The German Occupation of the Soviet Union:

The Long-Term Health Outcomes

January, 2016

Preliminary and incomplete

Olga Belskaya
UNC-Chapel Hill
belskaya@email.unc.edu

Klara Sabirianova Peter
IZA and CEPR
UNC-Chapel Hill
kpeter@unc.edu

Christian Posso
UNC-Chapel Hill
cpossosu@email.unc.edu

Abstract

This study examines the long-term health consequences of the early-life shocks caused by the Nazi occupation of the Soviet Union during WWII in 1941-1945. We focus on individuals who, at the time of the war, were in utero or in their early childhood until age 5, lived under occupation of Nazi Germany and survived until age 50 when we start observing their health outcomes. The study design relies on the precise timing of occupation, specific geographic location of occupied municipalities, and the unexpected and rapid advancement of the Nazy army into the Soviet Union. We test for the presence of the critical periods in child development affecting late-life health outcomes such as mortality, chronic heart conditions, diseases of respiratory system, diseases of digestive and genitourinary systems, spinal disorders, hypertension, depression, under- or over-weight, as well as subjective health assessments and life satisfaction. The average treatment effect is estimated using a difference-in-difference estimator with a full set of region of birth and birth year fixed effects and with accounting for attrition and selective mortality bias. The spatial regression discontinuity estimator is also implemented. In addition to the average treatment effect, we recover the heterogeneity of treatment by gender, by the length of individual exposure to shocks (from one month to a few years) and by the regional sub-division defined based on the level of destruction and human losses. Our preliminary estimates show that the German occupation of the Soviet Union had a large detrimental effect on health of survived children of the war, and that the effects are the largest if the exposure to shocks occurred in utero or in infancy.

Keywords: health, early childhood shocks, difference-in-difference, survival, WWII, Russia, Ukraine

JEL Code:

Acknowledgement:
We are thankful to Michele Mendez for useful comments. We also acknowledge helpful comments received from participants of the UNC Applied Microeconomics Workshop and Carolina Population Center Seminar Series.

Web appendix:

1. Introduction

There is a growing interest in understanding of how the formation of human capabilities in early stages of life influences adult outcomes (Conti and Heckman, 2010). Recent economics literature has built a theoretical life-cycle framework that draws attention to mechanisms through which the changes in child capabilities and investments in child development determine future life outcomes (see Cunha and Heckman, 2007; Heckman, 2007, Almond and Currie, 2011a and 2011b). Under this framework, the literature started seeking evidence on the role of early childhood in shaping late-life consequences by studying unexpected, disastrous events, such as wars and famines, experienced early in life.[1]

Our study examines the long-term impact of early life shocks from the World War II (WWII) using the case of the least studied and most harmed side of the conflict – the former Union of Soviet Socialist Republics (USSR or simply Soviet Union). Specifically, we focus on individuals who, at the time of the war, were in utero or in their early childhood, lived under occupation of Nazi Germany and survived until age 50 when we start observing their health outcomes.

Despite the non-aggression Molotov-Ribbentrop pact signed between the Soviet Union and Germany in 1939, Nazi Germany attacked on June 22, 1941 and within the next few months occupied Belarus, Moldova, Ukraine, Baltic republics, and a significant part of the central and southern Russia. Almost 85 million people or 44.5 percent of pre-war population of USSR lived in the territories that were occupied during WWII (Goskomstat, 2015). The Soviet Union suffered immensely. In just four years, 1941-1945, around 8.7 million Red Army soldiers died in the course of the war, 11.3 million civilian individuals died because of the military activity and crimes against humanity, 6.5 million Soviet citizens died because of the war-related famine and infectious diseases. In total, the Soviet Union lost 26.5 million people in World War II, or 15.5 percent of its 1939 population (Ellman and Maksudov, 1994). Based on historic evidence, we conjecture that Nazi occupation may have had a long-lasting effect on health outcomes of survived children of the war through a variety of channels, including malnutrition, diseases, stress, the reduced parental investment in child development, the lack of maternal and health care, and the destruction of housing and other forms of material wealth.

The occupation of the Soviet Union during WWII presents a unique quasi-experimental setup with several important properties. First, the occupation was sharply defined in time and space. The occupation borderline shown in Figure 1 splits the country roughly in half in terms of the size of pre-war population. Second, there is a significant spatial heterogeneity in terms of the timing and duration of occupation, which also can be seen in Figure 1. The duration of occupation varied anywhere from 3 days to 1413 days, with the mean of 690 days across municipalities. Third, occupation was not anticipated by the general population. Because WWII started unexpectedly and the German occupation occurred rapidly, most people could neither escape from its harsh conditions, nor prepare for it. Finally, living conditions on occupied territories were so severe that survived children are likely to carry the permanent impact of this adverse period to their adulthood.

By combining this quasi-experimental setup with unique data sources and rigorous empirical analysis, our study makes several important contributions to the existing literature on this topic. To start with, this is the first econometric study to evaluate the long-term health consequences of the negative shocks associated with the eastern front of the WWII. No other study has done it despite

---

[1] For example, previous studies evaluated the long-term impact of the Dutch famine of 1944 on numerous health outcomes (Lumey *et al.*, 2007), famine in Europe during WWII on height (van der Berg *et al.*, 2015), WWII in Europe on health (Kesternich *et al.*, 2014), bombing and destruction of infrastructure in Germany during WWII on height, mortality, obesity and chronic diseases (Akbulut-Yuksel, 2014 and 2015), the Korean War of 1950-1953 on disability and mortality (Lee, 2014), the Chinese famine of 1959-1961 on height (Chen and Zhou, 2007; Gorgens *et al.*, 2012; Meng and Qian, 2009), obesity (Fung, 2009), disability (Almond *et al.*, 2010), and mental disorders (St. Clair *et al.*, 2005), the Nigerian Civil War of 1967-1970 on height (Akresh, 2011), and the Ethiopian famine of 1984 on height (Dercon and Porter, 2010).

the fact that "the war assumed a far grander scale in the East than in any of the fronts where the Western Allies were involved" (Davies, 2006).

We utilize the detailed nationally representative surveys of the two largest countries of the Soviet Union, Russia and Ukraine,[2] and link these surveys with the novel dataset on the dates of occupation and liberation of USSR municipalities and also with the regional measures of destruction and human losses during WWII. The two survey sources have never been previously used in the WWII context, while the municipality-level data have been collected by the authors in the course of this research project.

In our identification strategy, we rely on the individual date of birth, municipality of birth (city or county), and exogenous variation in the location and timing of the German occupation of each municipality during WWII. Thus, we measure an *individual-specific exposure* to treatment at the highly disaggregate level, which is an improvement over previous studies that identify treatment at the country level (Havari and Peracchi, 2015; Kesternich *et al.,* 2014; Van ben Berg *et al.,* 2015, using the survey of European countries SHARELIFE) or region/province level (Akbulut, 2014, using German Socio-Economic Panel). Another important feature of our study design is a control group comprising of individuals born in non-occupied municipalities during the war as well as individuals born after the war in formerly occupied and non-occupied municipalities. This allows us to explore the disaggregated spatial and temporal variation in treatment simultaneously.

This study also contributes to the growing literature that compares the long-term effects of shocks across *different stages of childhood*. Most of the previous studies focus on specific stages, such as in-utero (Lee, 2014), infancy (Ampaabeng and Tan, 2013; Portrait *et al*., 2005), or children of school age (Akbulut, 2014). We distinguish four different stages: in-utero (from conception to birth), infants (from birth to 12 months of age), toddlers (1-3 years old), and preschoolers (3-5 years old). This allows us to test for the presence of the critical periods in child development affecting long-term health. We find astonishing heterogeneity in the treatment effect by age, with kids exposed to adverse events during the prenatal stage of life enduring the largest negative consequences.

Unlike many existing studies, our measure of the intensity of health shock varies with the date of conception and birth, which, together with the dates of occupation, determine the *length of individual exposure* to hardships in the early childhood under Nazi occupation. Additionally, we recover the heterogeneity of treatment by gender and by the regional sub-division defined based on the level of destruction and human losses.

2. Background

On June 22, 1941, Nazi Germany launches Operation Barbarossa – an invasion of the Soviet Union by three million German soldiers and half a million troops from German allies (Finland, Romania, Hungary, Italy, Slovakia, and Croatia) along a 1,800 miles front extending from the Arctic Ocean to the Black Sea. Barbarossa was the largest military operation in the world history in both manpower and casualties, and regions affected by Barbarossa were the site of some of the largest battles, highest casualties, and most horrific conditions. German forces were organized into four armies: army Norway operated in far northern Scandinavia; army North aimed at taking or destroying the city of Leningrad (now St. Petersburg) by marching through the Baltics into northern Russia; army Center was planned to march through present day Belarus and take Moscow; and army South aimed at striking Ukraine and taking the control over the oil-rich Caucasus. The main objective was to reach the so called A-A line (Arkhangelsk-Astrakhan):

"*In quick pursuit a line is then to be reached from which the Russian Air Force will no longer be able to attack the territory of the German Reich. The ultimate objective of the operation is to establish*

[2] The Russian Longitudinal Monitoring Survey – Higher School of Economics (RLMS-HSE) and the Ukrainian Longitudinal Monitoring Survey (ULMS), respectively.

a cover against Asiatic Russia from the general line Volga-Archangel." (German Archive, Hitler's Directive No. 21 Operation Barbarossa, December 18, 1940)

By early September 1941, German forces encircled Leningrad in the north and took Smolensk in the center and Dnepropetrovsk in the south. In early December 1941, they came close to the city of Moscow but were quickly driven away from the country capital by the Red Army. In the summer of 1942, German forces attacked in the direction of the city of Stalingrad (now Volgograd) on the Volga River and the oil fields of the Caucasus. Thus, in September 1942, the German occupation of the Soviet Union reached its furthest geographical extension, which was still far away from the desired A-A line, as we can see from Figure 1. The end of 1942 became the turning point of the German invasion of the Soviet Union. Six weeks of ferocious combat near the city of Stalingrad led to heavy casualties on both sides, the major victory of the Red Army, and the beginning of liberation of the Soviet territories. During 1943, the Red Army cleared German forces from most of the territory of Russia, eastern Belarus and Ukraine. In 1944, the western parts of Belarus and Ukraine, the northwestern part of Russia, Moldova, and most of the Baltic countries were liberated.

The German occupation of the Soviet Union was the most brutal and destructive occupation in Europe. According to Nazi plans, Jewish population and members of the Communist Party were to be annihilated, Slavic population significantly reduced as part of the Hunger Plan, and Soviet economic resources destroyed or expropriated. We have already mentioned the immense loss of many millions of lives. In addition, during the occupation of the Soviet Union, the Nazi invaders totally or partially ruined and burned 1710 cities and townships and more than 70,000 villages, burned and destroyed more than 6 million buildings and 31,850 industrial enterprises, rendered around 25 million persons homeless, and caused great damage to the infrastructure and public services (Goskomstat, 2015; most numbers are taken from the Nuremberg Trial).

Here, we only highlight some of the main channels through which the occupation may have had an effect on the long-term health outcomes of survived children of the war.

Malnutrition – According to the "Hunger Plan" developed by Germans in early 1941 and subsequently implemented, the Soviet Union was divided into two agricultural zones – a deficit zone (Belarus and Northern and Central Russia) and a surplus zone (Ukraine, Southern Russia and the Caucuses). Germany extracted food products from the surplus zone and sealed off the deficit zone from getting food supplies (Gerhard, 2009). Large industrial and urban centers in the surplus zone were also cut off from food supplies. Thus, famine on a mass scale was unavoidable and engineered with expectations that tens of millions of Soviet people would become superfluous and die from starvation. As a result of inadequate food rationing and blockading food supplies, urban civilian population plummeted.[3] Among survivors, the short-term health effects from undernourishment were immediately noticeable. For example, in Leningrad, the average birthweight was 100 grams less in 1941 compared to the birthweight in 1938-1940. In 1942, the birthweight declined further by 600 grams, average height – by 2 cm, chest circumference – by 1.5 cm, and head circumference – by 1.3 cm (Georgievsky and Gavrilov, 1975). These and other changes in human body during prenatal and early childhood stages of life may have produced long-lasting effects on human capabilities in adulthood and influenced subsequent life quality and duration. Unfortunately, scientific evaluation of the long-term health effects of the German hunger policy in the USSR has never been attempted on a large country scale.[4]

---

[3] Sieged for 872 days between September 1941 and January 1944, Leningrad experienced the longest and the most destructive sieges in the world history leading to 1,500,000 deaths due to the famine and artillery shelling (Salisbury 1969). Population of the Ukrainian city of Kharkov dropped from a million before the war to 250,000 two years later, which cannot be attributed to migration as entering and leaving the town was strictly prohibited (Mazower 2008).

[4] A few epidemiological studies examined the long-term consequences of the Siege of Leningrad on mortality and health outcomes using the sample of Leningrad residents (Koupil *et al*., 2009; Sparen *et al*., 2003; Stanner

Diseases – It is known that undernourishment can make a person more susceptible to infection and other diseases. In addition to poor nutrition, other war-related factors such as unclean water, poor sanitation, inadequate health care, and the lack of shelter provide conditions for the spread of infectious diseases during the wartime. The available scarce data on epidemiological situation in occupied territories of the Soviet Union indicate a dramatic rise in the incidence of diseases during the occupation period. For example, in 1944 compared to the pre-war year of 1940, the registered cases of illness in Belarus increased by 45 times for typhus, 5 times for sexually transmitted diseases, 4 times for malaria, 2.6 times for typhoid fever, and 2 times for tuberculosis (Logvinenko, 2010). Based on robust results in studies of the 1918 influenza pandemic (Almond 2006, Lin and Liu, 2014), it is plausible to expect that the increased exposure to infectious diseases in early childhood during WWII may also have long-lasting health consequences.

Stress **–** The war experience is undoubtedly traumatic on many levels; it may trigger stress reactions in children and be responsible for a variety of health problems later in life. The most likely stress triggers during the war include the death of parents, separation from parents due to their involvement in military service, witnessing public hangings and other forms of violence, experiencing physical abuse by German soldiers, starvation, watching parental stress reactions, etc. In our Ukrainian survey, almost 26 percent of the respondents born in 1936-1945 report not having a father alive at age 14; 4 percent did not have a mother when they were 14 years old. In later cohorts (born in 1949-1958), these numbers change dramatically – only 5 percent had a deceased father and 1 percent a deceased mother at age 14. A large number of experimental and epidemiological studies demonstrate that health in the later stages of life may be strongly influenced by stressful experiences in prenatal and early childhood stages of life (see review of studies in Vaiserman, 2015).

Reduced parental investment in child development – The loss of a parent or absence of a parent on war service not only triggers stress in children, but also prevents parental involvement during critical periods of early child development and thus diminish child's future capabilities. Even for children whose parents were present during the war, parental investment into child development is likely to be reduced due to deteriorated parents' health, insufficient parental income, starvation, and other reasons. Previous literature finds that children's health is negatively associated with parental death (Beegle *et al*., 2010; Corak, 2001; Krause, 1998), parental job loss (Lindo 2011), and adverse income shocks (Case *et al*., 2002)

Lack of maternal and health care – The destruction of hospitals and clinics along with barred medical supplies is another channel that may contribute to poor health outcomes later in life. During the German occupation of the Soviet Union, 40,000 health care facilities were destroyed or burned (Goskomstat, 2015). As of January 1, 1945, former occupied territories of the Soviet Union were left with 48 percent of pre-war birth care center beds and 52 percent of children hospital beds. The medical staff was severely downsized. By the end of 1944, the two republics that were fully occupied, Belarus and Ukraine combined, had only 46 percent of doctors from the pre-war level, 42 percent of gynecologists, 39 percent of pediatricians, 54 percent of nurses, and 46 percent of midwifes (Health Care in the Union of SSR, 1946). Even though a methodologically solid paper on the long-term health consequences of the destruction of the health care system is yet to be written, the causal link here is apparent.

Destruction of the material wealth – Many material assets that are essential for children health – such as water supplies, sewer, heating systems needed for survival in the winter, and housing – were severely damaged during the German occupation. Over 50 percent of houses and apartment buildings in occupied cities were destroyed (Voznesensky, 1948). In some cities like Pskov, Vitebsk,

et al., 1997; Vagero *et al*., 2013). The research design was not ideal by today's standard; for example, the control group inadvertently included individuals that could have suffered from malnutrition in other occupied areas of the country and also individuals exposed to the post-war 1946-47 famine. The findings from these studies are often inconclusive and sometimes conflicting.

Voronezh and others, the share of livable housing slumped to less than 10 percent of the pre-war housing stock. People had to live in sheds, bunkers and dangerous unheated housing conditions. The destruction of other forms of material wealth - electricity, communication, transport, manufacturing, agriculture, etc. – was also devastating (see Voznesensky (1948) for aggregate statistics on destruction). Production of material goods may have an indirect effect on children health, as it provides parents with jobs and income, produces and supplies food and medicine, and provides people with other basic essentials.

The above list of channels through which the Nazi occupation may have impacted children of the war is not complete, but it shows that the time of occupation was devastating for people and that the channels are numerous and intertwined with each other. For the disaster of this magnitude such as WWII, disentangling the impact of a specific channel is methodologically unfeasible due to channels' complementarity and reciprocity. What is feasible is to recover the treatment heterogeneity with respect to different margins, including child's age at which the impact occurred, the length of individual exposure to war hardships, and the extent of regional destruction and human losses. Therefore, in what follows, our treatment variable represents a combination of all channels, but it is interacted with various treatment margins.

3. Theoretical Framework

3.1. General Set-Up

In this section, we describe the theoretical framework that guides our research process. We model health outcomes using the health production function, in which human capabilities determine health outcomes at each stage of life. Following Conti and Heckman (2014), Cunha and Heckman (2007) and Heckman (2007), we assume that the health outcome is a function of a set of health and other human capabilities created in a multi-stage technology.[5] The key characteristic of the model is that each stage of life may have different technology and/or the set of human capabilities. Furthermore, the early stages of child development (e.g., in utero, infancy) affect later stages of life. In this model, the key mechanisms through which the Nazi occupation of the Soviet Union affects adult health outcomes are dynamic complementarity and self-productivity. *Dynamic complementarity* means that health and other capabilities produced at one stage increase the productivity of investment in health and other capabilities at subsequent stages. *Self-productivity* implies that health and other capabilities produced at one stage augment the capabilities generated at later stages, which in turn improves health and survival.

Our model is a simplified version of the overlapping generations model (Cunha and Heckman 2007). We model six stages of life: in-utero ($t = 0$), infancy ($t = 1$), toddler ($t = 2$), preschool ($t = 3$), young adulthood ($t = 4$), and adulthood ($t = 5$). The first four stages from conception to age 5 are the main focus of our research. As Almond and Currie (2011a, 2011b) argue, the adverse shocks experienced in utero and during early childhood may be more influential than shocks in later childhood. These four stages are often dubbed as the critical period of child development. According to Heckman (2007), a critical period is a stage in the lifespan when late investment is not able to compensate for the lack of early investment.

The kth health outcome at time t, H_t^k, is function of a vector of health and other individual capabilities, θ_t:

$$H_t^k = g^k(\theta_t) \tag{1}$$

For example, the kth health outcome in adulthood is given by $H_6^k = g^k(\theta_6)$.

The production technology of capabilities at stage $t + 1$ is

---

[5] An alternative model for two periods (childhood and adulthood) is presented in Almond and Currie (2011b). A multi-stage technology fits better with our empirical finding in that the different stages of childhood at which the shock occurred have differential effects on future capabilities.

$$\theta_{t+1} = f_t(\theta_t, I_t, p, z) \quad (2)$$

where $f_t(\cdot)$ is increasing and twice continuously differentiable in all arguments; I_t denotes investment in health and other capabilities made by parents in early stages, an individual himself in later stages, community, and government in period t, p denotes exogenous parental capabilities, such as genes or parent education, and z denotes constant community characteristics affecting capabilities such as climate, regional food practices, water sources, urban density, etc. θ_0 define the initial capabilities each individual is born with and which are formed in utero.

To make the technology function for capabilities more flexible, we can add a group-specific productivity shifter A_j for group j:

$$\theta_{t+1} = A_j f_t(\theta_t, I_t, p, z) \quad (2b)$$

For example, a productivity shifter can account for gender differences in the generation of human capabilities. In fact, our empirical exercise finds significant differences in the effect of early-life shocks on future adult capabilities between males and females.

If current health and other capabilities create higher capabilities in the next period, then technology $f_t(\cdot)$ is self-productive ($\frac{\partial f_t}{\partial \theta_t} > 0$). Similarly, if health and other capabilities accumulated until the previous stage complement investment at the current stage, then technology $f_t(\cdot)$ is complementary ($\frac{\partial^2 f_t}{\partial \theta_t \partial I_t'} > 0$). This complementarity is dynamic if early investment makes later investment more productive.

Substituting $\theta_0, \dots, \theta_5$ repeatedly for θ_6 in equation (2), human capabilities in adulthood can be written as a function of parental capabilities, health and other capabilities formed in-utero, θ_0, and all subsequent investments:

$$\theta_6 = m_6(\theta_0, I_1, I_2, I_3, I_4, I_5, p, z) \quad (3)$$

Using equation (3), the k^{th} health outcome H_6^k in adulthood is given by

$$H_6^k = n^k(\theta_0, I_1, I_2, I_3, I_4, I_5, p, z) \quad (4)$$

Through the channels discussed in Section 2, the formation of human capabilities during the initial stages of life ($\theta_t, t = 0, 1, 2, 3$) and investment at these stages are likely to be adversely affected by war-related exogenous shocks, ξ_g (i.e., $d\theta_t/d\xi_g < 0$ and $dI_t/d\xi_g < 0$). The subscript g indicates that the shock could vary at some geographic level, which may provide an exogenous variation in initial capabilities and early childhood investment. In particular, the intensity of shock, ξ_g, may depend upon the duration of occupation or the level of destruction that each city/region suffered. The technology function n^k does not need to be linear and may allow for non-linear effects of the intensity of shock, ξ_g, within each stage of life.

By linking the discussion in Section 2 with the above theoretical framework, we can highlight at least three long-term causal mechanisms through which the Nazi occupation of the Soviet Union may have affected health in the last stage of human life or adulthood in the model (H_T=H_6^k).

$$\xi_g \Rightarrow \begin{array}{l} (1) \downarrow \theta_t \Rightarrow \downarrow \theta_{t+1} \Rightarrow \cdots \Rightarrow \downarrow \theta_T \Rightarrow \downarrow H_T \\ (2) \downarrow \theta_t \Rightarrow \downarrow I_t \Rightarrow \downarrow \theta_{t+1} \ldots \Rightarrow \downarrow \theta_T \Rightarrow \downarrow H_T \\ (3) \downarrow I_t \Rightarrow \downarrow \theta_{t+1} \Rightarrow \cdots \Rightarrow \downarrow \theta_T \Rightarrow \downarrow H_T \end{array}$$

First, the occupation may lower health capabilities in-utero and in early childhood, for example, via malnutrition, war time diseases, physical abuse, injuries, and high stress levels. Since technology is self-productive, the deteriorated capabilities in early stages affect the formation of capabilities in later stages, including the last stage, θ_6, and through this channel reduces health outcomes in adulthood. The literature in epidemiology and medicine has shown that in-utero and infancy are critical periods for the physical, cognitive, sensory and motor development, and adverse environments during this time have lasting effects on health outcomes (Barker 1990, 1998, Karp et al 1995, Gluckman and Hanson 2005, Schultz 2010).[6]

Second, given the complementarity between the previous accumulation of human capabilities and current investment in health, a significant war-caused decline in health capabilities in early stages of child development reduces the productivity of investments in later stages (i.e., I_1, I_2, I_3, I_4, I_5), and through this channel negatively affects health outcomes during adulthood.

Third, the occupation directly reduces parental investment during childhood as a result of the loss of a parent, the absence of a parent on war service, deteriorated parents' health, the lack of parental jobs and income, and destroyed family housing. The community and government investments in health are also harmfully affected, leading to the lack of basic utilities and infrastructure, poor sanitation, and the lack of maternal and health care. These reduced investments during early childhood are likely to lower health outcomes in adulthood through the channels of self-productivity and dynamic complementarity.

3.2. Empirical Model

The main objective of the empirical model is to estimate the effect of shocks triggered by the Nazi occupation of the Soviet Union on health outcomes in adulthood. The empirical model is an approximation of equation (4). It is linear in parameters, but not necessarily linear in variables.

Our unique quasi-experimental setup exploits the individual date of birth (d), the municipality of birth (m), and exogenous variation of the location and timing of German occupation during WWII. The identification of the treatment effect relies on this quasi-experimental variation and on a full set of region of birth (r) and birth year (t) fixed effects, where r is a next-level aggregation of m, and t is an annual aggregation of d.[7]

The starting specification is the following:

$$Y_{irts} = \alpha + \boldsymbol{\gamma} O_{irt} + \rho_r + \delta_t + \rho_r \times \varphi(t) + \mu_s + \beta' X_{irts} + \epsilon_{irts} \quad (5)$$

where Y_{irts} is a late-life health outcome observed in survey year s for an individual i who is born in year t and birth region r, ρ_r is the birth region fixed effect, δ_t is the birth year fixed effect, $\rho_r \times \varphi(t)$ is the interaction of the birth region fixed effect with some functional form for t, μ_s is the year-of-survey fixed effect, X_{irts} is a vector of observable individual characteristics, and ϵ_{irts} is a random error term.

Our treatment variable O_{irt} is a dummy variable that takes the value of 1 for individuals who were born in an occupied municipality of the Soviet Union and who were in a specific stage of childhood under occupation. The control group includes two types of individuals: (1) individuals who were born in non-occupied municipalities and who were in a specific stage of life during the war period; and (2) individuals born after the war in both formerly occupied and non-occupied municipalities. Further specifics on definitions will be discussed in Section 4. Both parts of the control group are necessary, as they provide spatial and intertemporal dimensions for the difference-in-difference estimation. Without individuals born after the war, it would not be feasible to isolate the effects of Nazi occupation from other specific characteristics of regions such as factor z included in the production technology of capabilities at each stage given by equations (2)-(4). Including individuals born before the war into the control group will be inconsistent with our theoretical framework. Since investment and capabilities of pre-war cohorts are likely to be harmed by the war at later stages of child development ($I_2, I_3, \theta_2, \theta_3$), these war shocks, via self-productivity and dynamic complementarity, may generate an adverse effect on late-life health outcomes of pre-war cohorts.

---

[6] For instance, environmental conditions in utero modify the epigenome and through this channel increase the risk of chronic health conditions in adulthood (Thornburg *et al.* 2010, Petronis 2010).

[7] The combination of *m* and *d* in our data identifies a unique individual *i* for 99.6 percent of surveyed individuals. The coincidence of several respondents being born on the same day in the same city is negligible. Thus, using the subscript *i* is equivalent to using the subscript (*md*) in our case.

Here we are interested in estimating $\boldsymbol{\gamma}$. Region-of-birth fixed effects account for the systematic differences in health outcomes across regions, while year-of-birth dummies control for the permanent differences in health outcomes across birth cohorts. The interaction term $\rho_r \times \varphi(t)$ is included to account for time-varying region-specific policies (e.g., different speeds of post-war reconstruction and development across regions). Year fixed effects capture country-wide contemporary shocks in health outcomes. To avoid including endogenous individual-level covariates, we limit the vector X_{irts} to gender, ethnicity, type of the birth settlement, and parents' education. The latter variable serves as a proxy for parental capabilities p in equations (2)-(4).

Our main identifying assumption that is needed to consistently estimate γ is that once we control for the vector of observable individual characteristics X_{irts}, a set of region-of-birth and year-of-birth fixed effects (ρ_r, δ_t), the interaction term $\rho_r \times \varphi(t)$, and annual macro shocks μ_s, the error term ϵ_{irts} is uncorrelated with the Nazi occupation. In the end, our identification strategy relies on within-region cross-cohort variation to identify the effects of Nazi occupation.

Specification (5) is estimated for each of the four stages of life considered in our theoretical framework: in-utero, infants, toddlers, and preschoolers. We further test for gender differences in treatment effects by including a two-way interaction between the treatment variable and two gender categories, $F_i \times O_{irt}$:

$$Y_{irts} = \alpha + \boldsymbol{\gamma_F}(F_i \times O_{irt}) + \rho_r + \delta_t + \rho_r \times \varphi(t) + \mu_s + \beta' X_{irts} + \epsilon_{irts}. \quad (6)$$

Following our theoretical framework, we also estimate the model using three different proxies for the intensity of shocks, ξ_{md}: the duration of occupation, the length of exposure to shocks, and the level of destruction. We call these proxies as treatment margins.

$$Y_{irts} = \alpha + \boldsymbol{\gamma_\xi}(\xi_{md} \times O_{irt}) + \rho_r + \delta_t + \rho_r \times \varphi(t) + \mu_s + \beta' X_{irts} + \epsilon_{irts}, \quad (7)$$

where ξ_{md} is the intensity of shock experienced by a child born on day d in municipality m.

The intensity of shock may depend upon the duration of occupation in each municipality. This margin varies at the municipality level m but not at the d level. For instance, we can test if the late-life health outcomes for infants born during occupation in cities like Stavropol which was occupied for 169 days are the same than for infants born under occupation in cities like Pskov where occupation lasted for more than three years (1110 days).

The intensity of shock may also vary with the date of conception and birth, which, together with the dates of occupation, determine the length of individual exposure to adverse shocks.[8] This margin varies at both levels m and d. In the example of the city of Pskov, γ_ξ's might be different for individuals conceived at the beginning of occupation with a longer exposure to shocks compared to individuals conceived at the end of the occupation with a shorter period of exposure in the same city of birth.

Finally, a third margin captures non-linear effects associated with the spatial differences in the level of destruction and population losses that each city/region suffered. This margin varies at the level of m only.

The fact that we observe individuals affected by the war in later stages of their life (ages 50-77) has two important methodological implications. First, $\boldsymbol{\gamma}$ captures the treatment effect on survivors. It is likely to be biased if the probability of survival (and thus the selection of an individual into our estimation sample) is positively correlated with ϵ_{irts} in health outcomes. We can partly address this concern for those survived till the first year of the survey by modeling subsequent survival between ages 50 and 77 and re-estimating equation (5) with the inverse survival probability weights. Second, health effects of the past compensatory investment made by an individual, parents or communities to offset negative early-life shocks are going to be absorbed in the treatment effect. Even though the remediation of adverse shocks is generally found to be less effective after early

[8] The duration of the conflict is the most common margin used in the literature associated with the long term effects of wars or violent conflicts. This margin does not necessarily reflect the real exposure of an individual to the conflict.

childhood, some evidence of compensatory investment exists especially in non-cognitive skills such as perseverance, motivation, self-esteem, self-control, etc. (Conti and Heckman, 2014). Assuming some persistence in non-cognitive skills over the lifespan and using the same functional form (5), we can test if individuals in the treatment group have developed a different set of non-cognitive capabilities.

4. Data

In this section, we describe survey data, define "treated" municipalities, and discuss the composition of treatment and control groups. We also introduce various margins of treatment intensity, including municipality-specific duration of occupation, individual-specific length of exposure to occupation, and region-specific measures of destruction and population loss. Next, we provide the *t*-test results for the pre-treatment conditions at the regional level. We conclude the data section by presenting summary statistics on long-term health outcomes and model covariates.

4.1. Survey data

Our study draws on two longitudinal surveys of Russia and Ukraine – RLMS-HSE and ULMS, respectively. The RLMS-HSE is one of the longest nationally representative longitudinal surveys of households.[9] The main sample of households is surveyed annually from 1994 and onwards, with the exception of years 1997 and 1999. The number of respondents per year fluctuated between 10,500 and 14,700 individuals until 2009 and rose to more than 20,000 respondents in 2010-2013. We use all survey waves beginning year 2000 when the RLMS-HSE added many health-related questions.

The ULMS is a household and labor force survey based on a statistically representative sample of the Ukrainian working-age population between the ages of 15 and 72.[10] The Ukrainian survey started in 2003 and repeated in 2004, 2007, and 2012. The data from the 2012 survey is not yet publicly available and will not be used in our study. During the first three waves of the ULMS, 9,902 individuals from 4,232 households participated in the survey.

The most recent details on survey instruments, sample design, and attrition in both Russian and Ukrainian surveys can be found in Gerry and Papadopoulos (2015), Kozyreva and Sabirianova Peter (2015), and Lehmann *et al*. (2012). The attrition due to aging (i.e., reaching the upper age limit of 72) in ULMS does not represent a problem for us, as it concerns people born before 1935. The attrition due to mortality is an important issue in both surveys and will be addressed using a standard tool of inverse propensity weighting.

The research design described in the previous section requires having some non-occupied territories in the control group and thus can be applied either to Russia alone as the only Soviet republic that was partially occupied or to the combined Russia-Ukraine sample. The ULMS emulated many questions from the RLMS-HSE, hence making it easier to append the two surveys. We account for different sample sizes and different number of survey waves in two countries by re-weighting the country composition of the sample to match the country composition of the combined general population, of which 75 percent resided in Russia and 25 percent in Ukraine in 2000-2013.

---

[9] The RLMS-HSE is organized by the National Research University Higher School of Economics, Moscow together with the Carolina Population Center at the University of North Carolina at Chapel Hill and the Institute of Sociology at the Russian Academy of Sciences. The RLMS-HSE surveyed individuals in 32 out of 83 regions in all seven federal districts of the Russian Federation (according to the official classification of regions as of January 1, 2010).

[10] The ULMS is organized by the Institute for the Study of Labor (IZA) in collaboration with many other institutions, including the European Union, the World Bank, the William Davidson Institute of the University of Michigan, and others. The data collection was entrusted to the Kiev International Institute of Sociology. The survey was done in all 27 regions of Ukraine.

We start with the cohort of respondents who are born between 1935 and 1958 and whose municipality of birth (city or rural county) is known with certainty. In ULMS, the questions about the country, county, and settlement of birth were asked in an open-ended form in 2003 and 2007. This information is used to infer the birthplace for the same respondents in the 2004 survey. Thus, we have a nearly full coverage in Ukraine. The response rate is very high – 94 percent of observations from the 1935-1958 birth cohorts have complete information on the municipality of birth.[11] Unfortunately, in RLMS-HSE, an open-ended question on birthplace was only asked once, in 2013. In all other years, the municipality of birth can be inferred either from the 2013 survey if the respondent participated in earlier surveys or from the current residence if the respondent is born in the same place as her present municipality. Out of 68,605 observations from the 1935-1958 birth cohorts and from all RLMS-HSE survey years, 70 percent have the identifiable name of birth city or birth county; 5 additional percent can be unambiguously classified into either occupied or non-occupied areas based on the information provided on the region or republic of birth; the remaining 25 percent with unknown birthplace had to be dropped from the analysis.

4.2. Treatment at the Municipality Level and the Duration of Occupation

In our study, the word "treatment" has a negative connotation of being subjected to German occupation in childhood. The municipality where a person is born is classified as being "treated" if it was (i) fully occupied by German forces and their allies, (ii) partially occupied such as the city of Stalingrad (now Volgograd), or (iii) sieged such as the city of Leningrad (now St. Petersburg). Although Leningrad had not been occupied, its 900-day siege claimed 650,000 Leningrader lives in 1942 alone, mostly from starvation, disease, and shelling from German artillery (Encyclopedia Britannica). Based on these and many other horrific facts on the deadliest siege in history, Leningrad clearly belongs to the "treated category". Some ambiguity may arise with respect to other non-occupied cities, which were severely bombarded but not encircled, such as the city of Moscow. We decided not to include the bombarded non-occupied cities into the treatment group since their food supplies and evacuation routes remained intact.

We assign the occupation status to geographic units in the 2011 GADM spatial database of Global Administrative Areas of the former Soviet Union countries. The geographic units are chosen at the level of municipality that could be either a city or rural county.[12] Altogether we have 3,857 municipalities. This number excludes uninhabited geographic areas, within-city districts, and municipalities for which the occupation status is ambiguous such as Kaliningrad region, Klaipeda in Lithuania, Southern Sakhalin, and Transcarpathia, which were acquired by the Soviet Union in 1945 as a result of WWII.

Next, for each occupied municipality, we collect information on the beginning and end of occupation. We add rural municipalities and smaller cities to the previously published database on the dates of occupation and liberation of the USSR cities during the WWII (Dudarenko *et al*., 1985). Figure 1 plots the length of occupation in days on the contemporary map of the former Soviet Union countries. This map is summarized in Table 1. Almost 37 percent of 3,858 municipalities were occupied by the German army. The occupation lasted for less than one month in 43 municipalities, including some counties close to the never surrendered cities of Moscow and Tula; 1 to 6 months – in 200 municipalities, including Stalingrad and nearby areas and also locations in the North Caucasus; more than 6 months but less than a year – in 113 municipalities, such as the city of Rostov-on-Don

[11] Also, 1 percent are born outside the USSR and excluded from the estimation sample; 2.5 percent have missing information on the municipality of birth, but the answers provided on the country and region of birth are still sufficient enough to determine if the place of birth was occupied or not during the WWII, and only 2.5 percent of observations are discarded because of refusals or unclear answers.

[12] Further details on the GADM database and geographic coding used in this paper are provided in web Appendix 1.

and some areas on the Eastern coast of the Black Sea and the Sea of Azov. About two thirds of all occupied municipalities suffered under the occupation for more than a year – the occupation lasted for 12 to 24 months in 257 municipalities of the Central Russian Upland and Eastern Ukraine, including cities of Belgorod, Donetsk, Kharkov, Kursk, and Voronezh; 24 to 36 months – in 565 municipalities, including well-known cities of Kiev, Leningrad, and Sevastopol; and 3 years or more – in 241 municipalities of Western Belarus, Western Ukraine, the Baltics, Moldova, and Northwest Russia.

We use the same GADM classification of Global Administrative Areas to geocode the respondents' answers to open-ended questions on birthplace; the geocoding process is described in web Appendix 1. The geocoded individual observations in RLMS-HSE and ULMS are then linked to external databases, including the one on the dates of occupation and liberation. Table 1 provides the distribution of municipalities in the combined RLMS-ULMS sample from 2000-2013 surveys. The number of municipalities where respondents are born is fairly large (more than 2,000), and most of municipalities are located in Russia and Ukraine. The share of occupied municipalities in our sample is 45 percent, which is larger than that for the entire country of the Soviet Union, but the distribution of municipalities by the duration of occupation is similar.

4.3. Treatment at the Individual Level and the Length of Individual Exposure

The treatment group is defined at the individual level. It is constructed based on the birth date (day, month, and year) and the dates of occupation of the municipality of birth. An individual belongs to the "treatment group" if she is conceived, born or lived until age 5 in occupied municipalities during the period of Nazi occupation. The conception date is approximate – 266 days or 38 weeks subtracted from the birthdate.[13] The overlap between the age interval from conception to age 5 and the occupation interval determines the length of individual exposure to hardships in childhood under Nazi occupation, and it is one of the main measures of treatment intensity.

The details on children migration are not available in our data, so we have to assume that children stay in their municipality of birth until age 5. This assumption is not unreasonable. A number of WWII history books provide evidence that the German occupation of the USSR was unanticipated by general population, the retreat of the Red Army was rapid, the evacuation of civilians was poorly organized, and the escape routes were cut off once German forces moved into cities and villages (Ehrenburg and Grossman, 2003; Erickson, 1999; Manley, 2009).

To examine the role of different stages of child development in future health outcomes, we distinguish the following four age intervals within the treatment group: *in-utero* (from conception to birth), *infants* (from birth to 12 months of age), *toddlers* (1-3 years old), and *preschoolers* (3-5 years old). The treatment variable for each stage of life is equal to one if there is any overlay between the corresponding age interval and the occupation interval for the municipality of birth. We note that in the municipalities that were occupied for several years, the same individual may appear in multiple treatment groups by going through several stages of life under occupation.

Following the theoretical framework presented in Section 3, the controlled observations consist of two sets of individuals: (i) individuals who are conceived, born, or lived until age 5 during the war in non-occupied municipalities (or the war control group) and (ii) individuals who are born in all municipalities after the war (or the post-war control group). The war period for the first group is fixed between the beginning and end of German occupation in our survey sample, June 22, 1941 (occupation of the city of Brest) and October 15, 1944 (liberation of the city of Riga). Thus, children in the war control group are born in non-occupied municipalities between June 22, 1936 (for kids

---

[13] For example, the city of Leningrad was under siege from September 8, 1941 until January 27, 1944. Thus, the "treatment group" includes individuals who are born in this city between September 8, 1936 (for kids reaching age 5 on the first day of the blockade) and October 18, 1944 (for children conceived on the last day of the city blockade).

reaching age 5 on the first day of occupation) and July 8, 1945 (for kids conceived on the last day of occupation in our sample). Similar to the treatment group, we also distinguish the four stages of child development within the war control group: in utero, infants, toddlers, and preschoolers, as illustrated in Figure 2.

In defining the post-war control group, we follow the related literature by skipping birth cohorts born immediately after the war (Akbulut-Yuksel, 2014; Chen and Zhou, 2007). This is done to avoid the contamination of the control group with individuals affected in the aftermath of the war. For example, the major famine hit the USSR in 1946-1947 due to the drought and post-war devastations. Our post-war control group includes individuals born in the USSR in 1950-1953, i.e., conceived on as early as April 1949. The four-year period is chosen to approximately match the length of the war. Such approach to defining the post-war control group by choosing fixed birth cohorts is prevalent in the literature and will be adopted in our study. However, with this approach, the time intervals between the beginning of the treatment and the first year of the post-war control group are different across stages of life: up to 14 years for preschoolers and up to 9 years for babies in utero, as can be seen in Figure 2. To make the time spells comparable, we implement an alternative approach by shifting the entire time frame for 1936-1945 three years later for 1949-1958 (Figure 2).

4.4. Treatment Margins at the Regional Level: The Destruction of the Industrial Base and Population Change

In attempt to capture additional heterogeneity in the treatment effect in terms of the severity of shocks, we use two proxies for the extent of destruction at a more aggregate level. First, recently declassified documents contain regional data on electricity production in 1940 and 1944 (Central Statistical Directorate of the Soviet Union, 1946). The absence of a national grid system made it impossible to replace power losses suffered in one region of the Soviet Union by power imports from other areas (Hoeffding, 1970). Thus, the log difference in electricity production between 1941 and 1944 can serve as a reasonable proxy for the destruction of the industrial base and infrastructure at the regional level. We plot this indicator in Figure 3 for the European part of Soviet Union. The electricity production fell by 18.8 percent overall and by 72.6 percent in occupied regions. Some occupied regions lost practically all power generation capacity – in Belarus, Moldova, parts of the Ukrainian SSR (Nikolayev, Odessa, and Zaporozhye regions), and in parts of the Russian SFSR in 1945 borders (Crimea, Pskov, and Smolensk regions), electricity production fell by more than 95 percent from its pre-war level. To support the expanded military production, new power plants were built from scratch in non-occupied regions. For instance, regions of Western Siberia increased their electricity production by 107.6 percent between 1941 and 1944, and regions of the Urals – by 94.2 percent. The decline in electricity generation interacted with a dummy for treated municipalities is strongly correlated with the duration of occupation (–0.86)

The second proxy is the log difference in regional population (adjusted for pre-war mortality rates) between the two censuses 1939 and 1959. Specifically, we calculate a counterfactual population between the ages 20 to 70 in 1959, which is the region's population aged 20-70 that would have been achieved in 1959 if the mortality rates stayed at the pre-war level and the net migration was zero.[14] Then, for each region, we take the log difference between actual population aged 20-70 in 1959 and counterfactual population for the same age group. This difference shows the regional population change due to excess war mortality and also due to interregional migration during the war and post-war periods.[15] By no means, this measure represents the count of human losses, but it

[14] We take regional age-specific population aged 0-50 in the 1939 Census, adjust it for 20-year mortality with age-specific pre-war mortality rates, and recalculate it in 1959 borders of regions.

[15] Part of migration flows could be related to the war, including rushed and disorganized evacuations in the chaos of the Soviet retreat and the flight of people from destroyed unlivable areas right after the liberation. The

could represent another dimension of the level of hardship in occupied regions. The correlation between the 20-year population change interacted with a dummy for treated municipalities and the duration of occupation is –0.67.

The summary statistics on population change in excess of natural mortality over the 20-year period is reported in Table 2, Panel A. It indicates larger population losses in occupied regions (–0.233 log points) compared to non-occupied regions (–0.127). By further disaggregating the adjusted measure by age in 1939 and gender, we uncover shattering losses of male population in the 15-50 age groups in both occupied (-0.529) and non-occupied areas (-0.449). Young males in occupied regions seem also affected in a big way (–0.218). Relatively smaller changes in population are observed in non-occupied regions among young males (–0.048), adult females (–0.045), and young females (0.002). In Panel B of Table 2, we also report the raw differences in population between the two censuses without accounting for pre-war mortality rates. It is noteworthy that the largest decline in population occurred in a so called "deficit zone" outlined in the Hitler Hunger Plan in early 1941; see Section 2.[16]

4.5. Pre-Treatment Conditions

Before proceeding to the estimation, we compare the pre-war levels of economic development between the occupied and non-occupied territories. Table 3 shows that occupied and non-occupied territories prior to the WWII had similar levels of gross output per capita and per worker, energy power capacity, electricity production, employment participation, and wages of staff and workers. The two parts of the country exhibited similar investment levels in health care in 1935-1936, as the reader can see from the population-adjusted statistics on physicians, hospital beds, birth centers, inpatient clinics for maternal and infant health, and pediatric treatment centers. Two years before the war, the composition of population was also not that different between the occupied and non-occupied regions in terms of the share of females, the degree of urbanization, the attainment of secondary and higher education, female schooling achievements, and ethnic diversity.[17]

Only a few pre-treatment indicators exhibit statistically significant differences between the two groups, all favoring Western, soon-to-be-occupied territories. In 1939, population in these territories had a lower infant mortality rate (145 vs. 192 infants deaths under one year old per 1,000 live births), a lower crude birth rate (32 vs. 36 infants under one year old per 1,000 people), a smaller number of children per a woman of reproductive age (439 vs. 506 children under age 5 per 1,000 women aged 15-49), and a better adult illiteracy rate (19 vs. 24 percent of illiterate population age 15 and above).[18] If the population in occupied regions indeed had better health conditions prior to

---

escape during the occupation was practically impossible. Another part of migration flows is associated with post-war policies of labor reallocation to the East and mass deportations of entire ethnic groups.

[16] Regions that lost the most population aged 20-70 in excess of the pre-war levels of mortality are all located in a "deficit zone" and include Smolensk region (–50% drop in adjusted population), Vitebsk region (–42%), Novgorod region (–41%), Pskov region (–41%), Leningrad region (–40%), and other nearby regions.

[17] The Soviet Union was an ethnically diverse country. In 1939, 14 ethnic groups had population of almost one million or more, with ethnic Russians being the largest group (58.4 percent of total population). Even though both occupied and non-occupied territories had a similar average value of ethnic fractionalization, they had different ethnic composition. Before the war started, the four largest ethnic groups in the occupied territories were Russians (49 percent), Ukrainians (31 percent), Belarusians (12 percent), and Jewish (3.5 percent). 80 percent of Jewish population resided in the occupied territories, which also included the so called Pale of Settlement or the area of the Imperial Russia beyond which Jewish permanent residency was disallowed (Grosfeld et al. 2013). The residency ban was lifted in 1917, but the majority of Jewish population continued living closer to the Western USSR borders.

[18] A closer investigation reveals that statistically significant differences between the occupied and non-occupied territories in birth and fertility rates are largely driven by pre-dominantly Muslim regions with traditionally higher birth and fertility rates.

the WWII (as the infant mortality statistics suggests), then this would bias our results towards finding a smaller effect of the WWII on the long-term health outcomes, thus providing a lower bound estimate. To control for potential regional differences in pre-treatment conditions, we include regional fixed effects and the type of birthplace (city, township or village) in all specifications. In some specifications, we also focus on nearby municipalities in partially-occupied regions on both sides of the occupation boundary.

4.6. Outcomes

Both RLMS and ULMS offer a large array of potential health outcomes during the late-life years, including: 1) an indicator for having bad or very bad health; 2) having a chronic disease of heart, respiratory organs, liver, kidney, stomach, intestines, and/or spine); 3) had a heart attack or having a chronic heart disease; 4) having a chronic respiratory disease; 5) having a chronic disease of digestive and genitourinary systems; 6) having a spinal disorder; 7) suffered from hypertension; 8) had a surgical operation in the last 12 months (RLMS only); 9) had a depression in the last 12 months (RLMS only, 2003, 2004, 2011-2013); 10) individual height; 11) underweight, BMI<18.5; and 12) obese, BMI>30. Unless noted otherwise, all outcomes are available for the entire sample period, 2000-2013. Individual height is used to construct two proxies for child malnutrition, including the deviation of height from the mean for each gender aged 20 and above to account for biological differences in height (*residual height*) and having a height below the 25th percentile of gender distribution (*low height*). In line with Bertoni (2015) who shows the impact of famine in early life on life satisfaction in adulthood, we use two indicators as the outcome measures of subjective well-being: being satisfied or fully satisfied with life (*life satisfaction*) and being dissatisfied or fully dissatisfied with life (*life dissatisfaction*). The detailed definitions for all variables can be viewed in appendix Table A1.

Table 4 reports the descriptive statistics of the above outcomes for treatment and control samples as well as the mean comparison *t*-test. Compared to individuals in the control group, the treatment group has a significantly higher incidence of reporting bad health (39 percent vs. 21 percent), chronic diseases (67 percent vs. 56 percent), heart diseases (41 percent vs. 25 percent), and digestive system disorders (38 percent vs. 33 percent). Individuals in the treatment group are also shorter and more likely to be dissatisfied with their lives (51 percent vs. 43 percent). Other health outcomes do not exhibit statistically significant group differences in means.

RLMS provides additional opportunities to look at non-cognitive self-assessments on a 4-point scale in 2002-2005 and 2011 surveys, including 1) respondent determines own future; 2) has good characteristics; 3) feels like a loser; 4) has a positive attitude; 5) is satisfied with self; 6) feels useless; and 7) feels like a bad person, among others. One self-assessment question on a 9-pont scale for feeling well respected is available in all survey waves. The summary statistics for these variables (also reported in Table 4) does not paint a consistent picture: the treatment group appears to be dissatisfied with life, but satisfied with self; less likely to determine own future but more likely to have a positive attitude; to be critical of having good characteristics, but less likely to feel like a bad person. We will leave the interpretation of non-cognitive self-assessments till the results of model estimation.

5. Results

Re-introduce the model

Discuss Xs

To control for the differences in health outcomes between genders and ethnic groups, we include a female dummy and dummies for individual ethnicity (Russian-Ukrainian mixed, Russian, Ukrainian, Jewish, deported nationalities). Mother's and father's level of education (college, secondary education or missing) help us control for parental capabilities. Type of the birthplace (city, township, or village) is also likely to react differently to health shocks associated with German occupation.

Women comprise 64.8 percent of treatment and 61.3 percent of control sample. Treatment sample consists of 42.2 percent Russian, 45.2 percent Ukrainian, and 7.3 percent deported nationalities. In the control sample, 78.4 percent are Russian and 9.3 percent are Ukrainian. Individuals in the treatment group are more likely to be born in a city (64.3 percent) than in a township (5.8 percent), or a village (29.9 percent) while only 46.3 percent of the control group is born in a city (8.4 percent are born in a township and 45.3 percent – in a village). Finally, individuals in treatment group are less likely to have parents with college or secondary education (0.8 percent of mothers and 1 percent of fathers have college education) perhaps due to the fact that individuals born after WWII (hence part of the control group) are more likely to have educated parents as college education was becoming more available in late 1950s in the USSR.

Table 1 reports the estimates of our model for adults who were between in-utero and 5 years old at the time of the occupation as well as for 4 stages of childhood separately (in-utero, infant, toddler, and preschool). Living in an occupied municipality during the period of World War II is significantly associated with reporting one's subjective health in adulthood as bad (18.3 percent more likely), having one or more chronic diseases (7.5 percent more likely), having a respiratory disease (7.9 percent more likely), a digestive disease (21.8 percent more likely), and any health problems (7.6 percent more likely). Additionally, these results mask the heterogeneity of the effect of occupation at different stages of individual's childhood. We find the negative effects of occupation to be the most numerous and the largest in their magnitude for those adults who were exposed to occupation in-utero. Compared to adults who were not affected by occupation in-utero (whose mothers lived in unoccupied territories during the war or who were conceived after the war), affected adults are 37.5 percent more likely to report having bad health, 47.2 percent more likely to have a heart disease, 28.4 percent more likely to have respiratory problems, 29.2 percent more likely to have digestive disorders, 16.7 percent more likely to suffer from spinal problems, 8.7 percent more likely to have had a surgery in the last 30 days, 9.12 centimeters shorter, 37.4 percent more likely to be obese, and 24.5 percent more likely to be unsatisfied with their lives. We find the effect of occupation to be similar for adults who were exposed in infancy, with some effects (on chronic diseases, hypertension, and health problems) being even stronger. Individuals exposed in infancy are also 4.95 centimeters shorter and 4.9 percent more likely to be underweight. Our results suggest that the negative impact of occupation is less detrimental if the exposure occurs later in childhood. Thus, individuals who were exposed when they were a toddler or a preschooler are found to be more likely to report being in bad health, having chronic disease, respiratory problems, digestive disorders, spinal problems, and recent surgery, but no more likely (compared to unexposed adults) to suffer from the heart disease or hypertension. However, the effects on their health are found to be much smaller. There is also no statistically significant effect of occupation on their height in adulthood.

To explore the differences in the effects of occupation on health across males and females, we include an interaction between the occupation dummy and gender in our model and report the results in Table 2. To preserve the space, we choose to focus on five outcomes, including bad health, chronic diseases, height, low height, and life dissatisfaction. First, we find the negative effects of occupation on bad health, chronic diseases, and life dissatisfaction to be significantly larger for females. However, due to the shorter life expectancy of males (compared to females) in Russia and Ukraine, males that we observe in our sample are likely to be the healthiest in their cohort.[19] At the

[19] In 1995, male and female life expectancies in Russia were 58.1 and 71.5 years, respectively (13.4 years difference). In 2013, 65.1 and 76.3 years, respectively (11.2 years difference) (Goskomstat XXXX). In 1995, male and female life expectancies in Ukraine were 61.2 and 72.5 years, respectively (11.3 years difference). In 2013, 66.3 and 76.2 years, respectively (9.9 years difference).
http://www.gks.ru/free_doc/new_site/population/demo/demo26.htm

same time, female sample is less likely to be biased due to the selective mortality as many females survive to their older ages, which could explain why we are finding stronger effects of occupation on their health. If only the heathiest males survive to their older ages, the effect of occupation on males is likely to provide the lowest bound on the effect on their health. Second, we find larger effect of occupation on male height. Males exposed to occupation in-utero are 9.395 centimeters shorter while females exposed in-utero are 7.807 centimeters shorter. The effects are 5.214 and 4.695 centimeters for male and female individuals exposed to occupation in their infancy. This finding is consistent with some of the existing studies (Eveleth and Tanner 1990; Kuhn *et al.* 1991) and could be explained by a delay in menarcheal age as a result of the hunger which subsequently lengthens the period of female growth (Koupil *et al.* 2009).

---

http://ukrstat.gov.ua/operativ/operativ2007/ds/nas_rik/nas_u/nas_rik_u.html

Table 1: Distribution of Municipalities

	USSR	*Combined RLMS-ULMS sample*
Treatment status is known, *N*	3,857	2,021
Share by geographic area, %		
Russia	62.6	63.1
Ukraine	15.9	26.1
Baltic republics	2.2	1.0
Belarus & Moldova	4.0	2.8
Caucuses republics	4.0	1.7
Central Asia	11.3	5.2
Occupied during WWII, *N*	1,419	916
Occupied during WWII, %	36.8	45.3
Share by the length of German occupation, %		
< 1 month	3.0	3.1
1-6 months	14.1	11.5
7-12 months	8.0	7.2
13-24 months	18.1	21.4
25-36 months	39.8	44.8
More than 36 months	17.0	12.1
In 1939 borders, *N*	3,542	1,868

Notes: The list of USSR municipalities is taken from the GADM database of Global Administrative Areas. Web Appendix 1 describes adjustments to the list. The last column shows the distribution of municipalities in which respondents surveyed in 2000-2013 are born between 1935 and 1958.

Table 2: Change in Population between 1939 and 1959 Censuses

	Total		***Males***		***Females***	
	Occupied	***Non-occupied***	***Occupied***	***Non-occupied***	***Occupied***	***Non-occupied***
A. Adjusted						
Age 0-50 in 1939	-0.233	-0.127	-0.386	-0.251	-0.115	-0.025
	(0.038)	(0.027)	(0.042)	(0.030)	(0.036)	(0.026)
Age 0-14 in 1939	-0.164	-0.022	-0.218	-0.048	-0.113	0.002
	(0.061)	(0.042)	(0.062)	(0.043)	(0.060)	(0.042)
Age 15-50 in 1939	-0.281	-0.213	-0.529	-0.449	-0.114	-0.045
	(0.025)	(0.019)	(0.031)	(0.023)	(0.023)	(0.017)
B. Unadjusted						
Age 0-50 in 1939	-0.385	-0.277	-0.563	-0.425	-0.244	-0.152
	(0.038)	(0.027)	(0.042)	(0.030)	(0.035)	(0.025)
Age 0-14 in 1939	-0.281	-0.139	-0.343	-0.172	-0.224	-0.109
	(0.060)	(0.042)	(0.061)	(0.043)	(0.059)	(0.042)
Age 15-50 in 1939	-0.456	-0.389	-0.745	-0.664	-0.255	-0.184
	(0.026)	(0.019)	(0.031)	(0.023)	(0.023)	(0.017)

Notes: The table shows the average log difference in regional population between 1939 and 1959 weighted by the share of USSR region in 1939 population. Regional borders are taken as in 1959 Census. Panel A reports the log difference after adjusting for the pre-war age-specific mortality.

Table 3: Pretreatment Statistics – Economic and Social Indicators before the Second World War

	Occupied regions	*Non-occupied regions*	*t-test*	*N of regions*
Economic Development				
Gross output in millions of rubles per 1000 people, 1935	0.433	0.456	-0.137	64
Gross output in millions of rubles per 1000 workers, 1935	2.579	2.396	0.466	64
Power capacity in thousand kilowatt per 1000 people, 1935	0.056	0.047	0.387	64
Electricity production in million kilowatt-hours per 1000 people, 1935	0.232	0.175	0.528	64
Employment participation rate, %, as of Jan. 1, 1936	0.158	0.168	-0.270	64
Average wage of staff and workers, rubles, March 1936	193.668	204.549	-0.853	64
Health Care				64
Number of hospital beds per 1000 people as of Jan. 1, 1936	2.897	3.304	-0.747	64
Number of doctors per 1000 people as of Jan. 1, 1936	0.539	0.475	0.449	64
Number of birth centers and inpatient clinics for maternal and infant health per 100,000 people as of Jan. 1, 1936	2.411	2.191	0.764	64
Number of pediatric treatment clinics per 100,000 people as of Jan. 1, 1936	0.271	0.238	0.775	64
Coefficient of infant mortality - number of infant deaths to the number of born, 1939	145.360	191.613	-5.816	81
Census 1939				
Percent share of female population	52.494	51.976	0.936	111
Percent share of urban population	36.375	36.128	0.030	111
Crude birth rate - number of infants per 1000 people	32.074	35.692	-2.723	111
Child-woman ratio - number of children under 5 per 1000 women aged 15-49	439.339	506.352	-2.431	111
Percent share of illiterate population (age 15+)	18.983	23.858	-2.368	111
Number of people with secondary education per 1000 people	90.724	77.754	0.850	111
Number of females with secondary education per 1000 people	78.753	68.769	0.613	111
Number of people with higher education per 1000 people	7.137	7.645	-0.149	111
Number of females with higher education per 1000 people	4.384	4.773	-0.172	111
Ethnic fractionalization index, 0=homogeneous	0.321	0.321	0.010	111
Percent share of the largest ethnic group in region	79.183	79.145	0.010	111
Percent share of Jewish population	3.571	0.920	3.328	111
Percent share of deported nationalities	3.999	1.084	1.170	111

Notes: To save space, we only report means, t-test results for the mean difference, and number of observations. The complete table with standard errors is presented in web appendix. The data sources employ different levels of regional aggregation, which explains the varying number of regions (*N*). The regional averages are weighted by the share of the region of birth in the combined RLMS-ULMS sample. The sample includes the RLMS and ULMS respondents surveyed in 2003-2007 and born in the USSR (within 1939 borders) in 1936-1958. The USSR regions outside the Russian and Ukrainian Soviet republics are included in this table, but their contribution to the weighted average is insignificant due to the sample weights. Partially occupied regions are considered to be occupied if more than 20 percent of their territories were sieged or occupied. Using other cut-offs for partially occupied regions or dropping these regions entirely does not affect the results in any significant way.

Sources: The All-Union Census of Population 1939 (www.demoscope.ru); The USSR Country of Socialism (Statistical yearbook), 1936 (http://istmat.info/node/22521); Health and Health Care of USSR Workers, 1937 (http://istmat.info/node/22080).

Table 4: Descriptive Statistics of Health and Non-Cognitive Outcomes by the Treatment Status

	Percentage missing	***Treatment group***	***Control group***
Outcomes			
Bad health	0.005	0.392	0.214
Chronic condition	0.005	0.758	0.649
Chronic heart disease	0.005	0.038	0.035
Chronic respiratory illness	0.008	0.079	0.081
Chronic disease of digestive system	0.012	0.376	0.330
Chronic spinal illness	0.009	0.257	0.249
Hypertension	0.011	0.533	0.532
Surgery in the last 12 months	0.175	0.038	0.042
Health problems in the last 30 days	0.177	0.604	0.497
Residual height	0.026	-2.516	-1.220
Low height	0.026	0.398	0.302
Residual BMI	0.060	1.604	1.604
Low BMI	0.060	0.101	0.122
High BMI	0.060	0.355	0.368
Underweight	0.047	0.006	0.009
Obese	0.047	0.296	0.317
Ever smoked	0.003	0.256	0.384
Consume alcoholic beverages	0.005	0.342	0.497
Consumed alcoholic beverages last month	0.264	0.472	0.670
Life satisfaction	0.006	0.276	0.341
Life dissatisfaction	0.006	0.511	0.428
Loser	0.742	1.952	1.940
Proud	0.743	2.159	2.173
Useless	0.742	2.147	2.153
Controls			
Female	0.000	0.644	0.612
Ethnicity, Russian	0.000	0.413	0.784
Ethnicity, Ukrainian	0.000	0.458	0.093
Ethnicity, Jewish	0.000	0.007	0.003
Ethnicity, Deported nationalities	0.000	0.075	0.023
Ethnicity, other	0.000	0.029	0.088
Birth place, city	0.000	0.301	0.453
Birth place, township	0.000	0.058	0.083
Birth place, village	0.000	0.640	0.463
Mother’s education, college	0.000	0.008	0.036
Mother’s education, secondary	0.000	0.034	0.178
Mother’s education, missing	0.000	0.503	0.352
Father’s education, college	0.000	0.010	0.048
Father’s education, secondary	0.000	0.036	0.160
Father’s education, missing	0.000	0.508	0.392

Figure 1: The Length of Nazi Occupation by USSR Municipality

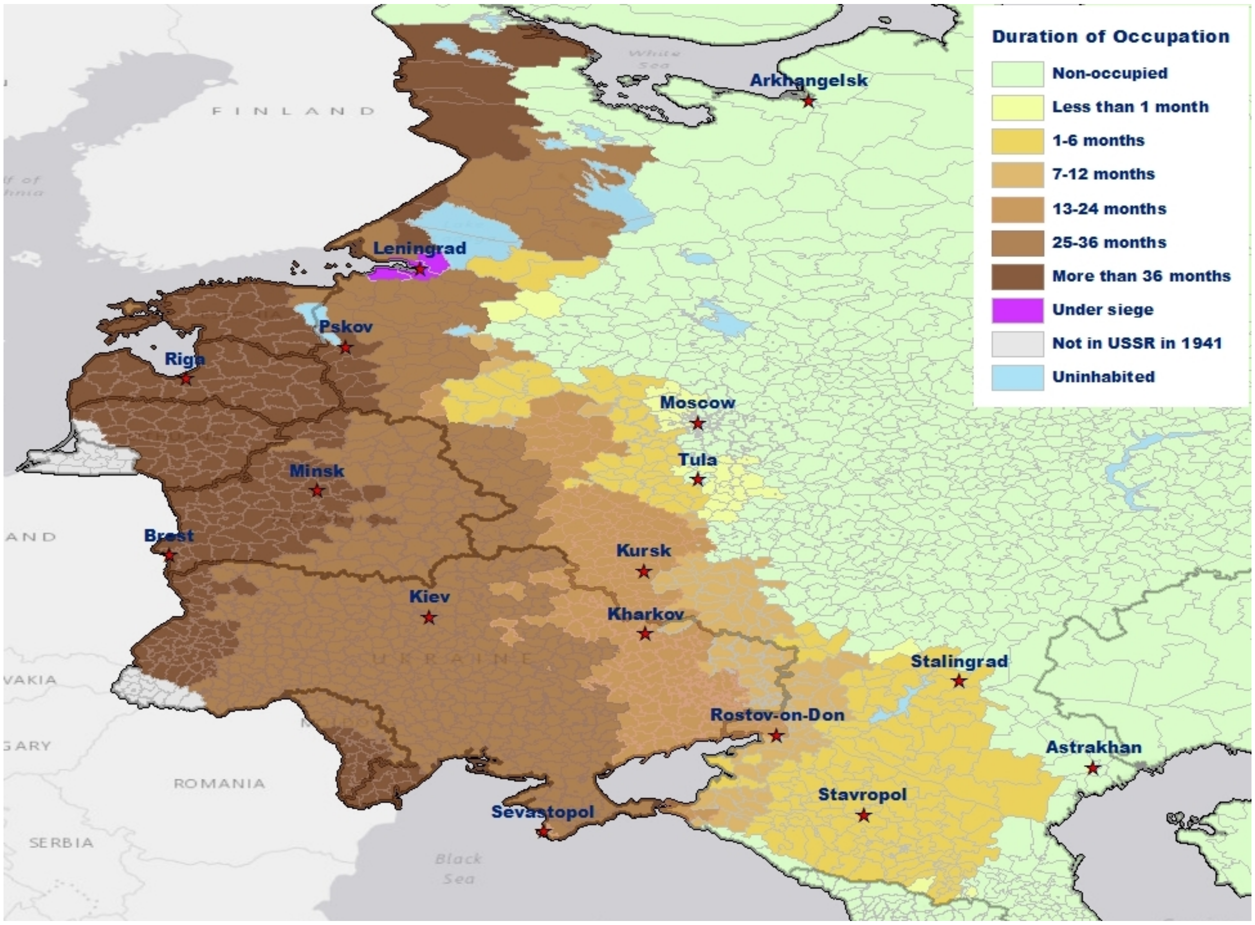
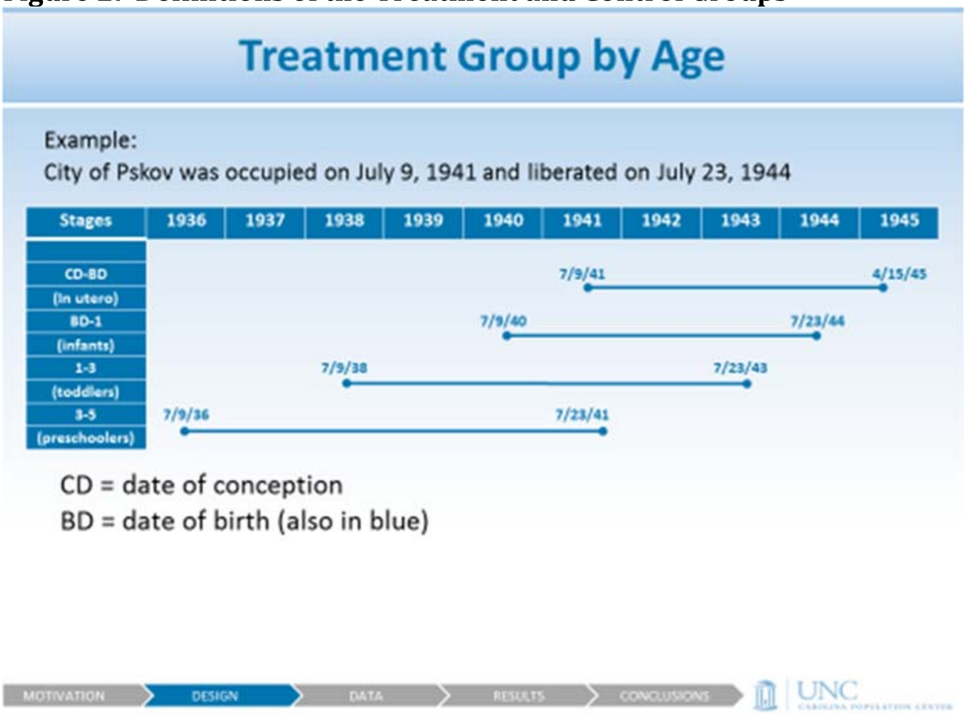


Figure 2: Definitions of the Treatment and Control Groups



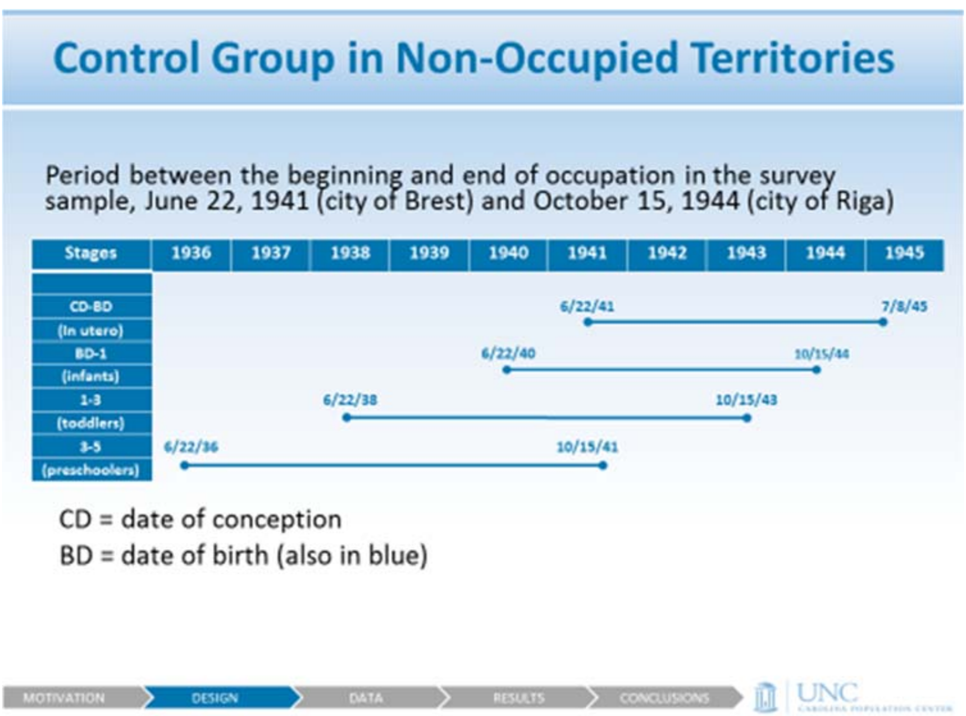


Table 5: Preliminary Results

Effect of Early-Life Shocks on Health Outcomes

	In-utero	Infants	Toddlers	Preschool	All kids
Bad health	0.259	0.167	0.221	0.241	0.220
Chronic illness	0.144	0.305	0.207	0.062	0.133
Heart disease	0.415	0.204	0.160	0.083	0.132
Respiratory disease	0.226	0.296	0.072	0.042	0.091
Digestive disorders	0.106	0.197	0.277	0.222	0.237
Spinal problems	0.175	0.205	0.219	0.113	0.118
Hypertension	0.037	0.267	0.123	0.004	0.073
Std. residual height	-1.298	-0.754	-0.358	-0.419	-0.369
Low height, < 25^{th} pct	0.774	0.580	0.294	0.218	0.281
Underweight, BMI <18.5	-0.002	0.064	0.010	-0.006	0.004
Overweight, BMI >30	0.428	0.352	0.048	-0.003	0.069

N fluctuates depending on outcome and stage of life from 15,400 to 23,000

Std. errors are robust and clustered at region of birth

red = sig at 5%; purple = sig at 10%; grey = insignificant

Appendix

Table A1: Variable Definitions

General notes:

1. *The source for all individual-level variables is RLMS and ULMS.*
2. *Estimation sample covers 2003-2007 time periods; variables are available for all years, unless noted otherwise*

Variable Name	Definition and Sources
	Economic Development
Gross output in millions of rubles per 1000 people, 1935	The USSR Country of Socialism (Statistical yearbook), 1936 (http://istmat.info/node/22521)
Gross output in millions of rubles per 1000 workers, 1935	The USSR Country of Socialism (Statistical yearbook), 1936 (http://istmat.info/node/22521)
Power capacity in thousand kilowatt per 1000 people, 1935	The USSR Country of Socialism (Statistical yearbook), 1936 (http://istmat.info/node/22521)
Electricity production in million kilowatt-hours per 1000 people, 1935	The USSR Country of Socialism (Statistical yearbook), 1936 (http://istmat.info/node/22521)
Employment participation rate, %, as of Jan. 1, 1936	The USSR Country of Socialism (Statistical yearbook), 1936 (http://istmat.info/node/22521)
Average wage of staff and workers, rubles, March 1936	The USSR Country of Socialism (Statistical yearbook), 1936 (http://istmat.info/node/22521)
	Health Care
Number of hospital beds per 1000 people as of Jan. 1, 1936	Health and Health Care of USSR Workers, 1937 (http://istmat.info/node/22080)
Number of doctors per 1000 people as of Jan. 1, 1936	Health and Health Care of USSR Workers, 1937 (http://istmat.info/node/22080)
Number of birth centers and inpatient clinics for maternal and infant health per 100,000 people as of Jan. 1, 1936	Health and Health Care of USSR Workers, 1937 (http://istmat.info/node/22080)
Number of pediatric treatment clinics per 100,000 people as of Jan. 1, 1936	Health and Health Care of USSR Workers, 1937 (http://istmat.info/node/22080)
Coefficient of infant mortality - number of infant deaths to the number of born, 1939	Health and Health Care of USSR Workers, 1937 (http://istmat.info/node/22080)
Percent share of female population	Census 1939
Percent share of urban population	Census 1939
Crude birth rate - number of infants per 1000 people	Census 1939
Child-woman ratio - number of children under 5 per 1000 women aged 15-49	
Percent share of illiterate population (age 15+)	

Number of people with secondary education per 1000 people		
Number of females with secondary education per 1000 people		
Number of people with higher education per 1000 people		
Number of females with higher education per 1000 people		
Ethnic fractionalization index, 0=homogeneous		
Percent share of the largest ethnic group in region		
Percent share of Jewish population		
Percent share of deported nationalities		
	Outcomes	
Bad health	=1 if a respondent evaluates his health as bad or very bad (very good, good, and average are other options)	All years
Chronic illness	= 1 if a respondent has one of the following chronic illnesses: 1) heart disease; 2) illness of the lungs; 3) liver disease; 4) kidney disease; 5) gastrointestinal disease	RLMS: 2000-2013 ULMS: all
Hypertension	= 1 if a doctor has told a respondent that he has high blood pressure	
Life dissatisfaction	= 1 if a respondent is unsatisfied or fully unsatisfied with his life	
	Control Variables	
Mother's education (secondary)	=1 if mother has secondary education or higher *Note: Available in 2006 and 2011 surveys; extrapolated to other years based on individual panel id for cases of consistent reporting of mother's education.*	
Mother's education is missing	=1 if mother's education is missing	
Jewish ethnicity	= 1 if	
Deported nationalities		
Other ethnicity		
Location of birth - Village	= 1 if respondent was born in a village, derevnia, kishlak, aul	
Location of birth - Township	= 1 if respondent was born in a city or an urban-type settlement	
Location of birth - missing	= 1 if location of birth is missing	
Federal districts (dummies)	1) Belarus and Moldova; 2) Armenia, Georgia, Azerbaijan; 3) Kazakhstan, Tajikistan, Uzbekistan,	

Appendix A: Location Coding and Data Linking (Incomplete)

A. Creating the Master List of Municipalities with ID

The data management in this project is complex, as it requires linking numerous data sets without the matching identifiers. We start with the list of geographic units provided by the 2011 GADM database of Global Administrative Areas (www.gadm.org), a spatial database of the location of the world's administrative areas. Administrative areas in the GADM database include countries, regions and lower level subdivisions such as cities and districts. We select all countries of the former Soviet Union and exclude uninhabited geographic areas (such as water bodies) and within-city districts.

The geographic unit in GADM is defined at the three levels of aggregation: 1) countries (ADM0), 2) regions and often country capitals (ADM1), and 3) counties and cities (ADM2). However, the level of aggregation in the GADM database is not consistent across countries. For example, the level two in Russia and Ukraine is slightly larger than a U.S. county, while in Estonia the level two corresponds to the neighborhood or small number of nearby villages. Two bordering geographic units, Pylva county and Pechorskiy district, have similar population size of 25,000 people, but the first one is classified at level one in Estonia, while the second one is at level two in Russia. To make the geographic codes consistent, we use the level two aggregations (ADM2) for larger countries such as Belarus, Kazakhstan, Russia, Ukraine, and Uzbekistan, and level one aggregation (ADM1) for less populous countries such as Baltic countries, Moldova, Armenia, etc. We refer to the selected levels of aggregation as municipality.

Each municipality in our master list is assigned a unique municipality ID. If the area of a large-to medium-sized city is included into the area of county in the GADM classification, then a new line with unique ID for the city is added to the master list. Altogether we have 3,857 municipalities, of which almost 80 percent are located in Russia and Ukraine; see Table 1 for the distribution of USSR municipalities by geographic area.[20]

B. Coding the birthplace

We use the unique municipality ID from the master list to geo code the place of birth of the RLMS and ULMS respondents who are born in the USSR and who report their city or county of birth. There were a few problems associated with geocoding of birthplaces. First, respondents often report their location of birth using the old names these places had in 1930s, 1940s, and 1950s. We had to find current names of their places of birth as they appear in GADM data.

To code location of birth of RLMS and ULMS respondents (*location of birth id*), we use constructed municipality id. We drop individuals whose country of birth is unknown (XX individuals) and who were born outside the Soviet Union (XX observations).

Finally, we merged GADM municipality name and id with individual data from RLMS-ULMS combined file. There were a few problems associated with this merging. First, survey respondents reported old names of their places of birth (names these places had in 1930s, 1940s, and 1950s) and we had to find current names of their places of birth (as they appear in GADM data). Second, some regions of the same country had multiple municipalities with the same name and if a respondent was born in one of them, we could not identify which municipality it was. Therefore, we had to choose randomly from the list of municipalities in a given region of birth. Finally, if respondent did not report his municipality of birth, we assumed that the duration of occupation of his municipality of birth was the same as his region of birth capital.

[20] This table excludes municipalities for which the occupation status is ambiguous such as Kaliningrad region, Klaipeda in Lithuania, Southern Sakhalin, and Transcarpathia, which were acquired by the Soviet Union in 1945 as a result of WWII.

We use the same GADM classification of Global Administrative Areas to geocode the respondents' answers to open-ended questions on birthplace; the geocoding process is described in web Appendix 1. The geocoded individual observations in RLMS-HSE and ULMS are then linked to external databases, including the one on the dates of occupation and liberation. Table 1 provides the distribution of municipalities in two survey samples. The number of municipalities where respondents are born is fairly large (more than 1,400), and most of municipalities are located in Russia and Ukraine. In the combined RLMS-ULMS sample, about half of municipalities were occupied, and of those, 47 percent were under the German occupation between 2 and 3 years. In the extended RLMS sample, the share of occupied municipalities is smaller (about 30 percent), but the variance in the number of days under occupation is significant; the occupation lasted for less than 6 months in 26 percent of the "treated" municipalities, 2 to 3 years – in 32 percent, and 3 years and more – in 13 percent.

C. Linking external data sets

To merge GADM municipalities with our external data on pre-treatment characteristics and demographic characteristics of population from Census 1939, we created four codes, namely, municipality id in 1935, 1939, 1945, and 1959. The main problem associated with merging contemporaneous GADM codes with characteristics of municipalities in 1935, 1939, 1945, and 1959 is changes in the geographic boarders of regions and municipalities as well as frequent changes of municipality names. Therefore, we check the histories of these changes and create municipality ids for 1935, 1939, 1945, and 1959 both in GADM and external datasets for merging.

D. Occupation of the Soviet Union

In our combined RLMS-ULMS sample, we coded 2,135 unique municipalities of birth located in all 15 republics of the Soviet Union. For each of these municipalities we collected information on whether they were occupied during WWII and if they were occupied how long their occupation lasted. Information on the dates of occupation of Soviet municipalities by the German forces and their liberation by the Soviet army was collected using the histories of WWII and maps displaying changes in the battlefield over time. Our primary source of information on the occupation of Soviet cities is Dudarenko *et al.* (1985) - one of the most comprehensive sources on German occupation of the Soviet Union prepared by a team of researchers from the Central Archive of the Ministry of Defense, the Institute of Military History of the Ministry of Defense and the Central Naval Archives of Russian Federation. Dudarenko *et al.* (1985) provides information on the beginning date of occupation and the date of liberation of 727 cities and towns of the Soviet Union including information on partially occupied cities as well as the current names and the names of the cities used during WWII if the names differ. For municipalities not listed in Dudarenko *et al.* (1985), such as rural municipalities, we used other respected sources.[21] If the exact timing of occupation could not be determined based on the available sources, we use the beginning of occupation and the liberation of the nearest geographic municipality (located within X miles) with known information. We calculated the duration of occupation as the difference between the beginning of occupation and the liberation in days. If a municipality was occupied more than once, we calculated the sum of all occupation periods.[22] In cases of missing information, we make the following assumptions. If the municipality of birth of RLMS-ULMS respondent is unknown, we assume that his municipality of birth's occupation

[21] Memory Book (1993, 1998), Newsletter (1995), Romanenko and Lovchakova (2002), history of Akhtubinsk, Alekseevka, Biryuch, Borisovka, Dubrovka, Gubkin, Ivnya, Kletnya, Klimovo, Komarichi, Navlya, Pogar, Rakitnoe, Rognedino, Veydelevka, Volokonovka.

[22] For example, Belgorod in Russia was occupied from October 1941 to February 1943 (473 days) and from March 1943 to August 1943 (140 days). Therefore, in our data Belgorod's total occupation lasted 613 days.

lasted as long as the occupation of the capital of the region of birth (X individuals). If respondent was born in a small country, such as Belarus, Estonia, Latvia, Lithuania, or Moldova, and his region and municipality of birth are unknown, we assume that the occupation of his municipality of birth is equal to the length of occupation of a capital city (X individuals).