Child Evacuees during WWII: Estimating Nurture Effects Using a Natural Experiment*

Torsten Santavirta†

JOB MARKET PAPER

November 18, 2009

Abstract

During WWII some 70,000 Finnish children were evacuated to Sweden and sorted in a random fashion to foster families. This quasi-random assignment produced a natural experiment whereby the temporary rearing environment of Finnish child evacuees can be considered exogenous to their biological family background and parental decisions. Using register and survey data I examine the extent to which the foster environment affected later life outcomes of Finnish child evacuees. The results show that nurture effects - the socioeconomic environment at early stages of life - have an important effect on schooling in the form of school track choice. Also significant nurture effects on long-run outcomes, such as labor market outcomes and risky behavior, are found.

---

*I am grateful for constructive comments by Pierre-André Chiappori, Per-Anders Edin, Lena Edlund, Peter Fredriksson, Bertil Holmlund, Markus Jäntti, Mikael Lindahl, Otto Toivanen, Roope Uusitalo and seminar participants at Columbia University, Helsinki School of Economics, Universitat Pompeu Fabra, and Uppsala University. Financial support from Yrjö Jahnsson Foundation (Grant No.) is gratefully acknowledged.

†Department of Economics, Helsinki School of Economics, P.O. Box 1210, 00101 Helsinki, Finland. E-mail: torsten.santavirta@hse.fi

1 Introduction

During WWII some 70,000 Finnish children aged between one and ten years were evacuated to Sweden and sorted in a random fashion to foster families. The evacuations were conducted between 1941 and 1946 by a large evacuation scheme with organized logistics on both sides of the border. A first-come first-served policy was applied where the children were assigned a running number and processed accordingly. The average age at evacuation was roughly five years and the average time spent in foster family care was roughly two years. In this paper I exploit the natural experiment that the evacuation scheme gave rise to in order to examine how the children's human capital formation is determined by a temporary switch in the socioeconomic family environment.

I use a random sample drawn from a war time register where detailed preintervention data on the evacuees is stored and complement this data with a survey collecting information about the foster and biological family environment and a wide range of outcomes, conducted in 2005. In total, 887 surveys were received back for a response rate of above 60 percent.

My principal outcome measure is school track choice, i.e., the choice of going to general secondary school (preparing for academic studies) after fourth grade of primary school or continuing the primary school track (preparing for vocational studies).[1] I find that a one standard deviation increase in foster father's occupation-based socioeconomic index (e.g. a move up from being a sheet metal worker to becoming an office clerk) increases the probability of going to secondary school by 3.8 percent. Also, the results show that a randomly assigned, temporary rearing environment during childhood accounts for up to 30 percent of a child's intergenerational transmission of schooling in the form of school track choice. This is a large impact considering the amount of time spent with the foster parents. Since school track choice is made at age eleven there are nine months prebirth and ten years postbirth during which parental input has an impact. The foster parents had an impact during two of the

[1] School starting age was seven years in Finland and Sweden in 1940s, so the children were tracked at an age of eleven years.

ten years postbirth and no impact via biology. Hence, assuming that the parental inputs add up in a linearly additional way, one would expect the foster family input to account for less than 20 percent of the environmental share of intergenerational transmission of schooling. Previous research suggests that postbirth environment accounts for less than half of the total intergenerational transmission (genetic factors included), implying a less than 10 percent contribution by the average foster family input to the total intergenerational transmission of school track choice. In light of this, I find a substantial impact of a temporary switch in the childhood rearing environment on school track choice. The results also show that parental input at ages below five years has a larger impact on schooling than it has at later stages of childhood. I further find that a temporary rearing environment during childhood has an important impact on long-run outcomes spanning over the entire working career, such as the risk of unemployment. More specifically, I find that, for men, a one standard deviation increase in foster father's socioeconomic index leads to a 5.3 percent lower probability of being unemployed.

The novel result that this paper provides is that a relatively short lasting switch in the early childhood environment may have substantial effects on later life outcomes. This result has important implications for policy. For instance, early childhood interventions may not necessarily need to be long lasting to have important effects: A two-year intervention into a child's environment may make a substantial difference for the cognitive development of the child.

By looking at how the family background affects later life outcomes this paper contributes to the intergenerational transmission literature that seeks to disentangle the parents' environmental influences from the prebirth environment (genetics and the fetal environment) of the child. This strand of literature (Dearden, Machin, and Reed 1997; Plug 2004; Björklund, Lindahl, and Plug 2006 (henceforth, BLP); Sacerdote 2007) uses data on adopted children to pin down the part of postbirth effects that are orthogonal to prebirth effects. The argument goes as follows: If adopted children share only their parents' environment but not their genes and fetal environment, any relationship between the child's later life outcomes and the adoptive parents is driven by the influence the parents have on their children's environment.

Potential selectivity in the placement of children makes it however difficult to establish a causal relationship between environmental effects and later life outcomes.

To this date, I know of two studies that control for potential selection effects. Sacerdote (2007) shows convincing empirical evidence for the random assignment of Korean American adoptees to adoptive families. BLP use information on both adoptive and biological parents to control for the impact of selective placements.[2] By estimating intergenerational transmission coefficients, these studies find that both adoptive mother's and father's schooling and income matter for the child's schooling outcomes, although more so for their schooling than for income. Rather striking is the finding in these studies, that biological mother's schooling and father's income have twice the effect of foster mother's schooling and father's income.

The analysis in this study is also related to the literature on neighborhood effects on the creation of human capital. Kling, Liebman, and Katz (2007) use data from an experiment where housing vouchers are offered in a random fashion to residents of disadvantaged neighborhoods. This social experiment creates exogenous variation in the neighborhood placement of households enabling the analysis of short-run and long-run effects on future outcomes, such as labor market outcomes, health and risky behavior.[3] The findings are that important neighborhood effects exist for some groups, e.g. teenagers. Gould, Lavy, and Paserman (2004) make use of a quasi-random assignment of temporary housing locations to Ethiopian immigrants (and thereby initial schooling environment of the children), who were airlifted to Israel during the so called "Operation Salomon", i.e., an emergency rescue of Ethiopian Jews in 1991. This random variation in schooling quality enables them to estimate the causal effect of schooling quality on future outcomes. The results show that the early schooling environment has an important effect on a number of high school outcomes.

By studying environmental effects on outcomes that span up to 65 years, I am able to address questions on the individual's health, labor market outcomes, fertility,

[2] BLP find in their data evidence for selective placement effects in the form of a positive association between background characteristics of biological parents and adoptive parents.

[3] Also Oreopolous (2003) and Edin et al. (2003) have done work along these lines.

and marriage with very little truncation in the data. Hence, this paper is also contributing to the study of long-run effects of the early childhood environment. To this date only one study has examined long-run outcomes that span over almost a life time. Gould, Lavy and Paserman (2009) exploit another airlift, the operation to airlift Yemenite immigrants to Israel in 1949 and randomly place them across the country, to study the long-run consequences of early life environment some sixty years after. They find that children who were placed in a good environment outperformed their less fortunate compatriots in a number of dimensions, they were, e.g., more likely to obtain higher education, marry at an older age and have fewer children.

Finally, this study is related to the extensive literature in economics that demonstrates the importance of early environmental conditions on the evolution of adolescent and adult cognitive and noncognitive skills (see Cunha et al. 2006 for an extensive survey). These skills largely determine the educational attainment, crime, earnings and participation in risky behaviors. Experimental interventions with long term follow up, surveyed in Blau and Currie (2006) and Cunha et al. (2006), confirm that changing the resources available to disadvantaged children improves their adult outcomes on a number of dimensions.

This paper contributes to the existing nature and nurture and early life environment - later life outcome literatures in the following five ways. First, I exploit register data on the age at evacuation, duration of stay with the foster parents, and survey data on whether the foster parents differentiated between foster children and foster siblings to control for these three factors when estimating environmental effects. This allows me to relax two strong identifying assumptions that BLP and Sacerdote (2007) make: (i) They assume that the children are adopted as newborns. Since adoption-at-birth is unrealistic and numerous studies point to the causal relationship between early life environment and later life outcomes, controlling for age at intervention is key; (ii) They assume that parents do not differentiate between adoptive children and their adoptive siblings.[4] Second, in identifying environmental

[4]BLP acknowledge the strength of the adoption-at-birth assumption and provide robustification checks that show, that the bias in pre- and postbirth estimates when not controlling for age at adoption seems to be small.

effects, I am able to both control for biological family characteristics in the estimating equation and to provide empirical evidence that the children were randomly assigned to foster families. Each of the two aforementioned adoption studies fail on either of these two (BLP find positive sorting into adoptive families and Sacerdote (2007), while being able to convince that the assignment is random, lacks data on the biological parent characteristics). Third, even though I am not able to decompose intergenerational transmission coefficients into prebirth and postbirth components due to the temporary nature of the intervention (all children in my sample returned to their biological family in Finland after WWII), it may, from a policy perspective, be of even more interest to look at how a temporary switch in the rearing environment affects later life outcomes. Most policy relevant interventions into children's environment are temporary by nature and do not last into adulthood, as adoption does, and in this sense, my estimates may give a more realistic upper bound estimate of an intervention into a child's early life environment. Fourth, due to the nature of the evacuation program and the external shock (i.e. the war) that suddenly caused adverse conditions for children who came from the whole range of pre-war socioeconomic backgrounds (many professionals sent their children away, as is seen from the descriptive statistics in Section (3.4)), both biological families and foster families are expected to be less a selected group of families as compared to the ones in adoption data. Ultimately, as already mentioned, this paper contributes by looking at very long-run outcomes of early life environment, in fact to my knowledge, the longest spanning outcomes examined in the literature.

The remainder of this paper is organized as follows. The next section presents the historical background of the operation to evacuate Finnish children to Sweden and describes in detail the evacuation scheme. It also provides anecdotal evidence that evacuees were randomly sorted into foster families. Section 3 describes the data and the empirical strategy. Section 4 tests for whether the data accord with the anecdotal evidence of random placement of the evacuees into foster families, as described in section 2. Section 5 presents the empirical estimates of the effect of the foster family environment on schooling and a number of alternative later life outcomes. Section 6 concludes the paper.

2 The Evacuation Scheme: The Case for a Natural Experiment[5]

2.1 Historical Background

Finland fought three wars during World War II: The Winter War (November 1939 – March 1940) against Soviet Union, the Continuation War (June 1941 – September 1944) against Soviet Union, and finally, reluctance of German troops to leave Finnish territory after the armistice with the Soviet Union led to brief encounters in Lapland. The last foreign troops had withdrawn from Finland by April 27, 1945. Despite Finnish collaboration with the Axis Powers during WWII, Finland's resistance against Stalin's Red Army gave rise to feelings of sympathy in many countries, in particular in neutral countries such as Sweden and Switzerland, who expressed their willingness to provide humanitarian help.

In Sweden, a civilian movement to help Finland emerged within days after the breakout of the Winter War. The politically most acceptable way for Swedish civilians to participate in helping the Finnish people was to express their willingness to act as foster parents to those Finnish children, who were most exposed to the adversities of war. A plan to evacuate Finnish children to Swedish families came from the private initiative of Maja Sandler, the wife of the then Swedish Minister of Foreign Affairs, Rickard Sandler. A unit called The Finnish Centre for Nordic Help was set up in Helsinki in December 1939 with the aim to help with all practical arrangements related to the evacuations. The first official transportation of children from Finland to Sweden took place on December 15, 1939. Altogether 9,000 children were evacuated to Swedish families during the winter of 1939-1940.

As the Continuation War broke out in June 1941, the idea of a large scale operation for evacuating Finnish children to Sweden was resumed. In Sweden, a voluntary organization called the Support Committee of Help for Finnish Children (henceforth

[5]The chief part of the documents that this section is based on are available in the Files for The Child Evacuation Scheme during WWII at The National Archives of Finland (NAF). When not specifically mentioned, we refer to Lomu (1974), who constructs a detailed report on the evacuation scheme based on the documents of NAF.

"the Placement Committee") was established, and a large network of families willing to accommodate child evacuees was created.[6] Shortly after, negotiations regarding the transfer of children to Sweden started and on September 2, 1941, The Committee for Transporting Finnish Children to Sweden (hereafter called "the Evacuation Committee") was established in Finland under the auspices of The Ministry of Social Affairs. The Evacuation Comittee was granted funding by the Finnish government for organizing the logistics of the evacuation scheme and for covering the travel expenses of the evacuees.[7] The Swedish government also covered part of the transportation expenses. However, the chief part of the financial support on the Swedish side came from private fund raising in the form of collections, lotteries, and campaigns.

The Evacuation Committee designed the evacuation scheme and carried out the major part of the evacuations made between 1941 and 1946.[8] The official motives for a mass evacuation of children were, as stated by the Evacuation Committee, that children who were particularly exposed to the various adversities of war should be given a better rearing environment.

In Sweden, the Support Committee established a placement scheme, and took over all administrative and financial matters on the Swedish side of the border. Fortunately (for me), documents such as all minutes of both aforementioned committees' meetings and descriptive statistics of the evacuation are stored in the National Archives of both Finland and Sweden. It is thus a fairly simple task to construct a picture of the evacuation scheme and the way the evacuations were conducted. Subsections describe the stages of the evacuation from children's "biological" home to their final placement in foster care in Sweden.

The child transfers during the Continuation War can be divided into two waves. During the first wave, from September 1941 to June 1943, 22,398 children were evacuated through the Evacuation Committee. When the war changed its character

[6] No financial compensation for accomodating Finnish children was ever promised to the families, and none was ever going to be rewarded either for that matter. In other words, accomodating evacuees was based on purely philantropic grounds.

[7] The word comittee is rather misleading in describing both the Finnish and the Swedish organizations, as these were as much executive bodies as performing the more traditional functions associated to the word comittee, i.e. governing and coordinating functions.

[8] Most evacuations were made in the winter between 1941 and 1942 and in 1944.

towards trench warfare in 1942, some of the pressure on the civilian population eased, and parents started to reclaim their children. As a consequence, some 9,000 children returned in 1943. The second wave of evacuations took place in spring 1944, first during the bombings of Helsinki, and later that same spring at the break of the massive Soviet offensive on the Karelian isthmus. During the second evacuation wave, in 1944, 29,268 children, some of which were re-evacuees who had returned in 1943, were evacuated to Sweden. In total 48,628 children were evacuated to Sweden by the Evacuation Committee. Apart from the evacuations supervised by the Evacuation Committee, roughly 10,000-15,000 children were sent to Sweden independently from the official evacuation program through private bilateral organizations and to family and acquaintances. In total, it is estimated that roughly 65,000-70,000 children spent between one to five years in a foster family in Sweden during WWII. The total amount of evacuees equals the size of one Finnish cohort during the 1930s, i.e., 70,000.

In this study, I restrict the analysis to the children who were evacuated within the official evacuation scheme, i.e. the population for the study consists of those 48,628 children who were evacuated by the Evacuation Committee. I will in the following subsections claim that the course of events that took place between separation from the biological parents to the final placement in foster families led to a random assignment of foster parents to the evacuee.

2.2 Eligibility Criteria

At first, each Finnish county was granted a quota of evacuees, but restricting the selection of participants to the stated quotas proved difficult as fear for air raids spread among the urban population, and food became scarce. The original eligibility criteria were the following: 1. Children of evacuated Karelians 2. children whose fathers were wounded in battle 3. children who had lost their home in bombings 4. children whose father's had died in war or who had lost their parents in bombings. In January 1942, the criteria were expanded to comprise children from large families, and children whose mothers were working; also children who resided in towns that

were potential targets for air raids. This latter criterion applied, in practice, to most eastern and southern towns of Finland, where hence most children were considered eligible to the program. At first the objective was to send preschoolers. The age criteria proved difficult to apply in practice due, for instance, to the difficulty to age-discriminate between siblings. Statistics show that roughly half of the children were past school starting age by the time of evacuation.

The lax criteria implied that a large part of the cohorts born within say 9 years before the war were eligible to the program. Clearly the most popular reason stated was "being evacuated from the conquered areas in Karelia" (28.2 percent). Also, the majority of the applications of children who did not pass the formal criteria were accepted.[9]

As the main form of placement in Sweden was with foster families, who themselves had no say in which child they would be assigned, it was considered just to restrict the selection to so called "normal" children. This means that no orphans were considered and neither were mentally retarded or psychologically unstable children. Physically sick children were, however, sometimes considered, and in fact, hospitals were set up in Sweden to treat these children.

2.3 Randomness in the Evacuation from Finland to Sweden

A large organization was set up on the foundations of the already established war time logistic organizations and volunteer institutions to carry out the evacuations.[10] The organization made itself known through nationwide broadcasting and advertising in local newspapers. Headquarters were established in Helsinki, employing a gradually growing administrative staff consisting of office clerks and volunteers, such as school teachers. Regional subsidiary offices were set up in each county headed by the person in the county's administrative board who was already in charge of

[9] From the applications, stored at the Archives for War Evacuees at National Archives in Finland, it is clear that 35.3 precent entered the evacuation program without any particular motivation (apart from being exposed to the general adversities of war).

[10] The Civilian Service of Finland contributed with an already existing network of volunteers. This organization was founded in 1939 for the purpose of distributing aid packages provided by foreign charity oragnizations to the civilian population.

the relocation program of migrants from the areas in Karelia, that had been conquered by the Red Army. The local county offices - which received funding from the MoS through the Evacuation Committee, but were, largely run by volunteers such as local nurses and representatives from Christian and feminist organizations and political parties (in particular the female branch of the Social Democrats) - started to process applications of parents wanting to send their children. The Evacuation Committee emphasized the importance of their unconditional possession of control over the evacuations in order to avoid incompleteness in the registers of evacuated children. As the evacuations began in September 1941 the county offices were converted into evacuation centers handling the selection process, transportation arrangements, documentation, accounting, correspondence between the evacuees and their families, and advertising of the program in the local media. Absorption centers were set up near the crucial ports of Turku and Vaasa, from where the evacuees were shipped to Sweden and in the border towns of Tornio and Kemi, from where the evacuees crossed the border by railway at the Tornio-Haparanda checkpoint. All travel expenses were covered by the Evacuation Committee.

Parents who heard of the evacuation program through mass media or local authorities first filed an application to the local evacuation office.[11] If the application was accepted, a health check at the local hospital followed and an identification document including a photograph was issued at the local police office or parish (as the majority lacked an official passport it was agreed that this document would work as a substitute).[12] The documents were sent to the local evacuation office where the child's file was put into a queue awaiting for information on the region's quota for the following evacuation round. As soon as the headquarter of the organization gave information of each region's quota for the next round, the county office gathered the children - according to the ordered date of application approval - to the nearest railway station from where they were sent to the absorption centers. The

[11] If the father was on duty the mother could file the application without his consent. However, in case the father should disapprove, he could have his child returned immediately from their placement.

[12] It took on average 21 days from the time of filing the application to embarking the evacuation transit to Sweden (author's calculation).

children were not allowed to carry any money and their ration cards (most grocery products were rationed in Finland during the war) needed to be handed in before departure.[13] The children brought with them a franked envelope, which the foster parents that would be assigned in Sweden, were urged to post to the Finnish county office with notification of the child's arrival and their name and address, to enable correspondence between the biological parents and their child.

Upon arrival at the absorption center, a brief health check was conducted and the information on the children was entered into an evacuee register. Each child was assigned a running number according to arrival order and given an identification plate to carry around the neck.

Until 1944, most transports were conducted by boat due to the capacity constraints of the highly congested Swedish railways. Also air planes were occasionally used during the winter season when an impenetrable ice layer prevented shipping. From 1944 onwards, train was considered the safest way of transportation due to the hydro-mines spread around the Finnish gulf.

Based on the above description of the evacuation, it is plausible that the evacuation created randomness in the order that each child was transported to Sweden. The strongest arguments in favor of random order of evacuation with respect to any background characteristics are that the children were processed according to a running number upon the arrival to the absorption centers (and boarded the vehicle of transportation accordingly). Also the fact that, on the ships, complete registers of the children on board were put together en route, suggests that the separation into small groups upon embarkment is likely to have taken place in a haphazard fashion. The fact that no money or ration cards were allowed on the trip is important for the sake of the random assignment argument in that it suggests, that the children could not possibly have affected the placement by bribing the officials. Furthermore, the biological parents' unawareness of both the final destination and the identity of the foster parents, for which the request for identity and address of the foster parents in the form of a franked envelope, is a concrete indicator, made any possible manipulation of the placement more difficult.

[13] This made any form of bribing of the officials difficult for the children.

2.4 Randomness in the Placement to Foster Families in Sweden

In Sweden, the structure of the Support Committee was literally a mirror image of the Evacuation Committee on the Finnish side. Its main office was located in the premises of the Anglais, a large hotel in Stockholm. Also each county had its own local committee led by an authorized representative who was in charge of the placement of the children into families. In practice, large local volunteer organizations conducted the major part of the placement, and the provincial offices handled registers of children and other administrative issues. So called quarantine centers, i.e. some kind of absorption centers, were established in geographically strategic Swedish towns, usually the capital of the county. Some of the staff in the quarantine centers was sent from Finland, and, in centers located in the northern parts of Sweden, e.g. the absorption center in the town of Boden, many volunteers from Finnish speaking areas of Sweden were employed in order to ease communication with the children.

At first, as the contingents arrived in Swedish territory, they were taken to so called sanitary centers, which were located in the near proximity of the arrival port or station (Stockholm, Umeå and Haparanda), where brief health checks and delousing was conducted. At the sanitary centers, the contingents were split into smaller groups, which were transported to quarantine centers. In these centers, the children were kept for a week[14] to make sure they were not carrying any contagious diseases. They went through careful health checks before being assigned a county of destination and finally a foster family. Although medical issues were of first priority, much weight was also put on nutrition and inventory of the children's luggage. Clothes were provided to poorly equipped children, and all children were cleaned thoroughly. When ready to leave the quarantine centers, the children were separated into smaller groups and transported via the county offices - where the group would be re-shuffled into smaller units - to their final destinations, that is, to towns and villages where the foster families-to-be were awaiting them.

[14]The children who were evacuated towards the end of the war in 1944 were generally in worse health and there were even some alleged death cases of foster parents due to contagion of disease. Thus the quarantaine period was prolonged to 14 days in spring 1944.

Little is documented about the final stage of the evacuation, that is, the placement in foster families, and thus one is principally referred to anecdotal evidence as recalled or retold by the evacuees.[15] Regarding the different stages of the journey until the local Swedish provincial offices, anecdotal evidence conforms with the information in official documents unusually well. The final stage of the trip to the eventual placement seems to have been completed chiefly in three different ways: 1. when several children arrived at the same time to a community, the distribution of children to foster families would take place at some temporary lodging, e.g. the local parish house or school premises, according to a first-come-first-pick process by the parents-to-be,[16] 2. in sparsely populated areas, or in areas where only one family was awaiting a child from the particular contingent, the assigned family would receive the child on a bus stop or train station without any possibilities to affect the choice of child, 3. in some cases the local ombudsman for the placement committee, often the local priest or school principal, assigned the children to families at their arrival on the train or bus station.

Anecdotal evidence reveals that foster parents were, in some cases, able to present preferences regarding children's age and gender. However, these wishes were not always met, a fact which, however rarely, according to the numerous first person accounts, led to the rejection of the child. Marja Leskipohja, aged only 9 months at the time of evacuation, recalls: "my foster parents have later told me that they were hoping to host a boy but as they arrived to pick up the child I was the only one left of the contingent that had arrived to Karlstad".

The description of the events during the journey to the final destination suggests that the children were processed anonymously according to their assigned running

[15] I make use of a compilation of 135 short stories of recollections by the evacuees edited by Lehtiranta (1996).

[16] Pirkko Bergman, a one time child evacuee, recalls "the children were taken to a room and told to sit on chairs with the identification plates visibly displayed. A group of people rushed in and among the first ones was a tall man who examined her, wrote something indicating "reserved" on the parcel she was carrying and continued to examine others. Others showed interest in me but noticed the reservation sign on the package. The tall man, who turned out to be the priest of the village of Åsunden, Gösta Rosen, returned to pick me up together with another child, who, I found out at arrival to the village, he had chosen for himself, I was assigned to a neighbour family" (Lehtiranta ed. (1996).

number and the purely demographic information on the identification plate, and sorted randomly at several stages of the journey. At departure from the quarantine center, the inequalities in clothing, cleanness, and nutrition are supposed to have been levelled out, and thus to have made any inference of social background based on the appearance difficult.

In conclusion, the evacuation scheme was carried out with exemplary orderliness taking into account that it was largely run by voluntary forces. The random assignment in this case was not achieved by chaotic conditions or a randomized draft, but rather by the deliberate objective to process the evacuees according to their assigned running number and not according to socioeconomic status or kin. The orderliness of the scheme also resulted in systematic documentation of the program, which is presumably also a consequence of the delicate nature of the program, dealing with the fates of young children. This documentation provides credible evidence for the random assignment of foster families to the evacuees. Casual evidence based on first hand accounts by the evacuees also supports this random nature of the assignment. This random variation is valuable in that it creates a temporary rearing environment for the children completely unrelated to the child's prenatal characteristics. There is, however, one caveat to the random assignment that may be innocuous when examined with closer scrutiny. Many of the numerous first hand accounts show that siblings would end up in families living close to each other.[17] One potential reason for this is that siblings, already at the beginning of the journey, were assigned consecutive running numbers. Another option is that the officials, at least in some cases, were trying to arrange it so that siblings would be placed in the same region. In order to be fully convinced that random assignment occurred, I will, in the next few sections, describe the data, and test whether the data supports the anecdotal evidence of random assignment.

[17]Some of the many first person accounts provide evidence that siblings could be separated from each other at any stage of the journey.

3 Data and Empirical Framework

3.1 The Econometric Model

This study is primarily interested in answering the question how the Finnish child evacuees' schooling is associated with the socioeconomic background of their Swedish foster families and with that of their Finnish biological families. The relation between socioeconomic family background and schooling is nontrivial due to the difficulty to sort out the underlying contributions of nature (prebirth environment) from the nurture effects (postbirth environmental factors). The natural experiment I exploit does not allow me to decompose intergenerational transmission into pre- and postbirth factors in the same fashion as the studies using adoption data (BLP and Sacerdote (2007)), which make the identifying assumption that the children move to their adoptive parents immediately after birth. However, since most interventions into childrens' environment are temporary, it may, in terms of external validity, be even more interesting to look at the size of the effect of a temporary switch in environment on a child's future outcomes.

I follow the same empirical strategy as BLP and Sacerdote (2007) in that I relate child outcomes to parental inputs and estimate intergenerational transmission coefficients. I include both foster and biological family characteristics in the estimating equation and show empirically in section (4) that the children were randomly assigned to foster families. My model explains schooling y of person i, who was assigned a foster family in Sweden with socioeconomic status F_i, and whose own biological family had socioeconomic status B_i, with the following equation:

$$y_i = \lambda_0 + \lambda_1 F_i + \lambda_2 B_i + \delta \left(F_i \times C_i\right)' + u_i \tag{1}$$

Here C_i is a vector of controls for age at, and duration of, exposure to foster family input. A dummy whether the foster family differentiated between the inputs to foster children and foster siblings is also included in C_i (non-differentiation=1). The error term u_i represents an unobserved child-specific characteristic assumed to be uncorrelated with both F_i and B_i. In case that the assumption of quasi-random

assignment of evacuated children to host families is correct, it guarantees that the key explanatory variable, F_i, is uncorrelated with both B_i and u_i. Thus, I obtain an unbiased estimate of λ_1 even without controlling for B_i. However, B_i is included since inference on the relative importance of the foster family and biological family inputs is key to the analysis that this paper contributes to.

I relate the magnitudes of λ_1 and λ_2 to the magnitude of the coefficient of B_i for the control group in regression $y_i = \alpha + \beta B_i + u_i$. This is the standard version of the transmission model, used widely in the literature on intergenerational mobility.[18] The estimated intergenerational transmission coefficient β represents a combined effect of different mechanisms, including genetic inheritance, fetal environment, and the family environment in which the child grew up. In line with the adoption data literature, my estimate for β is comparable with the sum of the estimates for λ_1 and λ_2.

The control group comprises children who did not experience any changes to their rearing environment during the war period. I construct the control group by matching three control units to every evacuee based on demographic characteristics (age within 1 months accuracy, gender, mother tongue and municipality of birth).

Even though I am able to sort out the foster family input from biological family input in the regression, I must not interpret the foster family parameter as causal. Evacuation to a foreign foster family imply that basically every aspect in the rearing environment changes abruptly. It is thus hard to sort out the causal pathways by which parent's socioeconomic status affects the children. Parental socioeconomic characteristics are correlated with neighborhood and school quality and a list of other socioeconomics factors, which are intertwined with each other. I follow the intergenerational transmission literature, and the recent adoption literature, and discuss the coefficients as associations instead of causal effects.

The model in its simplest form can be estimated with ordinary least squares. As my schooling measure is binary, more exactly, the school track choice, i.e. the choice of going to general secondary school (preparing for academic studies) after fourth grade of primary school or continuing the primary school track (preparing for

[18] See Solon (1999) for a review.

vocational studies), I am referred to a model of discrete choice.[19]

An issue that arises when comparing β with $(\lambda_1 + \lambda_2)$ is that the experience of evacuation may in itself form the treated children (see BLP). If it is the case, that any causal effects between participation in the evacuation program and later life outcomes with respect to, say mental health, exist, then the outcomes I look at might be affected by participating in the program itself. In a companion paper (Santavirta and Santavirta 2009), I evaluate the long-run causal effects of the evacuation programme on later life outcomes (schooling, unemployment, emotional well-being, depressive symptoms) and find a small positive and significant (5 percent significance level) effect on unemployment and schooling but no significant effect on depressive symptoms, emotional well being, life satisfaction or risky behavior. Furthermore, in the empirical analysis, I allow evacuees and the control units to be different by treating them as separate samples, and running the intergenerational mobility regressions separately in the same fashion as BLP. As long as the differences are unrelated to the parents' socioeconomic status, any real evacuation effect should be captured by the difference between the intercepts λ_0 and α.

3.2 Data and the Survey

In order to estimate specification (1) I need four types of information about each evacuee: (i) the socioeconomic background of the evacuee's biological family; (ii) the socioeconomic background of the evacuee's foster family; (iii) information about the duration of the evacuation and the exact age at evacuation; (iv) the evacuee's later life outcomes. To that end register data on the evacuees was combined with a survey conducted for a random sample of the child evacuees.[20]

Register data dating back to WWII are available on the biological family background and data on age at, and duration of, evacuation (points (i) and (iii)) at the National Archives of Finland in the files of the Evacuation Committee. In these files, an evacuee card is stored for each of the 48,628 Finnish war children who were sent

[19] The Finnish two-track school system is described in section (3.3).

[20] I together with a multidisciplinary team of researchers and research assistants from various departments at Helsinki University.

through the official evacuation scheme and were returned to Finland. We drew a random sample of 1,931 evacuee cards from the register.[21] Each card contains detailed information on both the evacuee's background, and the evacuation, including the following variables: name, age, gender, place of birth (municipality), mother tongue, family structure, biological parent's occupations, whether family was evacuated from the Red Army conquered areas (Karelia), whether father was wounded in battle, whether father has fallen in battle, whether exposed to air raids before evacuation, age at evacuation, date of departure, date of returning, number of times evacuated (in case of returning to Finland during seizure of fire and returning). In order to obtain data on the evacuee's foster family background and her later life outcomes, (points (ii) and (iv)) we conducted a survey for our sample.

For the survey we needed updated contact information of the evacuees. The contact information could only be obtained through identifying the social security number of each individual at Population Register Centre in Finland (PRCF). We identified 1,498 individuals from our original evacuee card sample still alive as of June 2005.[22] Of the identified individuals, 325 were living abroad, the most important group of expatriates being those living in Sweden (207 individuals), and the second largest being those without a specified destination country, reported to live "abroad" (75 individuals). A substantial part of the cohorts entering the work force in the 1960s emigrated to Sweden in search for work. It is very likely to be a selected group who found it worthwhile to emigrate in order to, generally, become hired in the expanding car and textile industries. Thus, in order to avoid attrition, it is of utmost importance to include the exiles in the data. We asked the Swedish Tax Agency (Riksskatteverket) to identify the ones who, by the PRCF, are reported to live in Sweden or abroad. Riksskatteverket was able to identify in total 233 individuals in our sample as living in Sweden as of August 2005.

[21] The sample does not contain any siblings due to the sampling method, i.e. we drew every twentieth card from the alphabetically ordered card register.

[22] The population registers and identification codes where introduced in Finland during the last years of the 1960s. This means that persons who had died or changed citizenship pre-1970 were not identified. In Table 12 in Appendix B I present results from a regression where identification is regressed on background characteristics of the evacuee. The only characteristics in the regression that significantly explain identificaiton are: being female and native Swedish speaking.

Another potential source of selection needs to be discussed, namely the adoption of evacuees by their foster families. In total 5,380 evacuees belonging to our base population were adopted after the war by their foster parents. However, the evacuee card register, from where the sample is drawn, only includes those children who were returned to their biological families after the war. A random sample ($n = 120$) of the adopted evacuees' evacuee cards (stored in a separate register at the National Archives of Finland) shows that they were on average from lower socioeconomic background, and more often from shattered families than their fellow evacuees who returned home.[23] Without information on these children's foster family characteristics, or on their outcomes, it is however hard to draw conclusions about the direction of the bias that this selection in our data may cause.

In total, 1,145 individuals from the evacuee card sample, residing in Finland or Sweden as of summer 2005, were identified for the second stage of the data collection, that is, the survey. Through the survey we obtained information on the rearing environment and the biological family, the evacuation, foster family background, the returning to Finland, assimilation to the biological families, and information on the person's outcomes, such as education, labor market outcomes, physical health, mental health, and risky behavior.[24] The survey was conducted in September 2005. After a second reminder, 752 questionnaires were returned for a response rate of roughly 65 percent.[25]

Table 13 in Appendix B presents the results from a probit regression of a dummy for responding against background characteristics. Most coefficients are small and statistically insignificant. The positive and significant coefficient of the female dummy points towards a gender response bias since roughly equal proportion of both genders were evacuated. The only other background characteristic that has a significant coefficient is a dummy whether the child was exposed to air raids pre-intervention,

[23] See Table 14 in Appendix C for a comparison of the background characteristics between the sample of returned evacuees and the adopted children.

[24] Of the 1931 persons who were included in the original sample we were unable to identify those, who had died or changed citizenship before the creation of a social security number system in the end of 1960s in Finland. The identified but excluded groups are those who had died before Fall 2005 and those who resided outside Finland or Sweden at that time.

[25] No money or financial compensation was offered in return for responding.

which is probably a symptom of an omitted variable. Although the response rate bias of my coefficients is likely to be modest due to the high response rate, I address the potential problem by weighting observations by the odds ratio of response probability, as proposed by Wooldridge (1999).[26]

A control group for the evacuee sample was constructed by exact matching (based on municipality of birth, age (within one month), gender and native language) of controls to each by the PRCF identified child evacuee. An identical questionnaire was sent to these persons, only excluding the questions regarding the evacuation. Of the 3,711 controls that were identified and sent a questionnaire, we received 1,991 questionnaires (response rate 53.7 percent). The possibility that a respondent of the comparison group belonged to the child evacuee population, was accounted for by including a question asking whether the person had participated in the evacuation program. Of the 242 comparison units, that responded positively on that evacuee question, 171 were identified from the aforementioned evacuee card register and were sent an identical evacuee questionnaire as was sent to the original evacuee sample.[27] Out of these, 135 persons responded to the survey. As these treated individuals belong to the same population from where the original treated sample was drawn, and were also randomly drawn, we add these observations to the original sample of child evacuees.

We thus obtain a sample of 887 individuals from the population of persons who participated in the child evacuation scheme. We also have 1,749 controls, for whom we included in the survey the same preintervention questions that are contained in the evacuee card. By definition we lack register data dating back to WWII for these comparison units as all pre-intervention register data come from the aforementioned evacuee cards. However, we have included in the control group's survey questionnaire the same pre-intervention questions as are contained in the evacuee card.

[26]The correction for the nonresponse bias delivers almost identical results as when ignoring the correction bias. The results of the estimations including response bias correction are available from the author.

[27]The somewhat low rate of identification (171/242) is not an indication of incomplete registers. The unidentified are likely to belong to the roughly 25,000 children who were evacuated outside the official program through family or personal connections.

The rich information contained in the evacuee cards drawn from the index allows us to obtain a good picture of the child's pre-intervention environment. By complementing this data with information on the family background, and characteristics of the foster family collected with the survey, we are able to form an almost complete view of the socioeconomic environment of both the biological family and the foster family. The outcome variables come mainly from the survey, except marital status info that come from the PRCF. The outcome variables are described in Appendix A. Also demographic data, e.g. place of birth, age, gender, and mother tongue are drawn from PRCF to double check the accuracy of the evacuee card data and to fill in missing values on demographic variables. By combining register data on each individual to the survey response I can relate the variation in later life outcomes to an exogenously generated variation in the rearing environment. I can further compare the importance of the temporary randomly assigned rearing environment with the importance of the biological pre- and postbirth environment.

3.3 Measures

As all data on the foster family's socioeconomic background is derived from the survey, I am limited to measures that can be constructed based on retrospective questions regarding foster parents' socioeconomic status during the respondent's childhood. My primary measure for socioeconomic status of the family will be based on occupation. The motivation for an occupation-based measure is that parent's occupation is arguably easier to recall for the respondents, than parent's income or education. This holds, in particular, for questions on socioeconomic characteristics of the foster parent's, with whom respondents only spent part of their childhood (and have not necessarily been in touch with since the war ended). A check for how accurately the survey respondents recall their parents' occupation is done by comparing twenty randomly drawn survey answers to pre-intervention data, obtained from the evacuee cards, on biological father's occupation. In all eighteen cases out of twenty where biological father's occupation was non-missing, the survey responses matched

the evacuee card information.[28] I use survey data on occupation for both foster parents and biological parents and substitute missing values on biological parents occupation with parents occupations as reported on the evacuee card.

The occupation-based measure I use for socioeconomic status of the family is the father's socioeconomic status, as measured by an occupation-based socioeconomic status scale that goes under the name International Socio-Economic Index of occupational status (SEI).[29] The SEI scale is derived by Ganzeboom et al. (1992) in the spirit of the Duncan index, (Duncan 1961), so that a weighted sum of the average schooling and the average income levels of occupations (coded into the International Standard Classification of Occupations (ISCO) released by International Labor Organization (ILO) in 1968, from now on, ISCO68), determine their status. The weights are obtained by relating occupational status to education, and occupational status to income. These weights are then used to produce scaled scores for each occupation group so as to maximize the indirect effect of education on income (via occupation) and minimize the direct effect. Occupation is thus treated as an intervening variable, that converts education into income. The scores are rescaled to a range from sixteen to ninety.[30] In the survey, the respondents are asked to report both their foster and biological parent's occupations. These are coded into the 1,506 occupational categories of the four-digit ISCO68.

To make sure that the balancing of socioeconomic status of the foster father is not driven by the coding of SEI, I additionally make use of a discrete class scheme established by Erikson, Goldthorpe and Portocarero (1979), and generally referred to as the EGP-scheme, for the balancing test in section (4). In this eleven-category (also commonly collapsed into seven categories) classification, class is determined by

[28] I drew 20 observations with alphabetically consecutive family names from my sample starting from a randomly drawn family name.

[29] The proportion of married mother's participating in the labor force was still low before WWII in both Finland and Sweden. Of the 709 respondents who reported non-missing values for foster mother's occupation, 496 had had foster mother's who were either "wives of a farmer" or "housewife" by occupation. The biological mothers where more actively participating in the labor force, probably because of the war, with 349 out of 874 biological mothers being housewives.

[30] See Ganzeboom, De Graaf and Treiman (1992) for the algoritm for estimating their scaled SEI variable.

the employment relation, and occupation (based on the ISCO68 occupational code) is used as an indicator of that relation. I apply the standard module generated by De Graaf et al. (1989) to derive the EGP-categories from ISCO68. A frequency table for the categories is shown in Table I. Even though nominal by its construction, this scheme has been widely applied in the intergenerational mobility literature since the 1980s.[31]

My primary dependent variable is schooling, which is measured as the discrete choice of school track. At the time around WWII Finland had a two-track school system. In this system, cohorts attended uniform education only the first four grades of primary school, after which they were divided into two tracks that differed both in terms of content of education, as well as eligibility to further education.[32] School starting age was seven years so the tracking was made at eleven years of age. The choice the children were faced with after fourth grade of primary school was whether to apply to general secondary school, or to continue in primary school. Those who continued another four years in primary school, were restricted to vocational professions, whereas those who were admitted to junior secondary school often continued to upper secondary school and were eligible to apply to university-level studies. If family background has a greater impact on early education choices than those taking place during late adolescence, then the school track choice in the Finnish two-track system should be a good outcome variable when estimating nurture effects in intergenerational transmission of socioeconomic status. There is recent evidence by Pekkarinen, Uusitalo, and Kerr (2009) from Finland, that early tracking to academic and vocational secondary education created stronger earnings correlations between fathers and sons than when the tracking was postponed until after grade nine, as a result of the Finnish comprehensive school reform in the early 1970s. In Sweden pupils attended primary school until the sixth grade and school starting age was seven years, implying that no evacuees had been tracked in Sweden before returning to Finland.

[31] See Erikson and Goldthorpe (2002) for a review of the use of EGP-scheme in the intergenerational mobility literature.

[32] The comprehensive school reform was implemented between 1972 and 1977, imposing a uniform academic curriculum for the entire cohort until age 16 .

3.4 Descriptive Statistics

As described above, I have register data available from the pre-intervention period from the evacuee cards filed before the evacuation and complemented during the evacuation. The evacuee card contains only sparse information about the foster parents, in fact, only their name and address was filed after information about the placement was received.[33] I thus use survey data for variables measuring foster parent's characteristics. All outcomes also come from the survey.

The Support Committee did not state any guidelines with respect to socioeconomic status or family arrangements of the foster family. Eighty-seven percent of the foster families were nuclear families and 65 percent of the evacuees had foster siblings who were biological children of the foster family. Table 1 reports the frequencies of children by family background (both biological and foster family background), categorized according to the EGP-scheme. Despite the lack of formal requirements on foster families, two socioeconomic classes were clearly overrepresented among the foster parents, namely farmers and professionals. This tendency is probably explained by the fact that farmers, even though at the lower end of the socioeconomic status scale, usually had more spacious dwellings, and were likely to have been less affected by the scarcity of groceries that also affected Sweden during WWII as the import of goods basically ceased in 1940. The professionals, i.e. the higher social classes, had the best economic resources to care for an additional family member. Interestingly though, 17 percent of the children were placed in working class families (i.e. where the father was a manual worker). Quite naturally, since low income was a legitimate reason to send ones child away, the distribution in socioeconomic background of the biological families is heavily skewed towards the lower end of the socioeconomic status scale, with 60 percent of the children coming from families where the father was a manual worker. Farmers were clearly less inclined to send their children away on the Finnish side than they were willing to receive the children on the Swedish side. This is probably due to the fact that, also on the Finnish side, farms were less affected

[33] Only in rare cases is there the information on foster father's occupation, usually in the form of a title in front of his name, the occurrance being highly biased towards professional titles such as Dr., Professor, or Member of Parliament.

by rationing. Even though almost all farming men were away from home serving their country for the chief part of the year, farms were virtually self-sufficient with respect to food. Also, one of the major threats that led the city population to send their children away was the fear for air raids. This was less of a threat in sparsely populated areas.

Table 2 reports the means and standard deviations of the main variables used in the analysis. I report these statistics for both the treated group and the control group, and for a subsample of interest: the child evacuees with no command of Swedish by the time of departure. The first panel of Table 2 reveals that the evacuees fare only slightly worse than the control group with respect to later life outcomes. The only exception is occurrence of cardiovascular disease, which is less prevalent in the treated sample.

The second panel shows that women outnumber men in all samples, which in fact represents differences in response rates, since equal proportion of boys and girls participated in the program (and the controls are matched based on demographic variables such as gender). The evacuees where on average five and a half years of age at the time of evacuation and spent slightly more than two years with their Swedish foster families. The ones who were native Finnish-speaking acquired a sufficient command in Swedish within eight months (although this variable is right skewed with the median learning Swedish within four months).

The third panel reports the characteristics of the biological family. It reveals that evacuees came from families of slightly lower social class but the substantial difference lies in how the war had affected the families pre-intervention. The evacuees came more often from shattered families and families that had fled from their homes.

The fourth panel presents the foster family characteristics. As already discussed earlier, the foster family had a higher socioeconomic status than the biological family of the evacuee, but did not differ much from the biological families of the control units with respect to socioeconomic status.

4 Empirically Testing for Random Assignment

I examine whether our data support the anecdotal evidence on the random assignment of evacuees to Swedish foster families, in particular with respect to family background. More exactly, I test whether there is a significant relationship between the socioeconomic status of the biological family and the socioeconomic status of the foster family. This is done by regressing F_i against B_i. If the assignment was indeed random, I would expect to find no significant relationship between the two variables. The measure used for socioeconomic status is the aforementioned SEI-score, i.e. a score on a socioeconomic index (SEI). I also test the random assignment using the EGP-scheme as a measure of socioeconomic background (making the simplifying assumption that the classes are ordinal and continuous). Note that a no-relation result does not entirely exclude nonrandom assignment, as the randomness assumption requires there to be no correlation between both observable and unobservable characteristics determining socioeconomic status of both families. I argue however, that the absence of a significant relationship between my measure of the socioeconomic status of the biological and foster fathers suggests that it is unlikely that such a relationship exists between the unobserved characteristics.

Table 3 presents the results from the balancing test. Each entry in the table represents the estimate for the effect of B_i on F_i for a particular specification. I have experimented with different specifications controlling for factors potentially causing omitted variable bias in the basic bivariate regression reported in the first row of column (1).

The first column reveals that the some positive association exists between the socioeconomic status of the foster father and that of the biological father. The magnitude of the coefficient in e.g. the first row of column (1) should be interpreted as a one unit increase in the biological father's score of the socioeconomic index leading to a 0.159 unit increase in the social status score of the foster father. Although small, the positive and significant association contradicts the rather convincing anecdotal evidence. One possibility is that the order of arrival plays a role in the association. More specifically, Finnish parents of higher socioeconomic status may have found out

about the program earlier and sent their children within the first contingents. In Sweden, families of higher socioeconomic status may have been the ones to first volunteer for the program. These both aforementioned patterns coinciding would have lead to positive sorting of families. To control for this, I include in the regression reported in column (2) a running number representing the order of evacuation as measured by day of departure. This does not remove the positive association. When looking separately at the non-Swedish speakers[34] in column (3), the positive association is mitigated substantially and not significant at conventional levels. The anecdotal evidence does not suggest among the native Swedish-speaking children. However, running the identical univariate regression for the native Swedish speaking subsample produces highly significant positive estimates, suggesting positive sorting within the Swedish subsample.[35] The probably most plausible explanation for this sorting is fairly mundane: The Swedish-speaking evacuees were simply able to communicate with the ones who came to pick them up during the occasion when the assignment to families took place. Swedish foster parents-to-be could ask the Swedish speaking children questions such as "What does your father do for living?" or "What city are you from?". Communication, even though limited, with the Swedish speaking evacuees is likely to have enabled the foster parents to form a view of the children's background.

In the strict sense of a natural experiment, where a process generating exogenous variation in the key variable is required, I am referred to the subset of the sample used in column (3). Due to the limited number of treated observations in my sample I alternatively control for the sorting by including the variable causing the sorting and its interaction with foster father's socioeconomic status in the estimating equation. In column (4) it is shown that this method produces essentially the same association estimates between the biological father's socioeconomic status and foster father's socioeconomic status.

[34] As reported in the evacuee card under the question "has a command in Swedish" correlating strongly with the observations' official mother toungue as reported in the population register.

[35] The bivariate balancing regression produces the following estimate ($t-statistic$) for the Swedish speaking subsample : 0.44 (4.97)

5 Estimating Intergenerational Transmission Coefficients

5.1 Benchmark Specification: The Importance of Family Environment on School Track Choice

Table 4 shows the results from the benchmark transmission model for how an occupation-based measure for a family's socioeconomic status affects the individual's school track choice. The coefficients are essentially transmission coefficients from parents to children as in (1). As already discussed in section (3.1), transmission coefficients should not necessarily be interpreted as the causal impact of parental input, in this study measured by socioeconomic status. However, as Sacerdote (2007) notes, transmission coefficients "...are a convenient and standard way to measure how changes in the child's outcome are associated with changes in the parental characteristics."

In columns (1) to (4) I begin by estimating the benchmark model using the subsample for which I find empirical support for random assignment. The entries in column (2) are the marginal effects of the probit model reported in column (1), evaluated at the means of the independent variables. The marginal effects imply that, for foster father's SEI, a ten score increase on the SEI-scale leads, in the "random assignment" subsample, to a 2.8 percent higher probability of going to secondary school after fourth grade of elementary school.[36] A one standard deviation (13.9) increase in foster father's SEI (e.g. a move up from being a sheet metal worker to becoming an office clerk), implies a 3.8 percent increase in probability of going to secondary school. The coefficients reported in column (3) are obtained by estimating the benchmark model using a linear probability model (LPM). The coefficients in the first three columns tell a fairly consistent story, the size of the coefficients are almost the same across the two models (probit and LPM) and the significance levels for the variables are also very similar. I will in the remainder of the analysis employ LPM for discrete models and OLS where appropriate.

[36] I will in the following use SEI when referring to the variable used as measure for the family's socioeconomic status (SES).

As discussed in section (3.1), the length of stay with the foster parents may interact with the foster family input. Also the variation in age at evacuation may lead to differential responses to the foster family input. A third factor that is important to control for is the potential differential treatment of foster children and their foster siblings by the foster parents. If one wants to relate the coefficient for foster family input to the one for biological family input, and further compare the 'decomposed' transmission coefficients of the child evacuee sample to the transmission coefficient of the control group, it is important that the family input was the same for all children. In order to get at the part of the variation of the foster family input that is not attributed to omitted variables, I control for all three aforementioned factors by including measures of them and interacting these with foster father's SEI. As the coefficients reported in column (4) show, the results are fairly robust to controlling for these factors. The interaction terms between foster father's SEI and the duration of the stay (duration of treatment), and between foster father's SEI and age at evacuation are not significant. The interaction term between foster father's SEI and a dummy whether the foster parents differentiated between foster children and foster siblings is positive and significant at 1 percent significance level.

There is in fact little previous evidence that would point towards any interaction between parental input and age at evacuation. Age is however a classical factor in adoption research and several adoption studies find that adoptees with long lasting pre-adoption adversity are likely to face delays in their cognitive and psychological development (Rutter and ERA Study Team 1998; O'Connor et al. 2000). Also a Swedish study using register data reports that an adoption age above four years is associated with lower schooling (Lindblad et al. 2003). Researchers differ in their views on the mechanisms by which age may mediate the effect of adoptive family input (or, drawing parallels between the adoption literature and this study, also foster family input) on later life outcomes. May it suffice to say that some of the adoption research interprets age at adoption in an interpersonal context. This means that a high age at adoption means a worse chance of developing a close relationship with the caregiver. Another channel is the already mentioned length of exposure to

adversity before adoption takes place.[37]

When dealing with such a small sample as the one at hand, one is reluctant to drop any observations. One way of departing from the non-sorting condition while still maintaining random variation is to control for the factor causing the sorting, in this case, most likely, the ability of the foster parents-to-be to communicate with the evacuees at the absorption center through a spoken language in common. In the estimations reported in columns (5) and (6), I include a dummy indicating whether the children had a command in Swedish pre-intervention, and interact it with foster father's SEI. The results are, as expected, essentially the same as in columns (3) and (4).

Table 5 estimates the intergenerational transmission coefficient for the control group. A one standard deviation increase in biological father's SEI (16.91) increases the probability of going to secondary school by 13.7 percent. I would, in line with BLP, like to find, that the sum of coefficients for the foster family input and the biological family input obtained in Table 4 equal the size of the coefficient for the biological family input obtained in the regressions reported in Table 5. This would provide support for the hypothesis that the prebirth and postbirth inputs add up in a linearly additive way.[38] Instead, I find that the size of the coefficient for biological family input of the control group (column (1) of Table 5) is essentially in the same ball park as size of the coefficient for biological family input in the estimations for the evacuees in Table 4. In column (2) of Table 5 I run the same regression but control for the possibility that the control unit was evacuated domestically from her parents for some period during WWII, e.g. to relatives or friends, and thus experienced a switch in the rearing environment. The results are robust to including this control variable.

[37] See Odenstad et al. (2008) for a discussion of mechanisms through which age at adoption may mediate the effect of adoptive parent's input on adoptive outcomes.

[38] Nonlinear intergenerational transmission is tested in Table 10. The results do not provide support for any nonlinearities in the intergenerational transmission.

5.2 Additional Outcomes

As already mentioned, it is difficult to separate out the causes that affect outcomes by using broad measures such as a socioeconomic index as explanatory variable. Nonetheless, Table 6 and Table 7 present the results of the transmission coefficient model for a list of outcomes to try to shed light on which outcomes are affected by the childhood family environment. The outcomes are recorded at the time of completing the survey, i.e. 2005, when respondents were in their early 60s and some even early 70s. The long time span and consequently small truncation is unusual; to my knowledge only one study has looked at early environment outcomes for individuals up to 60 years later (Gould, Lavy, and Paserman 2009).

5.2.1 Transmission Coefficients for Labor Market Outcome, Health, and Risky Behavior

Most estimates for foster father's SEI are insignificant in Table 6 and taken on the whole, a mixed picture emerges. In column (1), although imprecise, the positive and relatively large coefficient for foster father's SEI suggests that parental input has positive effect on employment throughout the whole working career.

The physical health related estimates have somewhat strange signs, i.e. negative for foster father's SEI, suggesting that the children's long-run health was negatively affected by placement into families belonging to higher social classes. The estimates are not significantly different from zero, but one observation that could explain the sign of the coefficient is that many children where noticeably fat when returning to Finland. A journalist witnessing the debarkation of a contingent of returning evacuees after a temporary armistice was reached wrote "Without exception, the children had gained weight, some even to the extent that their [biological] parents had difficulties in recognizing their child" [*Uusi Suomi*, October 10, 1942]. The long-run effects of a sudden switch of environment from one where food was scarce and malnutrition common to another where food was relatively abundant, and more so in the higher social classes, may have resulted in more pronounced obesity among the ones who were evacuated to families of higher social class. There is a large empirical

literature on the association between rapid infancy weight gain and later life obesity and most papers find that the association is strong and positive.[39]

Although foster father's SEI does not seem to significantly affect mental health, the coefficients for both foster and biological father SEI have a positive sign, intuitively suggesting that higher family input is associated with less depression and a higher emotional well-being score. I find evidence that foster family input had an adverse effect on risky behavior in the form of excessive smoking. The external validity of this finding is questionable though, as it is most likely explained by differences in smoking culture across social classes in Sweden in the early 1940s.

5.2.2 Transmission Coefficients for Marriage, Divorce, and Fertility Outcomes

I consider two marriage outcomes: having married at least once during ones lifetime and currently divorced. I also look at the fertility measured as the number of children during the entire fecund period, and whether the individual emigrated during his lifetime. The emigration measure suffers from selection since I only look at emigration to Sweden, the destination country for 72 percent of the expatriates in the data. Table 7 presents the estimates on these outcomes. I find mostly insignificant estimates for the family input variables. It is surprising that neither biological nor foster family input have any effect on marriage outcomes. One explanation may be that the measures used do not respond to differences in family input. Instead of "ever married", a perhaps better measure might have been "age at first marriage". Also the low rate of divorce in general within the studied cohorts may lead to imprecise results. I will look at the marriage outcomes separately for the female and male subsamples in section (5.4).

[39] See the evidence on the relationship between rapid infancy weight gain and later risk for obesity summarized in Ong and Loos (2006).

5.3 Transmission Coefficients by Age at Evacuation

Table 8 presents the estimated transmission coefficients separately for two age groups: children below five years at evacuation and children above five year at evacuation. One of the distinguishing features of this study is that I know the precise age at the date of evacuation. This enables me to split the data into subsamples by age at evacuation.

A comparison of the school track choice estimates in columns (1) and (5) shows that the slope coefficient of foster father's SEI is slightly steeper and more importantly, statistically significant for the younger age group when controlling for the duration of evacuation.[40] This is an important result, indicating that the effect of parental input on cognitive outcomes is stronger at early ages. The coefficients of the biological family input are positive and significant but vary somewhat in magnitude between the two groups. The difference in magnitude of the coefficients of biological family input may not be statistically different from zero for the whole population. The finding that parental investment affects cognitive skills, in this case schooling, more at earlier stages of development confirms the theory and evidence for the higher productivity of early life investments shown by Cunha and Heckman (2008) and in fact several other studies.[41]

Higher foster family input seems to be associated with poorer later life health for the group evacuated at a younger age. Although the health estimates are not significant in either group, the coefficients are substantial in magnitude and relatively close to being statistically significant at conventional levels for the group evacuated at a younger age. This finding gives further support for the consideration that a sudden positive nutrition shock and the consequent fast weight gain may have had an adverse effect on the younger evacuees' later life health. The finding that health estimates for biological parental input are not statistically different from zero in either age group is further evidence that the signs of the health estimates for foster family input are explained by nutrition shocks depending in size on social class of

[40] The younger age group spend on average 27 months in Sweden whereas the older age group spent on average 24 months in Sweden.

[41] See Knudsen *et al.* (2006) for a comprehensive survey and discussion.

the foster family.

In columns (4) and (8), the estimates point towards a negative relationship between foster family input at a early ages of childhood and long term emotional well-being, whereas at later ages of childhood, no such relationship is found. Instead, the ones who are sent at later ages seem to have been positively affected by biological family input (potentially the input that incided at early ages pre-intervention). A picture emerges, showing important associations between family input at early stages of childhood and emotional-well being later in life. It is outside the scope of this study, and the competence of the author, to analyze the mechanisms that may explain the negative sign of foster family input. However, relating to previous economic literature on parental investment early in life, the result that foster family input did not seem to matter for those evacuated at later ages of childhood (and neither did biological family input for those evacuated at early ages) is at odds with the results of Cunha and Heckman (2008) showing that parental investment is more important at later ages of childhood for the development of socio-emotional skills.

5.4 Transmission Coefficients by Gender

Table 9 studies gender differences in response to family input. Comparing the school track choice estimates in columns (1) and (5) show that girls seem to have been more responsive to foster family input than boys with respect to schooling. The labor market outcome in columns (2) and (6) shows that men were less likely to be unemployed as a result of higher foster family input than were women. The unemployment estimate for foster father's SEI is significant both in magnitude and statistically, suggesting that a one standard deviation increase in foster father's SEI leads to a 5.3 percent lower probability of being unemployed for men. The fertility and marriage outcomes show a similar pattern in the subsamples as in the regressions using the whole sample; estimates for parental input are not statistically different from zero.

5.5 Testing Robustness

I have provided evidence for random assignment of the foster children into foster families. Because it is challenging, if not impossible, to empirically exclude nonrandom sorting on all possible factors, I need to assess the robustness of my results against omitted variables. Omitted-variable bias occurs if F_i is correlated with unobserved biological family background characteristics or if B_i is correlated with unobserved foster family background characteristics. A conventional robustness check is to observe whether the coefficient of F_i remains stable to the inclusion of additional biological family background variables. Removing all biological family characteristics from the specification should also not affect the estimates on foster family input. In column (1) (column (2)) of Table 11 I show that the coefficient that corresponds to foster family input does not change noteworthily when excluding (including) biological family characteristics.

6 Conclusions

This paper exploits a unique event in which almost 70,000 Finnish children were evacuated during WWII and randomly placed in Swedish foster families. The operation creates a rare opportunity to study how truly exogenously generated variation in the rearing environment affects both short- and long-run outcomes. Some outcomes are observed more than sixty years after the evacuations took place, providing an unusually long time span for examining effects of early environment on long-run outcomes.

The results suggest that children who were placed in foster families of higher social class were more likely to continue to secondary school. This result is not sensitive to the inclusion of interactions between foster family input and the age at which the child was evacuated. Nonetheless, when splitting the data into subsamples, I find that children who were evacuated at the age of five years or below were more responsive to foster family input than the ones evacuated after they had turned five. This result indicates that children are more responsive to parental input at early

years of childhood with respect to cognitive outcomes. There is also evidence that higher foster family input had a stronger association with school track choice for girls than it had for boys in the subsample regressions by gender. Biological family input (including prebirth and postbirth input) had a strong positive association with probability of going to secondary school for both the whole sample and for the aforementioned subsamples by age and gender. A comparison of the size of coefficients for foster family input (part of postbirth effects) and biological family input (prebirth effects + (total of postbirth effects - foster family input)) suggests that up to 30 percent of parental input is accounted for by foster family input. Thus, the temporary random rearing family environment in foster care in Sweden accounts for a substantial part of the environmental effect on schooling, and more so at early childhood.

There is some evidence that lower foster family input elevated the risk of having been unemployed at some point during career, in particular for men. The physical and mental health estimates for foster family input are not significant, but the signs suggest that there may have been an adverse health effect of being assigned to a foster family of higher social class, perhaps through a larger nutrition shock and resulting rapid weight gain. Emotional well-being is negatively associated with early childhood (five years or below) family input at statistically significant levels and with an economically significant slope coefficient. This negative association suggests that the effect of parental input on cognitive and non-cognitive skills may not always work in the same direction. Long-run outcomes such as lifetime fertility, marriage outcomes and emigration are not affected by the random temporary rearing environment.

The findings in this paper complement the previous literature in three ways. First, by providing evidence that a randomly assigned temporary rearing environment has substantial impact on the individual's schooling, the results confirm the findings of BLP and Sacerdote (2007), i.e. that both prebirth and postbirth components are important. Comparing my results to the ones of these two adoption studies, which find that postbirth environment may account for up to 50 percent of the intergenerational transmission, it is striking how little less the, on average, two year lasting exposure to foster family input accounts for. Second, the finding, that parental input in early

childhood has a more important impact on schooling than at later ages, supports the results of Cunha and Heckman (2008), that early interventions into a child's environmental conditions are more important for a child's cognitive development than interventions at later ages. Third, it shows the importance of early life environment on long-term outcomes, such as labor market outcomes and emotional well-being.

The novel result that this paper provides is that a relatively short lasting switch in the early childhood environment may have substantial effects on later life outcomes. This result has important implications for policy. For instance, early childhood interventions may not necessarily need to be long lasting to have important effects: A two-year intervention into a child's environment may make a substantial difference for the cognitive development of the child.

References

[1] Björklund, Anders, Mikael Lindahl, and Erik Plug "The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data," *Quarterly Journal of Economics*, vol. 121 (2006), 999-1028.

[2] Blau, David, and Janet Currie, "Pre-School, Day Care, and After-School Care: Who's Minding the Kids?" in *Handbook of the Economics of Education, vol. 2,* Eric Hanushek and Finis Welch eds. (North-Holland: Amsterdam, 2006).

[3] Cunha, Flavio and James Heckman, "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Journal of Human Resources*, vol. 43 (2008), 738-782.

[4] Cunha, Flavio, James Heckman, Lance Lochner, and Dimitriy Masterov, "Interpreting the evidence on life cycle skill formation," in *Handbook of the Economics of Education, vol. 1*, Eric Hanushek and Finis Welch eds., (North-Holland: Amsterdam, 2006).

[5] Currie, Janet, "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development," NBER Working Paper Series, No.13987, 2008.

[6] Dearden, Lorraine, Stephen Machin, and Howard Reed, "Intergenerational Mobility in Britain," *Economic Journal*, vol. 107 (1997), 47-66.

[7] Edin, Per-Anders, Peter Fredriksson, and Olof Åslund, "Ethnic Enclaves and the Economic Success of Immigrants - Evidence from a Natural Experiment," *Quarterly Journal of Economics*, vol. 118 (2003), 329-357.

[8] Erikson, Robert, and John Goldthorpe, "Intergenrational inequality: A sociological perspective," *Journal of Economic Perspectives*, vol. 16 (3) (2002), 31-44.

[9] Erikson, Robert, John Goldthorpe, and Lucienne Portocarero "Intergenerational Class Mobility in Three Western European Societies: England, France and Sweden," *British Journal of Sociology,* vol. 30 (1979), 415-441.

[10] Ganzeboom, Harry, Paul De Graaf, Donald Treiman, and Jan De Leeuw, "A Standard Socio-Economic Index of Occupational Status," *Social Science Research*, vol. 21 (1992), 1-56.

[11] Gould, Eric, Victor Lavy, and Daniele Paserman, "Immigrating to Opportunity: Estimating the Effects of School Quality Using a Natural Experiment on Ethiopians in Israel," *Quarterly Journal of Economics*, 119 (2) (2004), 489-526.

[12] ————————, "Sixty Years after the Magic Carpet Ride: The Long-Run Effect of the Early Childhood Environment on Social and Economic Outcomes," NBER Working Paper No. 14884, 2009.

[13] Kling, Jeffrey, Jeffrey Liebman, and Lawrence Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica,* 75 (2007), 83-119.

[14] Knudsen, Eric, James Heckman, Judy Cameron, and Jack Shonkoff, "Economic, Neurobiological, and Behavioral Perspectives on Building America's Future Workforce," *Proceedings*

of the National Academy of Sciences, 103(27) (2006), 10155-10162.

[15] Lomu, Juhani, "Lastensiirtokomitea ja sen arkisto 1941-1949 (Engl. The Child Evacuation Comittee and its Archives 1941-1949)," Publication Series of the National Archives, Finland., No. 441:5a (1974)

[16] Lindblad Frank, Anders Hjern, Bo Vinnerljung, "Intercountry adopted children as young adults – a Swedish cohort study," American Journal of Orthopsychiatry 73 (2003), 190–202.

[17] Odenstad Anna, Anders Hjern, Frank Lindblad, Finn Rasmussen, Bo Vinnerljung, and Monica Dalen, "Does age at adoption and geographic origin matter? A national cohort study of cognitive test performance in adult inter-country adoptees," *Psychological Medicine,* 38 (2008), 1803–1814.

[18] O'Connor Thomas, Rutter Michael, Beckett Celia, Keaveney Lisa, Kreppner Jana, and ERA Study Team, "The effects of global severe privation on cognitive competence: extension and longitudinal follow-up," *Child Development,* 71 (2000), 376–390.

[19] Ong, Ken, and Ruth Loos, "Rapid Infancy Weight Gain and Subsequent Obesity: Systematic Reviews and Hopeful Suggestions," *Acta Pædiatrica*, 95 (2006), 904-908.

[20] Oreopolous, Philip, "The Long-Run consequence of Living in a Poor Neighborhood," *Quarterly Journal of Economics*, vol. 118 (2003), 1533-1575.

[21] Pekkarinen, Tuomas, Roope Uusitalo, and Sari Kerr, "Education Policy and Intergenerational Income Mobility: Evidence from the Finnish Comprehensive School Reform," *Journal of Public Economics*, 93 (2009), 965-973.

[22] Plug, Erik, "Estimating the Effect of Mother's Schooling on Children's Schooling Using a Sample of Adoptees," *American Economic Review*, 94 (2004), 358-368.

[23] Rutter Michael, and ERA Study Team (1998). Developmental catch-up, and deficit, following adoption after severe global early privation. Journal of Child Psychology and Psychiatry 39, 465–476.

[24] Sacerdote, Bruce, "How Large Are The Effects from Change in Family Environment? A Study of Korean American Adopters," *The Quarterly Journal of Economics*, 122 (2007), 119-157.

[25] Santavirta, Torsten, "Child Evacuees during WWII: Evaluating the Long-term Consequences of Evacuating Children from the Adversities of War," (2009) mimeo.

[26] Solon, Gary, "Intergenerational Mobility in the Labor Market," in *Handbook of Labor Economics, vol. 3*, Orley Ashenfelter and David Card, eds. (New York: Elsevier Science, 1999).

[27] Ware, John, and Cathy Sherbourne, "The MOS 36-Item Short Form Health Survey (SF-36): I. Conceptual Framework and Item Selection" *Medical Care,* 30 (1992), 473 - 483.

[28] Wooldridge, Jeffrey M., "Asymptotic Properties of Weighted M-Estimators for Variable Probability Samples", *Econometrica,* 117 (1999), 1385-1406.

7 Tables

Table 1: Socioeconomic Class-frequencies

	Foster parents	Rearing parents
Professionals, highe-grade	98 (13.46)	66 (7.65)
Proffessionals, lower-grade and technicians, higher-grade	102 (14.01)	55 (6.37)
Routine non-manual, higher-grade	3 (0.41)	18 (2.09)
Routine non-manual, lower grade	6 (0.82)	12 (1.39)
Small employers	14 (1.92)	8 (0.93)
Self-employed workers (nonprof.)	58 (7.97)	55 (6.37)
Farmers	319 (43.82)	86 (9.97)
Technicians, lower-grade, supervisors of manual workers	12 (1.65)	40 (4.63)
Skilled manual workers	54 (7.42)	276 (31.98)
Nonskilled manual workers	35 (4.81)	223 (25.84)
Agricultural workers	27 (3.71)	24 (2.78)
Observations	728 (100.00)	886 (100.00)

Entries in Table 1 represent the number of families belonging to each socioeconomic group. The fractions are presented in parentheses as percentage of total.

Table 2: Summary Statistics

	Treated group		Control group
	Whole sample	Non-Swedish subsample	
Number of observations	715	598	1721
Outcomes: Schooling			
School track choice	0.34	0.30	0.36
	(0.47)	(0.46)	(0.48)
Outcome: Labor Market			
Unemployed during	0.44	0.46	0.36
working career (yes/no)	(0.50)	(0.50)	(0.48)
Outcomes: Health & Risky Behavior			
Physical health	6.43	6.41	6.24
	(2.25)	(2.22)	(2.23)
Cardiovascular disease	0.29	0.30	0.36
	(0.45)	(0.46)	(0.48)
Beck Depression Inventory	8.78	8.78	8.69
	6.13	(6.15)	(6.52)
Emotional well-being	74.19	74.31	74.86
	(14.62)	(14.56)	(14.15)
Excessive drinking	0.20	0.21	0.17
	(0.40)	(0.41)	(0.37)
Smoking	0.43	0.42	0.33
	(0.50)	(0.49)	(0.47)
Outcomes: Marriage, Fertility, and Emigration			
Married at least once	0.88	0.86	0.92
	(0.33)	(0.34)	(0.27)
Divorced	0.17	0.16	0.15
	(0.37)	(0.37)	(0.36)
Number of children	2.26	2.26	2.19
	(1.33)	(1.39)	(1.21)
Emigrated to Sweden	0.17	0.16	
	(0.37)	(0.37)	
Demographic and Program Characteristics			
Female	0.55	0.55	0.54
	(0.50)	(0.50)	(0.50)
Native language*	0.16		0.113
	(0.37)		(0.32)
Age at evacuation (months)	65.60	65.11	
	(30.74)	(30.50)	

Table 2: Continued

	Treated group		Control group
	Whole sample	Non-Swedish subsample	
Duration of evacuation (months)	26.40	26.86	
	(15.25)	(15.57)	
Evacuated domestically to family			0.42
or friends for a period during WWII			(0.49)
Biological family characteristics			
ISEI-score of biological father	37.67	36.46	40.08
	(14.96)	(13.85)	(16.91)
Household financial	2.74	2.69	2.94
stability (1-5)	(0.80)	(0.80)	(0.78)
Evacuated from war zone*	0.26	0.31	0.25
(Karelian families)	(0.44)	(0.46)	(0.43)
Father died in war*	0.22	0.23	0.07
	(0.41)	(0.42)	(0.25)
Subject to air raids*	0.13	0.14	
	(0.34)	(0.35)	
Foster family characteristics			
ISEI score of foster father	40.49	39.46	
	(18.44)	(13.85)	
Non-differentiation btw	0.92	0.91	
foster child and foster siblings	(0.28)	(0.28)	

Entries in Table 2 represent the means of the relevant variables. Standard errors are in parentheses. For the variables marked with an asterisk, the entries of the treated group are pulled out from the evacuee card, i.e. are pre-intervention register data. The other variables are collected by the survey. In cases where a demographic variable (gender, native language) was missing in the survey we use data from the Populations Register Center in Finland. In rows 'alcohol abuse', 'Age at evacuation', 'Duration of evacuation', 'Household financial stability', 'Evacuated from war zone', 'Father died in war', 'Subject to bombings', and 'Non-differentiation btw foster and own birth children' the number of observations is less than the number reported for each column, namely 577-712 observations for the treated group and 1471-1678 observations for the control group .

Table 3: Test for Balancing - Regressing Foster Father′s Socio-Economic Status on Biological Father′s Socioeconomic Status

	Dependent variable: Foster father's socioeconomic status			
	(1) Whole sample	(2) Whole sample controlling for order of evacuation	(3) Non-Swedish subsample	(4) Whole sample controlling for native language
Biological father's	0.159	0.169	0.029	0.029
SEI (ISEI) (17-90)	(3.49)	(3.61)	(0.54)	(0.55)
Biological father's	0.115	0.118	0.032	0.032
SES (EGP-11)	(3.07)	(3.11)	(0.74)	(0.74)
Biological father's	0.118	0.119	0.042	0.042
SES (EGP-7)	(3.20)	(3.19)	(0.99)	(1.00)
Observations	717	692	600	717

Each entry in the columns of Table 3 represents the coefficient from separate regressions of the foster father's socioeconomic status on the biological father's socioeconomic status. The t-statistics are reported in the parentheses. In column (1) the estimates of a univariate regression are reported. In column (2) I include as a control the order of evacuation measured by the day of departure as reported in the evacuation card stored in the National Archives of Finland. In column (3) the native Swedish speaking evacuees are excluded. In column (4) native language (a dummy indicating if native Swedish speaking) is included and its interaction with biological father's socioeconomic status.

Table 4: Estimated Transmission Coefficients

Dependent variable: School track choice: Track = 0 (primary, civic or vocational school) vs. Track = 1 (secondary school or university)						
	Probit (MLE) (1)	(mfx) (2)	LPM (MLE) (3)	LPM (MLE) (4)	LPM (MLE) (5)	LPM (MLE) (6)
Foster father's SEI	.0082	.0028	.0025	.0020	.0024	.0018
	(2.56)	(2.55)	(2.28)	(2.49)	(2.28)	(2.77)
Biological father's SEI	.0240	.0082	.0080	.0080	.0087	.0082
	(5.32)	(5.29)	(5.42)	(3.41)	(7.04)	(5.34)
Interactions:						
With Forster father's SEI						
Swedish speaking					-	-
Age at evacuation				+		+
Time spent in Sweden				+		+
Non-differentiation btw foster child and foster siblings				+***		+***
With Biological father's SEI						
Time spent in Sweden				-		-
Observations	579		579	543	697	655
Percent correctly predicted	71.5		70.8	72.7	70.0	71.8
Log-likelihood	-320.27		-335.84	-300.28	-413.36	-374.69
Pseudo R-squared	0.110		0.131	0.152	0.159	0.174

The Non-Swedish (pre-intervention information of command in Swedish) speaking subsample is used in the regressions reported in columns (1)-(4) and the whole sample is used in regressions (5)-(6). The entries in column (1) are estimates from a Probit model. The entries in column (2) are the marginal effects of the Probit model of column (1) evaluated at the means of the variables. The entries in columns (3)-(6) represent the coefficients in a linear probability model. In columns (1) and (2) robust $z-$statistics in brackets and in column (3) -(6) robust $t-$statistics. The entries in column (4), (5) and (6) for Foster father's SEI and Biological father's SEI are $\partial P\left(y=1|F,B,C\right)/\partial F$ and $\partial P\left(y=1|F,B,C\right)/\partial B$ respectively, evaluated at means of the interacted variables. All the interacted terms are also included in levels. In columns (4)-(6) the $t-$statistics reported for the marginal effects of Foster father's SEI and Biological father's SEI refer to the coefficients of the level variable. An intercept, a gender dummy, 21 regional dummies, and 18 cohort dummies included in each regression. The sign of the interaction terms is reported and significance is reported with the conventional symbols(* for significant at 10 percent significance level, ** for significant at 5 percent significance level, and *** for 1 percent significance level).

Table 5: Estimated Transmission Coefficients for the Control Group

	Dependent variable: School track choice	
	(1)	(2)
Biological father's SEI	0.0081	0.0092
	(11.92)	(10.30)
Evacuated to family or friends		0.1024
during WWII (EVFAM)		(1.79)
(Biological father's SEI)×EVFAM		-0.0025
		(1.87)
Observations	1686	1686
Percent correctly		
predicted	69.6	77.4
R^2	0.118	0.142

All entries in Table 5 represent coefficients in a linear probability model. Robust $t-$statistics are reported in the parentheses. An intercept, a gender dummy, age, a Swedish speaking dummy, 18 cohort dummies, and 21 regional dummies are included in each regression.

Table 6: Estimated Transmission Coefficients for Evacuee Outcomes

	Unemployed during working career (yes/no) (1)	Physical health (2)	Heart disease (yes/no) (3)	Depression (BDI) (4)	Emotional well-being (5)	Smoking (6)	Excessive Drinking (7)
Foster father's SEI	-0.0016	-0.0070	-0.0002	-0.0039	0.0184	0.0024	-0.0002
	(1.31)	(1.26)	(0.21)	(0.29)	(0.54)	(2.02)	(0.23)
Biological father's SEI	-0.0053	0.0146	0.0002	-0.0368	0.0796	-0.0015	0.0007
	(4.18)	(2.43)	(0.19)	(2.38)	(2.05)	(1.15)	(0.64)
Swedish speaking	-0.1389	-0.8816	0.0691	2.0920	-5.5564	0.1824	0.0774
	(1.01)	(1.25)	(0.59)	(1.13)	(1.30)	(1.33)	(0.73)
Swedish speaking×	0.0012	0.0205	-0.0041	-0.0411	0.0755	-0.0028	-0.0027
Foster fathers SEI	(0.45)	(1.56)	(1.77)	(1.17)	(0.90)	(1.00)	(1.34)
Observations	660	660	660	660	660	660	660
R^2	0.09	0.05	0.09	0.06	0.06	0.13	0.15

All entries in columns (1), (3), (6) and (7) of Table 6 represent the coefficients in a linear probability model and all entries in columns (2), (4) and (5) represent OLS coefficients. Robust $t-$statistics are reported in the parentheses. An intercept, 18 cohort dummies, and 21 regional dummies are included in each regression.

Table 7: Estimated Transmission Coefficients for Marriage, Fertility, and Emigration Outcomes

	Divorced	Ever married	Emigrated to Sweden	Number of children
	(1)	(2)	(3)	(4)
Foster father's SEI	0.0006	-0.0001	-0.00002	0.0003
	(0.58)	(0.08)	(0.03)	(0.11)
Biological father's SEI	-0.0009	0.0001	-0.0015	-0.0033
	(0.82)	(0.09)	(1.73)	(1.01)
Swedish speaking	0.0649	0.0064	0.1717	-0.2099
	(0.53)	(0.10)	(2.22)	(0.73)
Swedish×Foster f's SEI	-0.0014	0.0008	0.0039	0.0073
	(0.56)	(0.65)	(2.04)	(1.23)
Observations	603	668	668	668
R^2	0.06	0.07	0.10	0.09

All entries in columns (1), (2), and (3) in Table 7 represent coefficients in a linear probability model and the entries in column (4) represent OLS coefficients. Robust $t-$ statistics are reported in the parentheses. An intercept, a gender dummy, 18 cohort dummies, and 21 regional dummies are included in each regression.

Table 8: Estimated Transmission Coefficients by Age Group

	Evacuees $\leq$ 5 years at evacuation				Evacuees $>$ 5 years at evacuation			
	School track choice (1)	Heart disease (yes/no) (3)	Physical health (4)	Emotional well-being (2)	School track choice (5)	Heart disease (yes/no) (7)	Physical health (8)	Emotional well-being (6)
Foster father's SEI	0.0026	0.0005	-0.0078	-0.0247	0.0023	-0.0010	-0.0102	0.0267
	(4.16)	(1.49)	(1.26)	(1.68)	(0.78)	(0.60)	(0.18)	(0.09)
Biological father's SEI	0.0088	-0.0009	0.0072	0.0473	0.0060	0.0015	0.0077	0.0266
	(3.32)	(0.35)	(0.36)	(0.35)	(3.05)	(0.23)	(2.24)	(3.50)
Interactions:								
With Forster father's SEI								
Swedish speaking	-	+	+	-	+	-***	+	+
Time spent in Sweden	-	+	-	-	-	+	+	+
Non-differentiation btw foster child and foster siblings	+***	+	+	+***	+***	+	-	-
With Biological father's SEI								
Time spent in Sweden	-	+	+	+	-	+	-	-***
Observations	306	306	306	306	351	351	351	351
R^2	0.25	0.16	0.11	0.14	0.17	0.14	0.12	0.17

All entries in columns (1)-(3) and (5)-(7) of Table 8 represent the coefficients in a linear probability model. All entries in columns (4) and (8) represent OLS coefficients. The entries for Foster father's SEI and Biological father's SEI are $\partial P\left(y=1|F,B,C\right)/\partial F$ and $\partial P\left(y=1|F,B,C\right)/\partial B$ respectively, evaluated at means of the interacted variables. All the interacted terms are also included in levels. Robust $t-$statistics are reported in the parentheses. The $t-$statistic for the marginal effect of Foster father's SEI refer to the coefficient of the level variable. An intercept, a gender dummy, 21 regional dummies, and 18 cohort dummies included in each regression. The sign of the interaction terms is reported and significance is reported with the conventional symbols (* for significant at 10 percent significance level, ** for significant at 5 percent significance level, and *** for 1 percent significance level).

Table 9: Estimated Transmission Coefficients by Gender

	Girl Evacuees				Boy Evacuees			
	School track choice (1)	Unemployed during career (yes/no) (2)	Divorced (3)	Number of Children (4)	School track choice (5)	Unemployed during career (yes/no) (6)	Divorced (7)	Number of Children (8)
Foster father's SEI	0.0038	0.0002	0.0007	-0.0030	0.0022	-0.0038	-0.0006	0.0046
	(2.48)	(0.10)	(0.59)	(0.71)	(1.30)	(1.89)	(-0.45)	(0.84)
Biological father's SEI	0.0053	-0.0050	-0.0002	-0.0018	0.0130	-0.0068	-0.0009	0.0008
	(2.85)	(2.88)	(-0.12)	(0.37)	(7.67)	(3.35)	(-0.70)	(0.16)
Swedish speaking	0.0415	-0.2346	0.0301	-0.4536	0.1293	0.0940	0.0661	0.3820
	(0.21)	(1.27)	(0.17)	(1.25)	(0.62)	(0.44)	(0.42)	(0.71)
Swedish speaking×	0.0010	0.0017	-0.0005	0.0078	-0.0015	-0.0004	-0.0016	-0.0005
Foster father's SEI	(0.26)	(0.50)	(-0.16)	(0.99)	(0.37)	(0.08)	(-0.50)	(0.05)
Observations	371	371	371	371	287	287	287	287
R^2	0.15	0.12	0.07	0.15	0.26	0.16	0.08	0.07

All entries in columns (1)-(3) and (5)-(7) of Table 9 represent the coefficients in a linear probability model. All entries in columns (4) and (8) represent OLS coefficients. Robust $t-$statistics are reported in the parentheses. An intercept, 21 regional dummies, and 18 cohort dummies are included in each regression.

Table 10: Estimated Transmission Coefficients
in Nonlinear Models

	Dependent variable: School track choice	
	(1)	(2)
Foster father's SEI	.0042	.0127
	(1.41)	(1.82)
Biological father's SEI	.0098	.0024
	(2.84)	(0.35)
(Foster father's SEI)^2 /100		-0.090
		(1.29)
(Biological father's SEI)^2 /100		.0081
		(1.03)
(Foster father's SEI×Biological	-.0046	-.0045
father's SEI)/100	(0.61)	(.47)
Observations	580	580
Percent correctly predicted	71.0	71.6
R^2	0.132	0.137

The Non-Swedish (based on pre-intervention information of command in Swedish) speaking subsample is used in the regressions reported in Table 10. All entries in the table represent coefficients in a linear probability model. Robust $t-$statistics are reported in the parentheses. An intercept, a gender dummy, 18 cohort dummies, and 21 regional dummies are included in each regression.

Table 11: Sensitivity Analyses

	Dependent variable: School track choice	
	(1)	(2)
Foster father's SEI	0.0022	0.0019
	(1.99)	(1.75)
Biological father's SEI		0.0079
		(6.22)
Swedish speaker	0.0213	0.0934
	(0.15)	(0.71)
Swedish speaker×	0.0023	-0.0004
Foster father's SEI	(0.95)	(0.16)
Father died in war		-0.0137
		(0.33)
Household financial		0.0715
stability		(3.17)
Evacuated from war		-0.1202
zone		(2.06)
Observations	675	675
R^2	0.097	0.132

All entries in Table 11 represent coefficients in a linear probability model. Robust $t-$statistics are reported in the parentheses. An intercept, a gender dummy, 18 cohort dummies, and 21 regional dummies are included in each regression.

Appendices

A Variable Definitions:

A.1 Outcomes

School track choice

The surveyees are asked to indicate their highest completed education on a five class categorization: 1. Primary school 2. Lower secondary school or civic school 3. Vocational school 4. Upper secondary school 5. Tertiary degree (university) education. Educational attainment is also obtained from census data of Statistics Finland. I collapse the 5 classes into a dummy variable measuring whether the person continued to secondary school after the fourth grade of civic school. Classes 1 and 3 obtain value zero and classes 2, 3 and 4 obtain value one.

Labor market

The surveyees were asked whether they at some point during their professional career had had a spell of unemployment.

Health

Physical health. Self-rated health compared to own age group was measured using a Visual Analogue Scale (VAS) with 0 indicating "very much worse" health compared to own age group, and 10 indicating "very much better health" compared to own age group.

Cardiovascular disease. The surveyees are asked whether they at some point during their life have suffered from any heart problems or cardiovascular disease.

Depressive symptoms. The respondent's depressive symptoms were measured using Beck Depressive Inventory, a 21 question instrument widely used for measuring the severity of depression.

Emotional Well-Being is measured using the items of the Short-Form-36 Health Survey of the RAND Health Insurance Experiment measuring emotional well-being (Ware and Sherbourne 1992).

Risky behavior

Smoking, drinking. The surveyees are asked, separately for each substance, whether they, in the course of their life, have been smoking or misusing alcohol (yes/no).

Marriage outcomes

Divorced. Register data from PRCF is used on the marital status of the individual. All individuals who never married are coded as missing values.

Ever married. Register data from PRCF is used on the marital status of the individual. All individuals who are either married, divorced, or live in separation from their legal husband are coded as married at least once.

Fertility

The surveyees are asked to indicate the number of children they have.

Emigration

The emigration measure is a dummy on whether the individual emigrated to Sweden. Register data from PRCF is used on the country where the individual is registered. All individuals who emigrated before the registers of PRCF were completed in the end of 1960s but were still Finnish citizens by 1970 are included in the registers of PRCF. We identified those, who emigrated to Sweden.

A.2 Independent variables

Household financial stability

The surveyees are asked to indicate on a likert scale from one to five how good the financial situation of the rearing family was pre-intervention.

B Test for Association between Identification for Survey and Positive Response and Backgroud Characteristics

Table 12: Test for Selective Identification of Treated Sample Units by the Population Register Center of Finland

	Dependent variable: dummy for sample unit identified by PRCF
Age at evacuation	0.003
	(0.91)
Swedish speaking	-0.099
	(3.35)
Female	-0.056
	(2.93)
Family evacuated from	-0.025
war zone	(1.04)
Father wounded in	-0.032
war	(0.75)
Father died in war	-0.019
	(0.64)
Out-of-wedlock child	-0.137
	(1.54)
Subject to bombings	0.043
	(1.46)
Family lived in town	-0.008
(countryside=0)	(0.40)
Observations	1912
F-test, background point estimates = 0	26.98
$p > \chi^2$	0.0014

The entries represent the marginal effects of a probit model evaluated at the means of the independent variables. All right hand side variables are pulled from the evacuee cards in the Child evacuee records at the National Archives. The whole sample drawn from the evacuee card register with nonmissing values on the relevant variables is included. A probit model is estimated with the dependent variable taking value one if the observation was identified by the Population Register Center, Finland. The marginal effects are reported evaluated at the mean of age at evaucation. The $t-$statistics are reported in the parentheses.

Table 13: Test for Selective Response to Survey with Respect to Pre-intervention Characteristics

Probit model with dependent variable: dummy for response to survey questionnaire	
Age at evacuation	-0.003
	(0.49)
Swedish speaking	-0.002
	(0.04)
Female	0.074
	(2.59)
Family evacuated from	0.006
war zone	(0.16)
Father wounded in	-0.090
war	(1.42)
Father died in war	-0.018
	(0.40)
Out of wedlock child	-0.004
	(-0.03)
Subject to bombings	0.108
	(2.56)
Family lived in town	0.036
(countryside=0)	(1.16)
Observations	1126
F-test, background point estimates = 0	18.08
$p > \chi^2$	0.054

All right hand side variables are pulled from the evacuee cards in the Child evacuee records at the National Archives. I include all observations that are identified and to whom we have sent the survey questionnaire. A probit model is estimated with the dependent variable taking value one for a positive response on the original survey or the reminder. The marginal effects are reported evaluated at the mean of age at evacuation. The $t-$statistics are reported in the parentheses.

C Background Characteristics of Adopted Evacuees

Table 14: Comparison of Background Characteristics between Sample of Returned Evacuees and Adopted Evacuees

	Children who Returned Mean	Adopted by foster family Mean
Both parents alive=1	.81	.55
	(.40)	(.50)
Family evacuated from	.27	.24
war zone (Karelian)=1	(.44)	(.43)
Single-parent	.01	.09
household=1	(.11)	(.29)
Biological father's SES	7.59	8.81
(EGP-10)	(2.94)	(2.14)
Age at evacuation (months)	65.07	51.19
	(31.31)	(28.70)

Standard deviations reported in the parentheses.