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

Matti Sarvimäki Roope Uusitalo Markus Jäntti

January 2016
(preliminary and incomplete)

Abstract

We examine the impacts of resettling 11% of the Finnish population after WWII. Farmers were given land and assistance to continue farming in areas resembling the origin regions. Nevertheless, a quarter of a century later, they were 10–16 percentage points more likely to hold a non-agricultural job and earned 11–28% more than plausible control groups. These effects suggest that farmers could increase their long-term income by 70–80% by leaving agriculture. Yet, almost three quarters of the non-displaced farmers stayed in their farms. We rationalize these findings with a simple habit formation model and provide further evidence supporting it.

JEL classification: O15, R23, J61
Keywords: Rural-urban gap, misallocation, migration

*This paper supersedes "The Long-Term Effects of Forced Migration". We thank Abhijit Banerjee, David Card, Kristiina Huttunen, Kaivan Munshi, Tuomas Pekkarinen, Daniel Sturm, 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 and VATT, matti.sarvimaki@aalto.fi. *Uusitalo*: University of Jyväskylä, roope.uusitalo@helsinki.fi. *Jäntti*: University of Helsinki and VATT, markus.jantti@iki.fi.

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 workers could double their income by switching to the modern sector, why do they stay in agriculture?

This paper examines transitions from agriculture to non-agriculture in the mid-20th century Finland—a poor, agrarian society at the time. We present empirical evidence suggesting that many farmers could have earned substantially more in the modern sector, but chose to forgo these opportunities because they were attached to their farms. This behavior can be formalized as habit formation, where the utility of living in a location increases with the time already lived in that location. Our results suggest that such habit formation is a quantitatively important factor affecting the allocation of labor across sectors and locations.

Any empirical investigation of habit formation faces the challenge that people who differ in their personal histories may also differ along other (unobserved) dimensions. In our context, farmers may choose to remain in agriculture because they love their farms, because their comparative advantage is in agriculture, because they lack networks or formal insurance, or because of a number of other plausible reasons. It is hard to distinguish between these possibilities using observational data. However, experimental manipulation of personal histories is typically infeasible.

We make progress by examining the long-term impact of forced migrations. World War II led Finland to cede its eastern parts to the Soviet Union and to resettle the entire population living in those areas to the remaining parts of the country. In total, 11% of the population were forced to migrate. For those working in agriculture—roughly one half of the population—the government attempted to reconstruct the pre-war situation 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 so-

cial networks. Once the resettlement was completed in 1948, however, 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 by estimating the impact of forced migration on long-term earnings 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 reports the distance from their pre-war municipality of residence to the post-war border. The vertical axis presents their annual income in 1971. The figure shows that a quarter of a century after being forced to migrate, displaced farmers earned substantially more than other men who worked in agriculture before the war.

The post-war differences between displaced and non-displaced farmers is likely to be caused by the resettlement. Selection into forced migration was determined by the 18th century Swedish-Russian borders (that were used as a reference point in the peace negotiations) and the unpredictable military success of the Finnish troops (that determined which of the many potential historical borders was used). The entire population living in the ceded area was evacuated in an orderly manner. Furthermore, there are no important 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 earnings by 11–28% among men working in agriculture before the war.

We next examine potential mechanisms behind the positive effect on income. The bottom panel of Figure 1 plots the share of men moving from agriculture to nonagriculture between 1939 and 1970 on their pre-war distance to the post-war border. Estimates from alternative specifications suggest that forced migration increased the likelihood of leaving agriculture by 10–16 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 rural-urban migration and that the impact for income and mobility closely mirror each other when we extend the analysis to other groups. Being displaced had no impact on the income or sector among rural men who worked out-

side of agriculture before the war, but decreased income and increased the likelihood of moving to rural locations among the urban population. Furthermore, the average income of displaced persons was similar in the 1970s as that of the non-displaced persons working in the same industries and living in the same locations.

Taken together, our results suggest that the impact of forced migration on income can be attributed to increased likelihood of leaving agriculture. Under the assumption that being displaced affected income only through this channel, our estimates imply that farmers could have increased their income by 70–80% by moving to the modern sector. We recognize that such an instrumental-variables interpretation may be too strong, because forced migration may have affected income also through other channels. We note, however, that the IV estimates are in line with OLS estimates using our data and with previous work documenting rural-urban (or agricultural-nonagricultural) wage differences in the 1940s United States (Caselli and Coleman, 2001) and in contemporary developing countries (Young, 2013; Gollin et al., 2014). Thus, we conclude that returns to leaving agriculture were likely high in the context we study.

If farmers could substantially increase their income by leaving agriculture, why did most of them decide to stay on their farms? Much theoretical work suggesting answers to this question has focused on the riskiness of urban labor markets (Harris and Todaro, 1970), local prices and amenities (Rosen, 1979; Roback, 1982) and sectoral differences in 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.[1]

We argue that a model of habit formation following Pollak (1970) and Becker and Murphy (1988) provides a compelling framework for rationalizing our results. In these models, contemporaneous utility derived from the consumption of a good (or activity) depends on the past consumption of this good. In our version of the same idea, the disutility of working in a task–location pair decreases with the time that the person has already

[1]Of course, our findings do not falsify these models, but merely suggest that they abstract away from mechanisms that are quantitatively important.

worked in this task or lived in this location. That is, farmers "consume" their farms and accumulate "consumption capital" that increases their period utility from this consumption over time. When they are forced to move to an otherwise similar farm, they lose the consumption capital tied to their former homes. Thus, they require a smaller compensating income differential for leaving agriculture, and many of them find it optimal to move to the modern sector.

We present several complementary pieces of evidence supporting the habit formation hypothesis. First, the impact of forced migration differs by previous migration history and by age in a way that is consistent with a habit formation model. Second, we review the results from large surveys conducted at the turn of 1950s, where the vast majority of the displaced persons expressed a strong desire to return to their former homes (Waris et al., 1952). Third, an most importantly, their revealed preferences are in line with they survey responses. A year after the displaced farmers had received their new farms, Finland reoccupied the ceded areas. Despite much destruction in the reoccupied areas and the genuine opportunity to remain on their new farms, 97% 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 ceded to the Soviet Union for the second and final time.

We end by considering alternative explanations. One possibility is that forced migration created new, or destroyed old, social networks in a way that facilitated transition to the modern sector (e.g. Banerjee and Newman, 1998; Karlan et al., 2009). We examine this network hypothesis using variation created by the initial evacuations and the size of the resettlement areas, but fail to find evidence supporting it. We also discuss how inertia, information updating, and quality differences between the old and new farms are consistent with some of our results, but inconsistent with others. Thus, only the habit-formation hypothesis seems to provide a unified explanation for the full pattern of our empirical findings.

This paper adds to the large body of work examining the possibility that the misallocation of labor across agriculture and other sectors constitutes

a major obstacle for development. The misallocation hypothesis goes back to at least Lewis (1955) and has recently been discussed by Caselli (2005); Restuccia et al. (2008); Vollrath (2009); Young (2013); Gollin et al. (2014); Munshi and Rosenzweig (2016).[2] However, empirical work examining the mechanisms behind the apparent misallocation of labor remains scarce.

Previous work closest to us include Bryan et al. (2014), who present experimental evidence showing large returns to temporary rural-urban migration in Bangladesh, and Munshi and Rosenzweig (2016), who focus on the role of informal insurance in preventing rural-urban migration in India. Furthermore, Galor et al. (2013) discuss a model where initially efficiency-enhancing inclusive institutions designed to restrict mobility can lead to misallocation of talent over time. They also present empirical evidence from the American Midwest to support this hypothesis. In comparison, we seem to be the first to present empirical evidence on the importance of habit formation in affecting the allocation of labor across agricultural and nonagricultural sectors and to use forced migration to examine the misallocation hypothesis.[3]

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. One half of the population was working in agriculture, typically owning small farms and working as hired labor in forest work during the winter. According to Maddison (2010), GDP per capita in 1938 was comparable to that in current North Africa.

Three months after the outbreak of World War II, in November 1939, the Red Army crossed the Finnish-USSR border, starting the Winter War.

[2] A related literature documenting productivity differences across firms includes, among others, Restuccia and Rogerson (2008); Hsieh and Klenow (2009).

[3] Previous work examining the impacts of forced migration include Kondylis (2008, 2010) and Bauer et al. (2013). However, none of these papers discuss the implications of their results for the misallocation hypothesis.

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 shelter. In the peace treaty ending the hostilities in March 1940, Finland ceded roughly a tenth of its territory to the Soviet Union and evacuated the remaining population from the ceded areas that had stayed under Finnish control during the war.

In July 1940, the Finnish Parliament enacted an Emergency Settlement Act (*Pika-asutuslaki*) guiding the resettlement policy, which is discussed in detail below. 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 the ceded areas. As we discuss in more detail in Section 6.4, 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 of 1944. The Finnish troops were pushed towards the west, but managed to stop the Red Army in early July. 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.

2.2 Post-War Borders

Figure 2 presents the pre-war and the post-war borders and the 1945 resettlement plan. As we explain in detail in Section 4, we will use those living west of the new border before the war as a control group for those living in the ceded areas. The validity of this approach would be called into question if the new border had been determined by regional economic differences. However, instead of economic considerations, historical borders were used as a reference point in the negotiations. The 1944 border closely followed the border set in the treaty of Nystad in 1721. Thus, the location of the 1721 border determined whether the persons living in Eastern Finland were

displaced or not.[4]

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. Rentola (2001) discusses archive material indicating that the US initially planned to propose to their Soviet allies a peace treaty with Finland based on the 1920 borders (roughly one hundred kilometers east of the current border). When the Soviet Union offered peace talks in March 1944, it was preparing to negotiate on the basis of 1743 borders (roughly sixty kilometers west of the current border). However, when the peace talks began in August 1944, the unexpected success of the Finnish troops—and the need to reallocate Soviet troops to the more important Baltic front—had improved Finland's position in the negotiations and thus moderated the Soviet demands.

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 Dutchy of Finland, which declared independence in 1917 and became the pre-WWII Finland. Below we will also show that average pre-war characteristics were roughly similar before the war on both sides of the post-war border.

2.3 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 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 (Tarkka, 1988).

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

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 compensations of only ten percent for very large ones. Those who had owned or rented land in the ceded areas were given agricultural and forest land. Those who had lost other kinds of property received their compensations 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 capital. Land for the settlers was first taken from the state, the local authorities (municipalities) and the church, but these institutions did not own a sufficient amount of land. Thus, roughly two thirds of the cultivated fields, one half of the land that could be cleared for cultivation and a third of forest land were seized from private owners. Land was seized using an explicit progressive expropriation schedule. Similarly, a heavy progressive tax was set on other forms of property.[5]

The implementation of the resettlement was entrusted to the Department of Land Settlement in the Ministry of Agriculture. The aim was to match the pre-war conditions as closely as possible. In order to preserve social connections, each ceded village was settled together to a designated target area. Furthemore, 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, where conditions for agriculture are unfavorable and very few were allocated to the Swedish-speaking municipalities on the western and southern coasts.

The non-agrarian population was free to settle wherever they could find

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

accommodation. Cities in Eastern Finland received flows of displaced persons that constituted almost ten percent of the population, while cities further west and cities with the most severe housing problems received much less. While those who had not owned land were not explicitly allocated, the settlement plan appears to have influenced their migration also due to family ties and employment opportunities with their former landowner 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 and no further policies targetting the displaced population were introduced. That is, the displaced and non-displaced population had equal legal status form 1948 onwards. In particular, everyone could sell and buy land and migrate everywhere in the country.

3 Data

3.1 Data Sources

The starting point of our data is the 1950 population census. Data were collected by personal interviews and the information on each dwelling unit was stored on a single form. These forms were sorted by municipality, within municipalities in alphabetical order and then filed in folders. Statistics Finland has drawn a sample from the full 1950 census by picking every tenth folder. Nearly all of the information on the census forms was manually inserted into a database. The resulting sample contains information about 114,000 dwelling units with 411,629 persons from 392 of municipalities (out of a total of 547 municipalities). These data are then linked to the 1970 census and the 1971 tax records.[6] The same adminstrative registers and survey instruments are used for displaced and non-displaced persons and hence all the information is fully comparable between the two groups.

The 1950 and 1970 censuses contain information on various household

[6]The link was done first to the Population Register using names, date of birth, place of birth and other characteristics. The Population Register includes social security numbers that can be used to link data to other sources. Thus a person had to be alive and be living in Finland in the mid-1960s when social security numbers were introduced in order to be included in our data. Statistics Finland (1998) provides details.

characteristics and person-level information such as place of birth and residence, education, occupation and sector of employment. Importantly, the 1950 census also included retrospective questions referring to September 1st, 1939—two months before the war began. This allows us to observe individual-level pre-war municipality of residence, occupational status and industry codes. Furthermore, we augment the indvidual-level data with municipality-level information on the pre-war income distribution and industry structure collected from publicly available statistical publications.

The earliest individual-level income data available to us comes from the 1971 tax register and provide us with an accurate measure of annual earnings. Importantly, agricultural profits were treated as taxable earned income and taxed according to the same rates as wage earnings, but production for own consumption was not taxed. However, by 1970s agriculture had become increasingly specialized and, for example, Pihkala (1975) 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. Thus, the tax records are likely to provide a comparable measure of income across sectors.

3.2 Estimation Sample and Descriptive Statistics

We limit our analysis to persons born between 1907 and 1925 in order to focus on individuals who were of working age throughout the period we examine. These birth cohorts were 14–32 years old when the war started in 1939 and 46–64 years old when we observe their income in 1971. This leaves us with information on 85,836 individuals of whom 8,528 were displaced. Within these birth cohorts, we examine separately those who worked in agriculture before the war, those who lived in rural areas but were not working in agriculture and those who lived in cities in 1939. Furthermore, we split each of these categories by gender.

Table 1 reports the average pre-war characteristics of the resulting groups. 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.

Among the rural population, the displaced persons were also more likely to have graduated from primary school (*kansakoulu*). These differences are driven by a larger share of the nondisplaced rural population living in the sparsely populated areas in the northern Finland. Once we condition for latitude, the differences in educational attainment disappear (see the appendix). Furthermore the displaced rural population was less likely to work as blue-collar workers and in manufacturing and tended to live in poorer municipalities.

4 Empirical Strategies

We evaluate the impact of forced migration by comparing the outcomes of displaced persons to control groups consisting of persons who were not displaced. As we discuss below, each of these comparisions may suffer from omitted-variables bias. However, alternative approaches are likely to suffer from opposite biases and thus we can provide plausible bounds for the causal effect of forced migration.

In practice, we estimate variants of the regression equation

$$y_{it} = \alpha + \beta D_i + \mathbf{X}'_{\mathbf{0i}}\gamma + \varepsilon_{it} \tag{1}$$

where y_{it} is the outcome of interest for individual i at time t, D_i is an indicator variable taking value one if the person was living in the ceded area just before the war and zero otherwise, $\mathbf{X}'_{\mathbf{0i}}$ is a vector of observed characteristics, measured before the war, and ε_{it} captures unobserved factors. We implement the various comparisons by estimating (1) for different subsamples and by varying the contents of $\mathbf{X}'_{\mathbf{0i}}$.

4.1 Controlling for Pre-War Characteristics

We start by using everyone living in the non-ceded areas as a control group. This comparison is unlikely to yield causal estimates, because the displaced and non-displacd populations differ in their pre-war characteristics. However, these differences were 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.

The comparision of unconditional and conditional estimates can also be used to bound the likely remaining omitted variables bias (Altonji et al., 2005). We follow the approach suggested by Oster (2015) and present bounds for the estimates under the assumption that selection on unobservables is as important as selection on observables.[7]

4.2 Spatial Regression-Discontinuity Design

Our second control group consists of farmers who lived just west of the post-war border. This spatial regression-discontinity design builds on the plausibility of locally random assignment into forced migration. Everyone who happened to live east of the new border had to move, and the historical borders used as reference points were unlikely to be correlated with the characteristics of the population living in these areas. In line with this hypothesis, Appendix Tables xx-xx show that the observable pre-war characteristics evolve smoothly over the post-war border.[8]

The caveat of the spatial RD design is that those living in the control areas may have been affected by the shift of the border. This hypothesis is supported by Redding and Sturm (2008), who 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 of the impacts of forced migration.

[7]Estimation of these bounds also require us to make an assumption on R_{max}, the R-squared of a hypothetical regression conditioning on all relevant observable *and* unobservable control variables. We follow Oster's recommended rule-of-thumb, based on an analysis of true randomized designs, and set $R_{max} = \min\{1, 1.3\tilde{R}\}$, where $\tilde{R}$ is the R-squared from the regression including the control variables observed in our data.

[8]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 $\mathbf{X}'_{\mathbf{0i}}$, restrict the estimation sample to persons who lived close to the post-war border before the war, weight the observations close to the border more than those further away using a triangle-shaped kernel, and use the Imbens and Kalyanaraman (2012) algorithm for choosing the optimal bandwidth.

4.3 Within-Resettlement-Area Comparisons

Our third approach compares displaced persons to the local population of their resettlement areas. 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 along soil quality and average temperature. That is, the resettlement 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.[9]

The caveat of the within-resettlement-area comparisons is that the destination areas tended to be richer and more industrialized before the war (see the appendix). Furthermore, the resettlement shock itself may have pushed rural municipalities to industrialize faster and thus increased local wages (Sarvimäki, 2011). Below, we will argue that forced migration affected long-term income by pushing the displaced farmers into sectors and regions characterized by better economic opportunities. Using the local population of the resettlement areas as a comparison group implicitly conditions on part of these possibly favorable economic conditions. Thus, the resulting estimates can be interpretted as lower bounds of the effects of forced migration.

5 Results

5.1 Long-Term Income

Table 2 presents estimates comparing real annual income of the displaced persons to the various control groups discussed in the previous section. Each entry comes from separate regressions that differ in the population examined (rows) and the comparison groups (columns). We have scaled annual income by a municipality level local price index (Statistics Finland,

[9]We implement this comparison by including resettlement area fixed-effects in $\mathbf{X_{0i}}$ and dropping the non-displaced persons living outside of the resettlement area from the sample. These fixed-effects are constructred 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.

1972) and include observations with zero income to account for the possible effects of the displacement on employment. In order to assess the magnitudes of the estimates, we also report the mean income among the non-displaced persons. All standard errors are clustered at the 1939 residence municipality level.[10]

The first row presents results for men working in agriculture before the war. The unconditional difference is 2,680 *markka*, corresponding to 18% higher earnings among the displaced farmers. Controlling for pre-war observable characteristics reduces the point estimate to 2,249 *markka*, or 15%. Assuming that selection-on-unobservables is as important as selection-on-observables, this reduction in the point estimates—together with an increase in the R^2 from 0.004 to 0.140—suggests a lower bound of 2,114 *markka*, or 14%.

The next two columns present the estimates from the spatial RD regressions. The estimates from specifications with and without controlling for pre-war characteristics are 4,146 *markka* and 3,534 *markka*, respectively, corresponding to 24–28% increase in real earnings in 1971. As we discussed in the previous section, we interpret these estimates as upper bounds.

The remaining two columns report the differences in income between the displaced farmers and their future neighbours in the resettlement areas. In line with our hypothesized direction of bias, the estimates are smaller than those from the other specifications. We interpret them as lower bounds, suggesting that being displaced increased the long-term impact of farmers by at least 1,677 *markka* or 11%.

The remaining of Table 2 examines whether these results extend also to other groups. Interestingly, we find no statistically significant effect for rural men who were not working in agriculture before the war and a roughly 10% negative impact for urban men. The estimates for women are smaller in absolute magnitude than those for men, but larger relative to the baseline earnings of non-displaced women. The qualitative differences between

[10]This choice of the level of clustering is motivated by the assumption that persons coming from the same areas might have been exposed to common shocks, in particular because the resettlement was conducted based on pre-war residence municipality. It also accounts for the measurement error due to using municipality mid-points (rather than exact location) when measuring distance to border in the spatial RD specification (see Lee and Card, 2008 for a discussion).

men and women are that the spatial RD specification yields statistically insignificant estimates for women working in agriculture before the war, and that estimates for non-agricultural rural women are large and statistically significant.

5.2 Sectoral and Spatial Mobility

Table 3 examines to impact of forced migration on the likelihood of holding a nonagricultural job in 1970. We find exactly the same pattern as for income. The estimates suggest a 10–16 percentage points increase in the likelihood of working outside of agriculture for displaced male farmers and a 2–3 percentage points decrease for the urban population. Similar to the income results, there is no a statistically significant effect for rural men who did not work in agriculture, but the baseline and within-resettlement-area estimates for all rural women are positive and statistically significant.

In the appendix, we report similar estimates for working in a non-agricultural job already in 1950 and examine the impact of being displaced on post-war geographical mobility. The results show that the gap in non-agricultural employment was already present in 1950. Forced migration had a large positive impact on urbanization among the rural population and an equally strong negative effects among the urban population. Furthermore, forced migration increased rural-to-rural migration in the period between 1950 and 1970.

5.3 Conditioning on Post-War Sector and Location

The results discussed thus far are consistent with the hypothesis that the increased sectoral mobility, often accompanied with geographical mobility, led to higher earnings among the displaced farmers. Table 4 presents complementary evidence for the importance of mobility by comparing the real annual earnings between displaced and 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 for working outside of agriculture (columns 2 and 6), fixed effects for 1970 residence municipality and 2-digit industry (columns 3 and 7) and fixed effects for 1950 residence municipality

(columns 4 and 8). Among rural population and urban women, the point estimates for displacement status approach zero as we add further post-war control variables and become statistically insignificant. However, the estimates for urban men remain negative and significant in all specifications.

An important limitation of the results reported in table 4 is that the estimates do not have a causal interpretation. The reason is that we are now conditioning on post-war outcomes that we have already shown to be affected by forced migration. Nevertheless, the estimates can be interpreted as informative descriptive statistics showing that the long-term income of displaced persons—with the exception of urban men—did not differ from the income of non-displaced persons who worked in the same industries and locations in 1970.

5.4 Returns to Leaving Agriculture

Taken together, our results suggest that moving to the modern sector led to a high pecuniary return in mid-20th century Finland. We now ask whether our data and research design allow us to be more precise about the magnitude of these returns.

We start by presenting OLS results from regressing real annual income in 1971 on a dummy for the person working outside of agriculture in 1970. Table 5 reports the results for individuals who worked in agriculture in 1939. For reference, column 1 shows that men who still worked in agriculture in 1970 earned roughly 14,000 *markka* and women only 1,000 *markka*. The estimates reported in columns 2–3 show that men who had a nonagricultural job in 1970 had roughly 8,000 *markka* higher annual earnings than those who did not. This corresponds to 55–59% higher annual income in comparison to those working in agriculture. For women, the earnings difference is roughly 10,000 *markka*, but given the very low earnings for women outside of the modern sector, the relative difference is an order of magnitude larger.

The OLS estimates would measure the true returns of 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 comparative advantage.

We next turn to an instrumental-variables approach, where we use displacement status as an instrument for working outside of agriculture in 1970. These estimates measure returns to leaving agriculture if the impact of forced migration on long-term income was mediated entirely through the transition to the modern sector.[11] Clearly, other potential mechanisms exist. For example, the trauma of being displaced may have had a direct effect on the labor market prospects of some farmers. On the other hand, some may have responded to their partial loss of wealth by increasing their labor supply. Thus, we emphasize that the IV approach is based on stronger identifying assumptions than the results discussed in the previous sections.

Interestingly, the IV estimates reported in Table 5 present a very similar picture as the OLS estimates. Since IV approaches are informative only about a subpopulation of "compliers"—those who left agriculture because they were displaced and would have stayed in agriculture otherwise—we first report estimates of what the compliers' would have earned if they had stayed in agriculture (columns 5–6).[12] In comparison to this baseline, the 2SLS estimates suggest that leaving agriculture increased the income of men by 65–83% (columns 7–8). Again, the estimates for women are similar in levels, but much larger in comparison to the baseline income in agriculture.

6 Habit Formation and Mobility

The results reported above 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 farmers into the modern sector. As we discussed in the introduction, the models explaining rural-urban (or almost equivalently, agricultural-nonagricultural) income differences have mostly focused on mechanisms which are unlikely to explain our results. In this section,

[11] In addition, identification requires that the likelihood of working outside of agriculture does not decrease for anyone if they are forced to migrate.

[12] We estime these complier averages in the absence of treatment with 2SLS using $y_i (1 - D_i)$ as an outcome and $(1 - D_i)$ as an "treatment". See Imbens and Rubin (1997) for discussion.

we introduce an explanation based on a simple habit formation model and report further empirical results supporting it.

6.1 Predictions of a Habit Formation Model

In the appendix, we draft a variant of the Becker and Murphy (1988) model that provides a simple way to rationalize our results. It builds on the key assumptions that (i) an agent's contemporaneous disutility of working in a certain task–location pair depends on her work history, (ii) agent's productivity in a certain task depends on her task-specific innate ability and work experience, and (iii) the children of farmers have accumulated work history at the farm by the time that they make their occupational choices. Everyone is fully rational and maximizes their life-time utility taking into account that their decisions today will affect their future productivity and disutility from work.

In our model economy, workers select into agricultural and non-agricultural jobs based on their comparative advantage and initial "hours capital" accumulated through living in a certain location and working in a certain job during childhood. If individuals growing up in a farm accumulate work experience at this farm during childhood, they are more likely to remain working in that farm due to both productivity advantages (through learning-by-doing) and lower disutility from work. When farmers are forced to migrate, they will loose the "hours capital" tied to their previous locations. In comparison to otherwise identical non-displaced farmers, they are thus more likely to move to the modern sector. However, since their productivity advantage is in agriculture, many will remain in agriculture also after the forced migration.

Our model illustrates that the empirical results discussed above can be given a fully rational interpretation. However, we note that the empirical results could equally well be rationalized, for example, with a model of history dependent preferences and myopic behavior. Thus, our results are unlikely to be informative about the precise way that habit formation may emerge.

6.2 Effect Heterogeneity

We now turn to the additional testable hypotheses suggested by the model described above. We start with the prediction that, other things equal, the impact of forced migration should increase with the time that the displaced persons had lived on their farms or home towns before the war. This is because, everything else constant, a person with a larger initial stock of consumption capital requires a larger compensating differential for leaving agriculture.[13] Importantly, however, the impact of forced migration does not need to increase with age, because while the older displaced farmers lose more consumption capital (pushing them away from agriculture), they have also accumulated more experience in farming and are closer to retirement (making it less profitable for them to move to the modern sector).

Table 6 examines the heterogeneity of the impact of forced migration by age and previous migration history. It reports estimates from

$$y_{it} = \alpha + \beta D_i + \gamma Z_{i0} + \delta \left(D_i \times Z_{i0} \right) + \theta \mathbf{X}_{\mathbf{0i}}^{'} \gamma + \varepsilon_{it} \tag{2}$$

where D_i is a dummy for individual i having been forced to migrate, Z_{i0} is either a dummy for her living outside of his municipality of birth in 1939 (panel A) or her age in 1939 (panel B), and $\mathbf{X}_{\mathbf{0i}}^{'}$ is a vector of pre-war observable characteristics.

In line with the predictions of a habit formation model, we find that the impact of forced migration on long-term income was stronger for those rural residents who were still living in their municipality of birth in 1939 (with the exception of women working in agriculture). Furthemore, the effects are larger for younger individuals. Interpreting this result through the lens of our model suggests that the "pull factor" from productivity increases due to learning-by-doing was stronger than the "push factor" of losing consumption capital. In the appendix, we show that exactly the same pattern of results emerges when we examine the likelihood of working in a nonagricultural job in 1970.

[13] Conversely, displaced and non-displaced persons with zero consumption capital should have identical post-war outcomes if they have similar pre-war work experience and inate ability.

6.3 Survey Evidence

Another prediction of a habit formation model is that the displaced persons will start to accumulate a new stock of consumption capital in their new locations. However, this process is slow and there can be a long period during which they would have been willing to give up part of their income in order to return to their previous homes.

The first piece of evidence supporting these predictions comes from Waris et al. (1952), who surveyed 1,982 displaced and 1,150 non-displaced persons in 1949 and 1951.[14] The surveys included question on the future migration intentions and why they were planning to move (or to stay). Two thirds of the displaced persons stated that they expected to remain in their current location. Waris et al. (1952, p. 314) 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.*

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*". 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".

[14] The research project "The Social Adjustment of Displaced Persons in Finland" was launched in 1948 with funding from the Rockefeller Foundation. It contained two general surveys of both the displaced persons and local populations supplemented with in-depth interviews in two rural municipalities and in one industrial town. The results were published (in Finnish) in Waris et al. (1952).

6.4 Return Migration

While the survey results discussed above are in line with a model of habit formation, their limitation is that they are based on stated preferences expressed in low-stake interviews. An unique feature of this historical episode, however, 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 it for three years. Once a reconquested municipality was sufficiently secured, the displaced persons could apply for a permit to return. If they were granted the permit, they would give up the land 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 their 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 the areas were closer to the frontline. The conditions were often harsh. For example, many of those returning in 1941 endured food shortages. Furthermore, the outcome of the war was anything but certain and, indeed, those who returned to the ceded areas were forced to migrate again in 1944.

Despite all the cost and risk, return migration proved extremely popular—particularly among the farmers. Out of the 38,872 farms allocated to the displaced persons as part of the 1940 resettlement policy, 97% were returned in an exchange for the remainings of the old farms (Waris et al., 1952, p. 110). In total, more than two thirds of the displaced population returned, even though not everyone were granted a permit to return due to housing shortage, proximity of the frontline or being considered a politically unreliable person (i.e. known or suspect communists).

Figure 3 plots the share of 1939 population living in the rural reconquested municipalities in 1944 against the share of housing that had survived the battles. While there is a statistically significant association between return migration and housing is statistically significant, almost all municipalities are above the 45° line.

7 Alternative Mechanisms

Above, we have shown that our results are in line with the predictions of a habit formation model. However, they could also be consistent with alternative explanations such as the forced migration affecting social networks, learning, or the new farms being of lower quality than those left to the ceded area. In this section we argue that none of these alternative mechanisms provides a compelling unified explanation for the full pattern of our results.

7.1 Expansion of Dispersed Networks

We start with the hypothesis that forced migration affected income through creating valuable social networks. In particular, the initial evacuations could have created geographically dispersed networks that would have facilitated the flow of information about job and business opportunities.

We examine the role of the "evacuation networks" using plausibly exogeneous variation created in the way that the government solved the logistic challenge of evacuating a tenth of the population in a few weeks. The displaced population of each ceded municipality were transported into a designated evacuation area and the local population was oblidged to provide them shelter. As a consequence, most displaced persons were hosted by a local family for the winters of 1940–41 and 1944–45. During the spring and summer of 1941 and 1945, the displaced farmers received their new farms.

In the appendix, we examine the impacts of these evacuation networks on post-war outcomes. We find that they affected migration patterns in the period between 1950–1970. However, these migrations do not seem to have had any impact on long-term income or the likelihood of leaving agriculture. Thus we conclude that while spatially dispersed networks were not irrelevant in the mid-20th century Finland, they are unlikely to explain our results.

7.2 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). Loosing access to such informal arrangements would reduce migration costs and could thus explain our results.

The importance of local networks was not lost by the 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 villages. As Figure 2 illustrates, even neighboring municipalities ended up to be settled to areas that differed vastly in size.

In the appendix, we exploit the variation in the size of the resettlement areas to examine whether the destruction of local networks can explain our results. We find no association between the size of the resettlement area and post-war outcomes. 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, these findings suggest the destruction of local networks is unlikely explain why forced migration affected income and mobility.

7.3 Intertia and Learning

Another potential explanation is that being forced to migrate once may have reduced the cost of migrating again. For example, 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. This hypothesis would explain the higher mobility of the displaced persons after the war. However, it fails to explain why the vast majority of the displaced persons returned to their former homes in 1941–44.

7.4 Quality of the New Farms

The final alternative explanation we consider is that the new farms provided to the displaced farmers were smaller than the farms left to the ceded area. The local farmers may have offered their worst agricultural land to the authorities in charge of the resettlement, who may not have been able to sufficiently look after the interest of the displaced farmers.

However, the displaced farmers could have acquired more or better land relatively easily. They were entitled to get subsidized loans from the State Settlemen Fund for purchasing (or clearing) land, buildings and machines (Pihkala, 1952). Furthermore, the quality of land is relatively easy to assess and Finland has an abundance of forest land, which can be cleared into fields. Thus it is unlikely that credit constraints or asymmetric information about the quality of the land would have prevented trade in this context. Of course, some other type of market frictions could prevent trade. In particular, the habit formation hypothesis implies that the local farmers would value their land over the market price. Such frictions could be imporant, but they would only strengthen the case for the importance of habit formation. In any case, in the absence of some type of preference to working in agriculture, frictions would prevent the displaced farmers from utilizing their comparative advantage and thus forced migration would have had a negative effect on income. We find the opposite.

8 Conclusions

In this paper, we have argued that part of the apparent misallocation of labor can be explained by models of habit formation. At some level, this conclusion is obvious. Everyone has a story of someone, who chose the humble life of her home town instead of a lucrative career elsewhere. We all know—or are—people who have grown addicted to their work. However, while anecdotes are abundant, systematic evidence remains scarce.

The displacement and resettlement of 11% of the Finnish population provides an informative opportunity for examining the mechanisms behind rural–urban (or agricultural–nonagricultural) income gaps. Our results suggest that forced migration increased the likelihood of leaving agricul-

ture, which in turn led to a large increase in their long-term income. These results are consistent with a combination of high returns to leaving agriculture and high migration costs. We argue that a substantial part of these migration costs are due to farmers feeling attached to their farms, i.e. habit formation.

To be sure, habit formation alone is unlikely to explain the persistence of agricultural-nonagricultural income differences. We do not advocate a view that other mechanisms are irrelevant, but merely suggest that habit formation models provide additional insight. In particular, they illustrate that even if labor is misallocated in the sense of lost production, this does not need to imply lost welfare. In the language of Becker and Murphy (1988), farmers may be "addicted" to their farms, but this is "rational addiction" in the sense that given their initial consumption capital, farmers maximize their lifetime utility by staying in the low produtivity traditional sector.

However, the welfare implications are subtle, because people do not choose where to accumulate their initial hours capital. The choices of the parents may be particularly important in agriculture, where children typically start working at a young age and family traditions may be markedly important. Thus it is possible that while the forced migation we examine was surely a tragedy for those who had already grown attached to their homes, it may have increased the welfare of their small, or yet unborn, children by liberating them to enjoy the benefits of working in the modern sector.

References

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

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

Bauer, Thomas K, Sebastian Braun, and Michael Kvasnicka (2013) 'The

Economic Integration of Forced Migrants: Evidence for Post-War Germany.' *The Economic Journal* 123(571), 998–1024

Becker, Gary S., and Kevin M. Murphy (1988) 'A theory of rational addiction.' *Journal of Political Economy* 96(4), 675–700

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

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

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

Galor, Oded, Kaivan Munshi, and Nicholas Wilson (2013) 'Inclusive Institutions and Long-Run Misallocation'

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

Harris, J.R., and M.P. Todaro (1970) 'Migration, unemployment and development: a two-sector analysis.' *American Economic Review* 60(1), 126–142

Hsieh, Chang-Tai, and Peter J. Klenow (2009) 'Misallocation and manufacturing tfp in china and india.' *The Quarterly Journal of Economics* 124(4), 1403–1448

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

Imbens, Guido W., and Donald B. Rubin (1997) 'Estimating outcome distributions for compliers in instrumental variables models.' *Review of Economic Studies* 64(4), 555

Karlan, D., M. Mobius, T. Rosenblat, and A. Szeidl (2009) 'Trust and Social Collateral.' *Quarterly Journal of Economics* 124(3), 1307–1361

Kondylis, F. (2008) 'Agricultural Households and Conflict Displacement: Evidence from a Policy Intervention in Rwanda.' *Economic Development and Cultural Change* 57, 31–66

_ (2010) 'Conflict Displacement and Labor Market Outcomes in Post-War Bosnia and Herzegovina.' *Journal of Development Economics* 93(2), 235–248

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

Lee, D.S., and D. Card (2008) 'Regression Discontinuity Inference with Specification Error.' *Journal of Econometrics* 142(2), 655–674

Lewis, W Arthur (1955) *Theory of economic growth* (Routledge)

Lucas, R.E. (2004) 'Life earnings and rural-urban migration.' *Journal of Political Economy* 112(1), 29

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

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

Oster, Emily (2015) 'Unobservable selection and coefficient stability: Theory and validation.' Working Paper 19054, National Bureau of Economic Research, May

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

Pollak, Robert A (1970) 'Habit formation and dynamic demand functions.' *The Journal of Political Economy* pp. 745–763

Redding, S.J., and D.M. 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., D.T. Yang, and X. Zhu (2008) 'Agriculture and aggregate productivity: A quantitative cross-country analysis.' *Journal of Monetary Economics* 55(2), 234–250

Restuccia, Diego, and Richard Rogerson (2008) 'Policy distortions and aggregate productivity with heterogeneous establishments.' *Review of Economic Dynamics* 11(4), 707–720

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 *Current Issues in Urban Economics,* ed. P.M. Mieszkowski and M.R. Straszheim (Johns Hopkins University Press)

Sarvimäki, M. (2011) 'Agglomeration in the Periphery.' SERC DP 80

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

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

Tarkka, Juha (1988) 'Paasikiven aika.' In *Suomen Historia 8,* ed. Juha Tarkka (Espoo: Weilin ja Göös) pp. 1–2

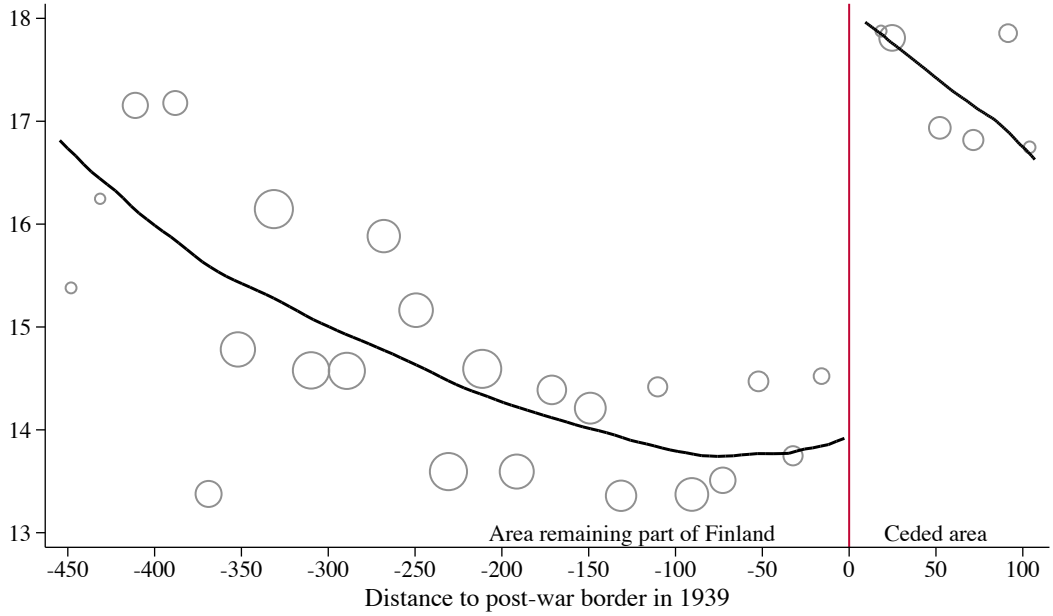
Vollrath, Dietrich (2009) 'How important are dual economy effects for aggregate productivity?' *Journal of development economics* 88(2), 325–334

Waris, Heikki, Vieno Jyrkilä, Kyllikki Raitasuo, and Jouko Siipi (1952) *Siirtoväen sopeutuminen* (Otava)

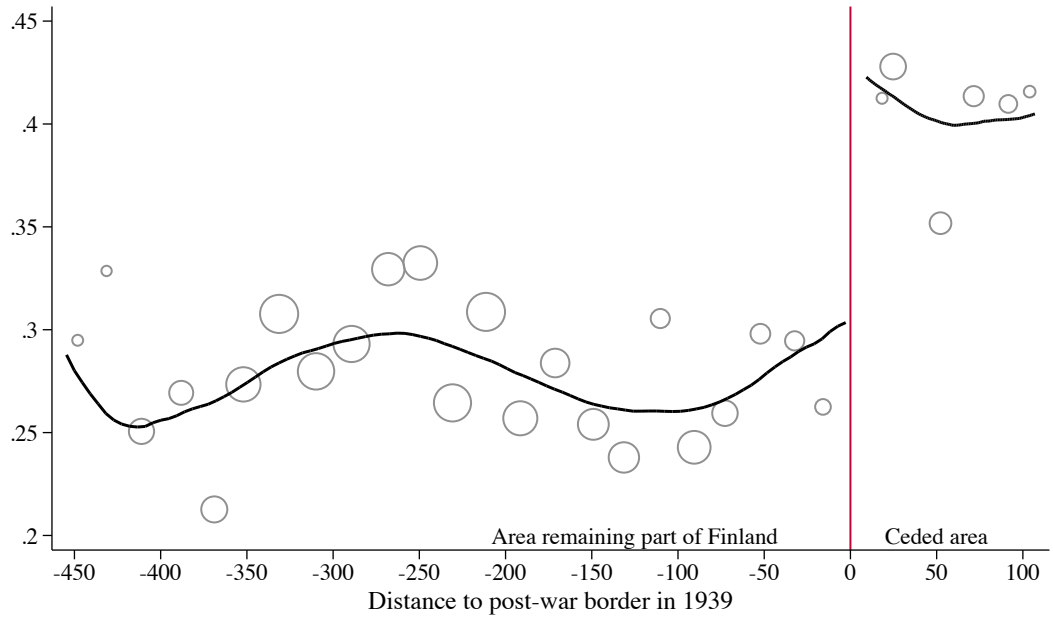
Young, Alwyn (2013) 'Inequality, the Urban-Rural Gap and Migration.' *Quarterly Journal of Economics* 128(4), 1727–1785

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

(a) Annual Real Income in 1971 (including zeros)

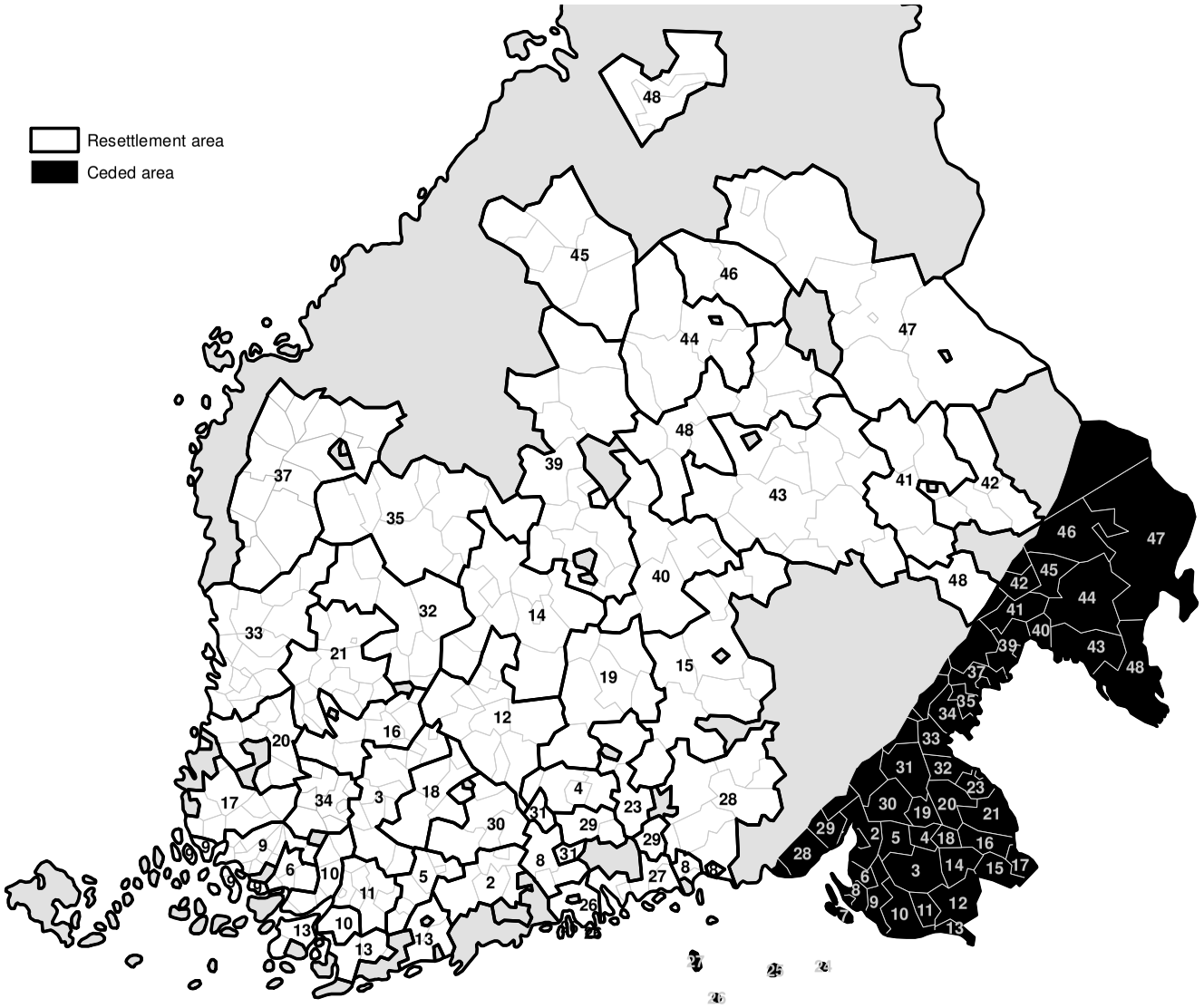


(b) Non-Agricultural Employment in 1970



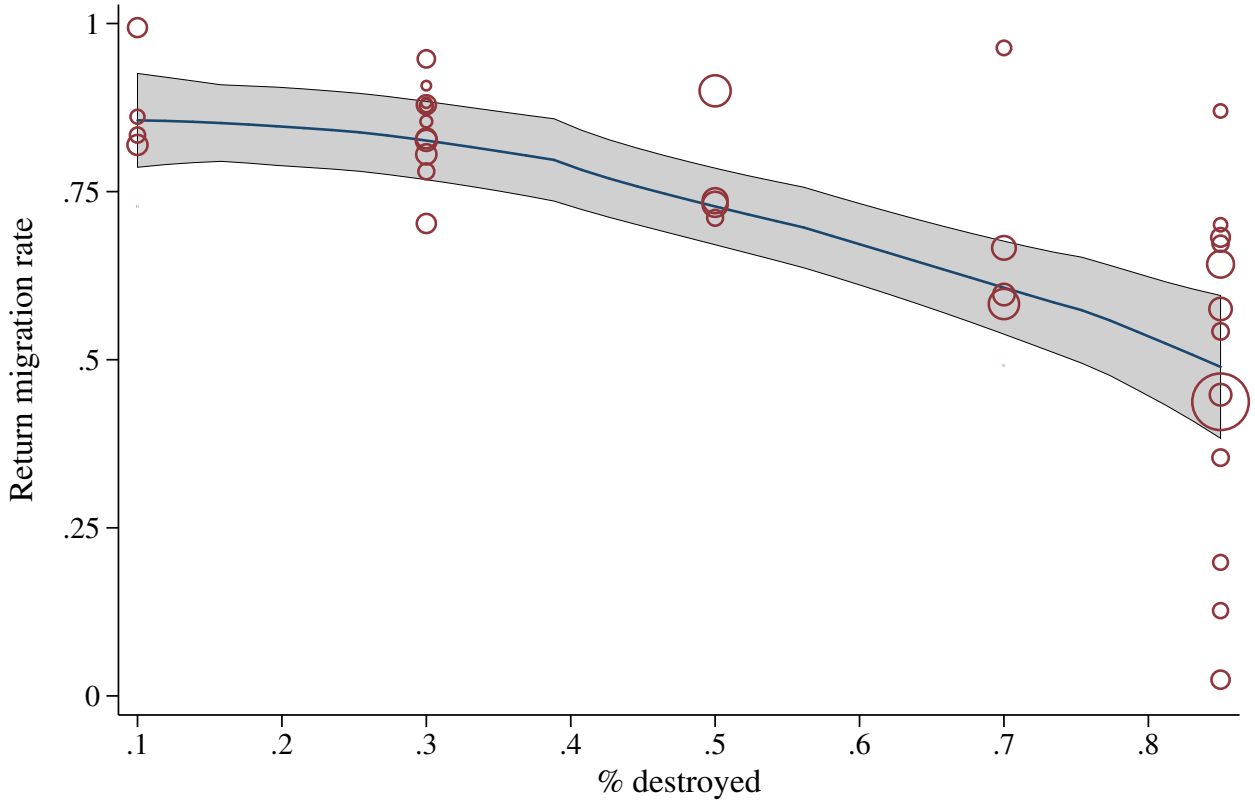
Note: The top panel plots annual earnings in 1971 (including zeros) by distance of the 1939 residence municipality to the post-war border (in kilometers). Positive numbers on the x-axis refer to areas that were ceded to the Soviet Union in 1940/45. The sample used in this analysis consists of 13,987 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 selection algorithm of Imbens and Kalyanaraman (2012). The dots correspond to the sample means by 20km bins. On average, each dot represent 478 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: Return Migration and the Destruction of the Housing Stock



Note: 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: Souce: The share of existing housing stock destroyed by December 31st, 1941. Source: Waris et al. (1952, Appendix Tables 7 and 9)

Table 1: Pre-War Characteristics

	Men						Women					
	Agricultural		Other rural		Urban		Agricultural		Other rural		Urban	
	Displaced (1)	Non-disp. (2)	Displaced (3)	Non-disp. (4)	Displaced (5)	Non-disp. (6)	Displaced (7)	Non-disp. (8)	Displaced (9)	Non-disp. (10)	Displaced (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
B: Education (1950)												
Less than primary	0.21	0.13	0.15	0.13	0.07	0.08	0.19	0.16	0.16	0.13	0.09	0.08
Primary	0.78	0.86	0.76	0.77	0.70	0.71	0.79	0.83	0.76	0.78	0.67	0.70
More than primary	0.01	0.01	0.08	0.10	0.23	0.21	0.01	0.01	0.08	0.09	0.23	0.23
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	0.48	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	1.40	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
Manufacturing LFS	0.06	0.06	0.11	0.07	0.29	0.27	0.06	0.05	0.09	0.07	0.29	0.27
Construction LFS	0.01	0.00	0.01	0.01	0.05	0.06	0.01	0.00	0.01	0.01	0.06	0.06
Services LFS	0.03	0.04	0.04	0.05	0.18	0.24	0.03	0.04	0.04	0.04	0.19	0.24
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	12,610	1,377	10,239	1,258	7,940	889	7,165	831	18,223	2,259	11,393	1,191

Note: Sample means of pre-war covariates.

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

	Cont. Means	Baseline		Oster's Bound	Spatial RD		Resettlement Area FEs	
Status in 1939	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				A: Men				
Agricultural	14.7	2.68	2.22	2.07	4.14	3.71	1.68	1.65
		(0.52)	(0.50)	(0.46)	(1.22)	(0.98)	(0.47)	(0.63)
Other rural	22.4	1.24	1.66	1.78	1.27	2.88	0.56	1.78
		(0.79)	(0.80)	(0.68)	(3.36)	(3.07)	(1.09)	(1.22)
Urban	31.5	-3.00	-3.97	-4.26	.	.	.	.
		(2.28)	(0.64)	(1.07)				
				B: Women				
Agricultural	2.7	0.73	0.81	0.83	0.64	0.36	0.57	0.90
		(0.21)	(0.24)	(0.30)	(0.67)	(0.50)	(0.20)	(0.32)
Other rural	6.7	1.39	1.78	1.90	1.84	1.50	1.11	1.54
		(0.26)	(0.26)	(0.28)	(0.92)	(0.70)	(0.27)	(0.36)
Urban	11.7	-0.82	-1.36	-1.52	.	.	.	.
		(0.85)	(0.25)	(0.46)				
Controlling for:								
Pre-war char.		no	yes	.	no	yes	no	yes
Resettlement Area FE		no	no	.	no	no	yes	yes

Note: OLS estimates for an indicator variable for living in the ceded area in 1939. Standard errors (in paretheses) are clustered at the 1939 municipality of residence level. Each estimate stems from a separate regression. Pre-war characteristics: year of birth dummies, indicators for speaking Swedish as one's mother tongue, living outside of one's municipality of birth, education, sector of employment, socioeconomic status, residence municipality's taxable income per capita (quintile dummies) and latitude.

Table 3: Impact of Forced Migration on Working Outside of Agriculture in 1970

	Cont. Means	Baseline		Oster's Bound	Spatial RD		Resettlement Area FEs	
Status in 1939	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				A: Men				
Agricultural	0.28	0.12	0.14	0.14	0.13	0.14	0.10	0.16
		(0.02)	(0.02)	(0.02)	(0.05)	(0.04)	(0.02)	(0.02)
Other rural	0.63	0.03	0.02	0.01	-0.11	-0.07	0.01	0.00
		(0.02)	(0.02)	(0.02)	(0.06)	(0.05)	(0.02)	(0.02)
Urban	0.76	-0.03	-0.03	-0.03	.	.	.	.
		(0.01)	(0.01)	(0.02)				
				B: Women				
Agricultural	0.15	0.04	0.05	0.06	0.04	0.02	0.03	0.07
		(0.02)	(0.02)	(0.02)	(0.05)	(0.04)	(0.02)	(0.02)
Other rural	0.34	0.08	0.10	0.11	0.06	0.07	0.06	0.08
		(0.01)	(0.01)	(0.01)	(0.04)	(0.04)	(0.01)	(0.02)
Urban	0.55	-0.04	-0.04	-0.04	.	.	.	.
		(0.02)	(0.01)	(0.02)				
Controlling for:								
Pre-war char.		no	yes	.	no	yes	no	yes
Resettlement Area FE		no	no	.	no	no	yes	yes

Note: OLS estimates for an indicator variable for living in the ceded area in 1939. Standard errors (in paretheses) are clustered at the 1939 municipality of residence level. Each estimate stems from a separate regression. Pre-war characteristics: see the notes for Table 2.

Table 4: Real Annual Income in 1971 Conditional on Post-War Sector and Location

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Agricultural	2.22	0.84	0.37	0.12	0.81	0.23	-0.09	0.00
	(0.50)	(0.45)	(0.41)	(0.40)	(0.24)	(0.17)	(0.28)	(0.27)
Other rural population	1.66	1.41	0.54	0.30	1.78	0.47	0.20	0.13
	(0.80)	(0.73)	(0.75)	(0.77)	(0.26)	(0.19)	(0.25)	(0.24)
Urban population	-3.97	-3.37	-4.10	-3.87	-1.36	-0.75	-0.55	-0.48
	(0.64)	(0.65)	(0.72)	(0.61)	(0.25)	(0.21)	(0.23)	(0.17)
Controlling for:								
Pre-war char.	yes	yes	yes	yes	yes	yes	yes	yes
1970 non-agricultural job	no	yes	yes	yes	no	yes	yes	yes
1970 industry and location FE	no	no	yes	yes	no	no	yes	yes
1950 industry FE	no	no	no	yes	no	no	no	yes

Note: Estimates for an indicator variable taking value of one if the person was displaced after WWII and zero otherise. Standard errors (in paretheses) are clustered at the 1939 municipality of residence level. Each estimate stems from a separate regression. Pre-war characteristics: see the notes for Table 2.

Table 5: 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	14.0	8.3	7.7	16.7	16.8	13.8	10.9
		(0.0)	(0.0)	(2.5)	(2.2)	(3.5)	(3.0)
Women	1.0	11.4	11.1	1.0	0.9	16.4	14.2
		(0.3)	(0.3)	(1.4)	(1.4)	(3.1)	(2.7)
Controlling for							
Pre-war characteristics		no	yes	no	yes	no	yes
Resettlement area FEs		yes	yes	yes	yes	yes	yes

Note: Column 1 shows the real annual earnings in 1971 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 clusterer at the 1939 residence municipality level. Pre-war characteristics: see the notes for Table 2.

Table 6: Impact of Forced Migration on Income by Age and Pre-War Migration

	Men			Women		
	Agri-cultural (1)	Other rural (2)	Urban (3)	Agri-cultural (5)	Other rural (6)	Urban (7)
A: Impact by Pre-War Migration						
Displaced	2.7	2.2	-4.3	0.6	2.2	-0.4
	(0.5)	(0.8)	(1.0)	(0.3)	(0.3)	(0.4)
Migrated before 1939	0.8	1.7	0.6	0.1	1.2	0.4
	(0.3)	(0.5)	(0.5)	(0.2)	(0.2)	(0.2)
Displaced ×	-3.0	-1.5	0.7	0.8	-1.1	-1.6
Migrated before 1939	(0.9)	(1.1)	(1.7)	(0.5)	(0.5)	(0.4)
B: Impact by Age						
Displaced	6.0	5.2	-6.7	4.0	4.1	-2.0
	(1.7)	(2.5)	(2.6)	(1.2)	(0.8)	(0.9)
Displaced ×Age	-0.2	-0.2	0.1	-0.1	-0.1	0.0
	(0.1)	(0.1)	(0.1)	(0.0)	(0.0)	(0.0)

Note: The main effect of age in panel B is captured by year of birth fixed-effects.