (3)	se. (4)	Diff (5)	se. (6)	Diff (7)	se. (8)
A: Demographics								
Age	22.7	22.5	-0.21	(0.16)	-0.50	(0.91)	-0.63	(0.18)
Swedish-speaker	0.07	0.00	-0.06	(0.01)	0.00	(0.00)	-0.05	(0.02)
Migrated prior to 1939	0.39	0.35	-0.05	(0.03)	-0.17	(0.15)	-0.10	(0.03)
Orthodox	0.00	0.12	0.12	(0.04)	-0.08	(0.06)	0.08	(0.02)
C: Socioeconomic status								
Entrepeneur	0.02	0.02	0.00	(0.00)	-0.01	(0.01)	0.00	(0.00)
White-collar	0.10	0.10	0.00	(0.01)	-0.02	(0.03)	0.00	(0.01)
Blue-collar	0.16	0.13	-0.03	(0.01)	0.00	(0.03)	-0.07	(0.02)
Assisting family member	0.72	0.75	0.03	(0.02)	0.04	(0.04)	0.08	(0.02)
D: Sector of employment								
Manufacturing	0.08	0.06	-0.02	(0.01)	-0.06	(0.03)	-0.05	(0.02)
Construction	0.00	0.00	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
Services	0.18	0.18	-0.01	(0.01)	0.04	(0.03)	-0.03	(0.01)
E: Characteristics of the municipality of residence								
Average taxable income	1.81	1.58	-0.23	(0.16)	-1.80	(0.91)	-0.85	(0.16)
Agricultural LFS	0.76	0.78	0.02	(0.03)	0.01	(0.10)	0.09	(0.03)
Latitude	69.0	67.8	-1.24	(0.16)	-0.49	(0.74)	0.00	(0.00)

Note: Pre-war differences between displaced and non-displaced rural population. Columns 1–2 report averages, columns 3–4 baseline differences, columns 5–6 control for the distance of the 1939 residence municipality to the post-war border (and its interaction with future displacement status), and columns 7–8 condition on resettlement area fixed-effects.

Table A6: Impact of Forced Migration on Annual Income in 1971 Inflated by Local Prices

	Cont. Means	Baseline		Oster's Bound	Spatial RD		Resettlement Area FEs	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				A: Men by 1939 status				
Agri-cultural	10.9	2.00	1.91	1.87	2.89	2.86	1.16	1.52
		(0.39)	(0.38)	(0.38)	(0.94)	(0.82)	(0.35)	(0.49)
Other rural	16.5	0.94	1.58	1.91	1.08	0.44	0.61	1.91
		(0.59)	(0.70)	(0.61)	(2.49)	(2.77)	(0.78)	(1.16)
Urban	23.3	-2.41	-4.19	-4.81	.	.	.	.
		(1.70)	(1.10)	(0.83)				
				B: Women by 1939 status				
Agri-cultural	1.9	0.57	0.68	0.74	0.78	0.65	0.45	0.66
		(0.16)	(0.19)	(0.22)	(0.44)	(0.39)	(0.15)	(0.26)
Other rural	4.9	1.03	1.37	1.53	1.60	2.11	0.78	1.42
		(0.19)	(0.22)	(0.26)	(0.70)	(0.60)	(0.20)	(0.29)
Urban	8.6	-0.57	-1.12	-1.32	.	.	.	.
		(0.64)	(0.55)	(0.35)				
Controlling for:								
Pre-war char.		no	yes	.	no	yes	no	yes
Resettlement area		no	no	.	no	no	yes	yes

Note: This table reports estimates similar to those in Table 2 of the main paper, but using income in 1971 scaled by local price index (Statistics Finland, 1972) as the outcome variable.

Table A7: Impact of Forced Migration by Year of Birth

	Income in 1971			Non-agriculture, 1970			Secondary degree, 1970		
	Agri-cultural	Other rural	Urban	Agri-cultural	Other rural	Urban	Agri-cultural	Other rural	Urban
	(1)	(2)	(3)	(6)	(7)	(8)	(9)	(10)	(11)
					A: Men				
Displaced	1.40	0.92	-4.02	0.11	-0.01	-0.11	0.01	0.02	-0.02
(born 1907–12)	(0.42)	(0.77)	(1.13)	(0.02)	(0.03)	(0.01)	(0.02)	(0.02)	(0.02)
Displaced ×	1.79	2.32	-3.64	0.18	0.03	0.01	0.04	0.06	0.04
born 1913–18	(0.58)	(0.95)	(1.59)	(0.03)	(0.03)	(0.01)	(0.02)	(0.02)	(0.02)
Displaced ×	2.74	2.14	-5.94	0.16	0.05	-0.04	0.04	0.09	0.01
born 1919–24	(0.58)	(0.98)	(1.52)	(0.03)	(0.02)	(0.01)	(0.02)	(0.02)	(0.03)
					B: Women				
Displaced	0.15	0.81	-0.80	0.03	0.07	-0.03	-0.01	-0.01	0.02
(born 1907–12)	(0.19)	(0.33)	(0.70)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)
Displaced ×	0.85	2.08	-1.86	0.04	0.14	-0.07	-0.03	0.03	-0.04
born 1913–18	(0.33)	(0.31)	(0.71)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Displaced ×	1.26	1.50	-1.25	0.10	0.08	-0.05	0.07	0.05	0.01
born 1919–24	(0.42)	(0.28)	(0.63)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)

Note: Point estimates and standard errors (in parentheses) from regressing annual income in 1971 in thousands of 2010€ (columns 1–3), an indicator for working outside of agriculture (columns 4–6) and an indicator for holding a secondary degree in 1970 (columns 7–9) on an indicator for displacement status, it's interaction with year of birth group and observable characteristics measured before the war (year of birth dummies, indicators for speaking Swedish as one's mother tongue, member of the Orthodox church, living outside of one's municipality of birth, sector of employment, socioeconomic status, quintile dummies for residence municipality's taxable income per capita, labor force share in agriculture and latitude).

Table A8: Industry Mix in Non-Agriculture, 1970

	Agricultural		Other rural		Urban	
	Dis-placed (1)	Non-disp. (2)	Dis-placed (3)	Non-disp. (4)	Dis-placed (5)	Non-disp. (6)
A: Men						
Manufacturing; Mining and quarrying	34.4	31.7	29.0	35.1	28.5	34.6
Construction; Electricity, gas and water	29.3	32.5	23.6	22.4	16.5	14.6
Trade, restaurants and hotels	6.7	8.2	9.8	9.4	16.5	13.4
Transport, storage and comm.; Finance, insurance, real estate and bus. services	14.4	14.1	19.7	16.4	16.3	16.7
Community, social and personal services	15.3	13.6	18.0	16.8	22.2	20.6
B: Women						
Manufacturing; Mining and quarrying	43.3	31.2	30.7	28.0	21.9	28.9
Construction; Electricity, gas and water	3.2	3.1	2.1	2.6	2.5	1.9
Trade, restaurants and hotels	19.7	20.1	24.3	25.5	28.2	28.9
Transport, storage and comm.; Finance, insurance, real estate and bus. services	3.2	8.2	7.6	9.3	9.4	10.2
Community, social and personal services	30.6	37.3	35.3	34.7	38.0	30.1

Note: 1-digit industry shares among those working outside of agriculture in 1970.

Table A9: Farms size distributions in the 1930 and 1941 Agricultural Censuses

	Hectares of agricultural land											
	Likely part-time farms				Likely full-time farms							
	.25-.5	.25-1	1-2	2-3	3-5	5-10	10-15	15-25	25-50	50-100	>100	Total
A: Ceded area, 1930												
# farms	928	1,785	3,885	3,932	6,581	8,376	3,078	1,404	354	64	28	30,415
share (all)	0.03	0.06	0.13	0.13	0.22	0.28	0.10	0.05	0.01	0.00	0.00	1.00
share (likely full-time)					0.33	0.42	0.15	0.07	0.02	0.00	0.00	1.00
B: Partly-ceded area, 1930												
# farms	1,040	1,874	3,416	2,477	3,319	4,430	1,677	879	272	54	14	19,452
share (all)	0.05	0.10	0.18	0.13	0.17	0.23	0.09	0.05	0.01	0.00	0.00	1.00
share (likely full-time)					0.31	0.42	0.16	0.08	0.03	0.01	0.00	1.00
C: Rest of Finland, 1930												
# farms	11,740	20,044	32,364	25,573	36,223	49,300	23,978	20,417	11,599	2,741	790	234,769
share (all)	0.05	0.09	0.14	0.11	0.15	0.21	0.10	0.09	0.05	0.01	0.00	1.00
share (likely full-time)					0.25	0.34	0.17	0.14	0.08	0.02	0.01	1.00
D: Total, 1930												
# farms	13,708	23,703	39,665	31,982	46,123	62,106	28,733	22,700	12,225	2,859	832	284,636
share (all)	0.05	0.08	0.14	0.11	0.16	0.22	0.10	0.08	0.04	0.01	0.00	1.00
share (likely full-time)					0.26	0.35	0.16	0.13	0.07	0.02	0.00	1.00
E: Ceded area, 1941												
# farms	72	108	252	56	82	80	18	6	1	0	0	675
share (all)	0.11	0.16	0.37	0.08	0.12	0.12	0.03	0.01	0.00	0.00	0.00	1.00
share (likely full-time)					0.44	0.43	0.10	0.03	0.01	0.00	0.00	1.00
F: Partly-ceded area, 1941												
# farms	758	1,500	2,899	2,212	2,854	3,639	1,412	739	208	40	15	16,276
share (all)	0.05	0.09	0.18	0.14	0.18	0.22	0.09	0.05	0.01	0.00	0.00	1.00
share (likely full-time)					0.32	0.41	0.16	0.08	0.02	0.00	0.00	1.00
G: Rest of Finland, 1941												
# farms	14,360	22,400	33,259	26,140	40,318	60,155	30,269	23,722	12,108	2,480	724	265,935
share (all)	0.05	0.08	0.13	0.10	0.15	0.23	0.11	0.09	0.05	0.01	0.00	1.00
share (likely full-time)					0.24	0.35	0.18	0.14	0.07	0.01	0.00	1.00
H: Total, 1941												
# farms	15,190	24,008	36,410	28,408	43,254	63,874	31,699	24,467	12,317	2,520	739	282,886
share (all)	0.05	0.08	0.13	0.10	0.15	0.23	0.11	0.09	0.04	0.01	0.00	1.00
share (likely full-time)					0.24	0.36	0.18	0.14	0.07	0.01	0.00	1.00

Sources: Agricultural Censuses 1930 and 1941. See section "Additional empirical results" for discussion.

Table A10: Impact of Forced Migration by Pre-War Farm Size

	Income in 1971		Non-agri-culture, 1970		Secondary degree, 1970	
	Agri-cultural	Other rural	Agri-cultural	Other rural	Agri-cultural	Other rural
	(1)	(2)	(3)	(4)	(5)	(6)
			A: Men			
Displaced	2.47	3.29	0.15	0.08	0.04	0.06
	(0.60)	(1.02)	(0.04)	(0.03)	(0.02)	(0.02)
Share of large farms in the pre-war municipality	5.73	2.94	0.13	0.15	0.03	0.02
	(0.95)	(1.87)	(0.05)	(0.06)	(0.03)	(0.04)
Displaced × share large farms	2.53	-11.39	0.11	-0.39	-0.04	0.08
	(4.43)	(9.27)	(0.31)	(0.25)	(0.14)	(0.20)
			B: Women			
Displaced	0.71	1.09	0.08	0.08	-0.03	0.03
	(0.30)	(0.41)	(0.04)	(0.03)	(0.03)	(0.01)
Share of large farms in the pre-war municipality	0.79	-0.26	0.07	0.01	-0.04	0.05
	(0.50)	(0.59)	(0.04)	(0.04)	(0.04)	(0.03)
Displaced × share large farms	0.53	3.68	-0.20	0.22	0.28	0.10
	(2.24)	(2.93)	(0.25)	(0.23)	(0.25)	(0.12)

Note: Point estimates and standard errors (in parentheses) from regressing annual income in 1971 in thousands of 2010€ (columns 1–3), an indicator for working outside of agriculture (columns 4–6) or an indicator for holding a secondary degree in 1970 (columns 7–9) on an indicator for displacement status, share of farms larger than 15 hectares in 1930 in the person's 1939 residence municipality and their interaction. Controlling for observable characteristics measured before the war (year of birth dummies, indicators for speaking Swedish as one's mother tongue, member of the Orthodox church, living outside of one's municipality of birth, sector of employment, socioeconomic status, quintile dummies for residence municipality's taxable income per capita, labor force share in agriculture and latitude).

Table A11: Average Yields Relative to National Average by Region in 1930

	Region								
	Viipuri	Turku ja Pori	Häme	Uusimaa	Kuopio	Mikkeli	Vaasa	Oulu	Weight
Oats	1.08	1.02	0.98	1.10	1.04	1.01	0.84	0.83	0.49
Rye	0.95	0.99	1.07	1.04	1.11	1.08	0.93	0.85	0.26
Barkley	1.04	1.01	0.97	1.12	1.08	1.03	0.94	0.97	0.09
Potato	1.00	0.98	0.98	0.89	1.09	0.92	1.06	1.01	0.08
Root crops	0.95	1.13	1.05	1.02	1.00	0.96	0.93	0.69	0.02
Spring wheat	0.91	1.07	0.92	1.00	0.99	1.02	0.79	0.78	0.01
Green fodder	1.06	0.97	1.02	1.08	0.95	0.93	0.92	0.99	0.01
Hay seed	1.19	1.06	1.05	1.03	1.26	1.15	0.89	0.81	0.01
Mixed grain	1.09	0.99	0.89	1.10	1.03	1.01	0.95	0.87	0.01
Winter wheat	0.90	1.04	0.95	0.97	0.83	0.73	0.79	0.59	0.00
Peas	0.97	1.06	0.78	0.96	1.07	0.99	0.77	0.69	0.00
Flax and hemp	1.02	1.04	0.97	0.98	1.26	1.13	0.86	0.99	0.00
Yield index	1.03	1.01	1.00	1.06	1.07	1.02	0.89	0.86	
Share of new farms	0.07	0.26	0.23	0.14	0.12	0.07	0.09	0.04	

Sources: 1930 Agricultural Census (Yleinen maataloustiedustelu vv. 1929–30, Osa 1, pages 12 and 15) and Resettlement Statistics (Asutustilastoa, Asutustoiminta 1948–1950, Appendix Table 17). The weights refer to the share of agricultural land used for each crop (excluding hay) in Viipuri region in 1930. See section "Additional empirical results" for discussion.

Table A12: Impact of Forced Migration by Religion

	Income in 1971			Non-agriculture, 1970			Secondary degree, 1970		
	Agri-cultural	Other rural	Urban	Agri-cultural	Other rural	Urban	Agri-cultural	Other rural	Urban
	(1)	(2)	(3)	(6)	(7)	(8)	(9)	(10)	(11)
				A: Men					
Displaced	2.06	1.66	-4.51	0.15	0.02	-0.05	0.03	0.05	0.01
	(0.37)	(0.69)	(1.37)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)
Member of the	-1.84	-4.38	-4.81	-0.06	-0.21	-0.01	-0.04	-0.12	-0.08
Orthodox church	(1.06)	(1.54)	(2.08)	(0.06)	(0.06)	(0.04)	(0.02)	(0.04)	(0.03)
Displaced ×	-0.27	4.75	-0.25	0.01	0.20	0.01	0.02	0.17	0.02
Orthodox	(1.15)	(2.30)	(2.44)	(0.06)	(0.08)	(0.04)	(0.03)	(0.05)	(0.04)
				B: Women					
Displaced	0.69	1.54	-1.27	0.06	0.10	-0.05	0.01	0.03	0.00
	(0.19)	(0.22)	(0.63)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
Member of the	-0.08	1.51	0.22	0.13	0.06	0.07	0.13	-0.03	-0.01
Orthodox church	(0.96)	(1.45)	(0.50)	(0.08)	(0.05)	(0.02)	(0.09)	(0.03)	(0.03)
Displaced ×	-0.98	-2.08	-0.74	-0.22	-0.08	-0.07	-0.17	-0.01	0.03
Orthodox	(1.01)	(1.48)	(1.62)	(0.09)	(0.06)	(0.05)	(0.09)	(0.04)	(0.04)

Note: Point estimates and standard errors (in parentheses) from regressing annual income in 1971 in thousands of 2010€ (columns 1–3), an indicator for working outside of agriculture (columns 4–6) or an indicator for holding a secondary degree in 1970 (columns 7–9) on an indicator for displacement status, being a member of the Orthodox church and their interaction. Controlling for observable characteristics measured before the war (year of birth dummies, indicators for speaking Swedish as one's mother tongue, living outside of one's municipality of birth, sector of employment, socioeconomic status, quintile dummies for residence municipality's taxable income per capita, labor force share in agriculture and latitude).