The Impact of Forced Migration on Mortality: Evidence from German Pension Insurance Records

Thomas K. Bauer* Matthias Giesecke† Laura M. Janisch‡

June 8, 2018

Abstract

We examine the long-run effects of forced migration from Eastern Europe into post-war Germany. Existing evidence suggests that displaced individuals are worse off economically, facing a considerably lower income and a higher unemployment risk than comparable non-displaced Germans even twenty years after being expelled. We extend this literature by investigating mortality outcomes across the entire life cycle. Using social security records that document the exact date of death and a proxy for pre-retirement lifetime earnings, we estimate a significantly and considerably higher mortality risk among forced migrants compared to non-displaced West-Germans. The adverse displacement effect persists throughout the earnings distribution except for the top quintile. Although forced migrants are generally worse off regarding mortality outcomes, those with successful labor market histories seem to overcome the long-lasting negative consequences of flight and expulsion.

Keywords: Forced Migration, Differential Mortality, Lifetime Earnings, Economic History

JEL-Classification: I12, J61, O15, R23

We are grateful to Ronald Bachmann, Gerard van den Berg, Julia Bredtmann, Bernd Fitzenberger, Joseph-Simon Görlach, Elke Jahn, Rainer Kotschy, and Regina Riphahn for helpful suggestions. We also thank the participants of the meetings of the DFG-SPP 1764 priority program (Essen, 2016), RWI research seminar (2016), RGS conference (Dortmund, 2017), ESPE conference (Glasgow, 2017) and the DFG SPP1764 international conference (Nürnberg, 2018) for insightful discussions. We further thank the team of the research center of the German federal pension insurance (FDZ-RV), in particular Ute Kirst-Budzak, Torsten Hammer and Ingmar Hansen, for supporting the data processing. Fabian Dehos and Jan Wergula provided excellent research assistance. Financial support from the German research foundation (DFG-SPP 1764) is gratefully acknowledged. All correspondence to RWI, Hohenzollernstr. 1-3, 45128 Essen, Germany, EMail: matthias.giesecke@rwi-essen.de.

*RWI, Ruhr-University Bochum and IZA

†RWI and IZA

‡RWI, RGSEcon and Ruhr-University Bochum

1 Introduction

In the aftermath of World War II, approximately eight million Germans were forcibly displaced from Eastern Europe to West Germany. This mass expulsion is one of the largest population movements in history, leaving the bulk of forced migrants to arrive in a country that suffered from major destruction. Against this background and the fact that displaced individuals had to restart with little or no possessions, the integration of this large inflow of migrants was a challenging task. Even though tremendous efforts were taken by the German government to reduce the burden of war-related financial losses, the integration of forced migrants remained a societal challenge. Empirical research has shown that these migrants exhibit considerably lower incomes and a higher unemployment risk even 20 years after the mass exodus (Lüttinger, 1986; Bauer et al., 2013). Emphasizing the role of the local environment, recent evidence further indicates that counties with high industrialization and low inflows were much more successful in integrating forced migrants within two decades after the war (Braun and Dwenger, 2017). An open question that remains is whether flight and expulsion have long-run consequences across the entire life cycle.

This article aims at filling this gap by investigating lifetime outcomes of forced migrants. In particular, we analyze the effect of displacement on mortality, comparing displaced persons predominantly from the former Eastern territories of the German Reich to non-displaced West-Germans. We further examine effect heterogeneity across completed labor market biographies, i.e. pre-retirement lifetime earnings, to learn more about mechanisms that transmit the effects of displacement over the life course.

The identification of the displacement effect is based on exogenous variation from the large-scale inflow of forced migrants into post-war Germany between 1944 - 1950. This migration flow involves several features that allow us to isolate the effect of flight and expulsion on the mortality of forced migrants relative to the indigenous West-German population.

First, the predominant part of the displaced population under study had almost no choice to stay in their home region and migrated in a short period of time, which refrains our analysis to suffer from self-selection problems and cohort effects inherent to many empirical studies on the integration of migrants (Borjas, 1985, 1987, 1999). Second, selective return migration (see Lubotsky, 2007, for a discussion) was virtually impossible due to the iron curtain until 1989. Even after the fall of the wall it remains unlikely that the elderly population under study remigrated after settling in West Germany for more than 40 years. Third, language abilities and transferability of skills should not play a role for the group of migrants analyzed in this study, since most of the displaced persons were German speakers, received their education at German schools, and grew up under the same institutional setting. Finally, it has been shown that pre-war characteristics of displaced and non-displaced persons were largely similar in terms of socio-economic characteristics (Bauer et al., 2013), providing confidence that the indigenous West-German population is a valid control group. To avoid biased estimates due to pre-war regional differences potentially affecting mortality outcomes among displaced and non-displaced, we use information on regional disparities in mortality rates at the time of birth, i.e. we compare displaced to non-displaced individuals in West-German regions with similar pre-war regional-specific birth and mortality rates. Note, however, that we might not identify the average treatment effect on the treated (ATT) because the huge influx of forced migrants into post-war West Germany could have had general equilibrium effects which may also affect the mortality rates of our control group.

Studying the elderly population at risk from age 68 onwards, we find that displaced individuals exhibit a considerably and significantly higher mortality risk than comparable non-displaced persons. Our estimates of the mortality hazard ratio indicate that the mortality risk of displaced men is 15% to 27% higher than for non-displaced men and 4% to 11% higher for displaced women if compared to non-displaced women. Using a measure of pre-retirement lifetime earnings, we further show that the adverse displacement effect persists through large parts of the earnings distribution. Within the top quintile, the adverse effect not only disappears; our estimates even indicate a slightly lower mortality risk for those migrants. This suggests that displaced individuals with exceptionally successful labor market biographies are able to overcome the long-lasting negative consequences of flight and expulsion.

Although more than 65 million people were forcibly displaced worldwide in 2015 UNHCR (2016), little is known about the consequences of forced migration yet.[1] A growing literature examines the effects of forced migration on host countries' labor markets (Ruiz and Vargas-Silva, 2015) or its impact on economic outcomes and integration of displaced persons (e.g. Sarvimäki et al., 2009; Falck et al., 2012; Bauer et al., 2013; Fiala, 2015; Braun and Dwenger, 2017), finding predominantly negative effects[2] of forced migration in the short- and medium-term. Our contribution to this literature is to examine the effects of forced migration across the entire life-cycle, focusing on mortality as an objective health outcome.

Investigating the long-run consequences of forced migration also contributes to the demographic literature regarding the impact of forced migration on mortality (e.g. Saarela and Finnäs, 2009; Saarela and Elo, 2016; Haukka et al., 2017). These studies investigate a specific forced migration episode in Finland, where some areas were ceded to the Soviet Union in the aftermath of World War II. In terms of historical timing and territorial losses, the Finish case is very similar to the German one. Our central finding, that forced migration increases long-run mortality outcomes even decades after being displaced, is consistent with this literature. Although our estimates of the mortality risk ratio (internally displaced to

[1]An extensive literature has examined voluntary migration. However, the process behind forced migration differs in many respects from voluntary migration decisions such that many of the policy implications and conclusions from studies on voluntary migrants may not be applicable to forced migrants (see Ruiz and Vargas-Silva, 2013, for an overview on the forced migration literature).

[2]Positive effects on long-term income of males in Finland have been argued to be due to an accelerated transition from traditional to modern occupations (Sarvimäki et al., 2009). A similar result is documented for those forced migrants in Germany who worked in the agricultural sector before migrating to West Germany (Bauer et al., 2013).

non-displaced) are considerably larger, this seems highly plausible when taking into account that post-war displacement in Finland was rather well-organized and generous in terms of compensations (see e.g. Haukka et al., 2017), while the displacement from Eastern Europe to post-war Germany took place under arguably harsher conditions.

Furthermore, our study relates to the mortality literature that largely agrees upon conditions and experiences early in life to have profound consequences later in life.[3] Since the population that we study is predominantly aged 18 - 45 by the time of expulsion, our results closely reflect what has recently been termed the *cumulative disadvantage* over life (Case and Deaton, 2015, 2017). Although Case and Deaton refer to a different study group[4], the mechanisms behind the cumulative disadvantage are strikingly close to what we find with respect to mortality outcomes of forced migrants in Germany: adverse conditions at the outset of their living in West Germany accumulated to disadvantages over the entire life-cycle. Estimating higher mortality rates in the elderly displaced population indicates that the experience of expulsion has negative effects that are largely irreversible.

Finally, the vanishing displacement effect at the upper margin of the earnings distribution indicates an inverse relationship between lifetime earnings on the one hand and the mortality risk differential between displaced and non-displaced persons on the other hand. This insight is not only in line with the inverse relationship between income and health (or mortality) as one of the most robust empirical findings in economics and medical science.[5] It is also highly

[3]Individuals born in a recession die earlier than individuals born in times of economic prosperity (van den Berg et al., 2006). Recent evidence suggests that entering the labor market in a recession ("unlucky start") increases mortality later in life even if initial losses in earnings have faded (Schwandt and von Wachter, 2017).

[4]The term *cumulative disadvantage* is used by Case and Deaton (2017) to describe recently increasing mid-age mortality rates among non-hispanic whites without college degree in the U.S. (1998 - 2015), triggered by increasingly poor labor market conditions at the time of labor market entry.

[5]For recent empirical evidence on the link between income and health, see Chetty et al. (2016). What remains controversial until today is measuring the causal impact of income on health (or mortality) and vice versa. Low income could explain poor access to health care or less healthy nutrition. One could also think of poor health to reduce the ability for gainful employment or even to prevent people from investing in human capital. See Smith (1999) and Deaton (2003) for discussions.

policy relevant because it highlights the importance of successful economic integration.

The remainder of this article is organized as follows. Section 2 outlines the historical background of the paper. Section 3 describes the data, introduces the empirical strategy and discusses selection issues. Section 4 presents the estimation results and several sensitivity checks. Section 5 concludes.

2 Historical Background

West Germany experienced a major population inflow in the aftermath of World War II. Between 1944 and 1950, about 7.9 million displaced persons or more than 16% of the total West German population at that time entered the country (Reichling, 1958).[6] Figure 1 shows the regions of origin of these migrants. Among them were 4.4 million 'National Germans' (*Reichsdeutsche*) that used to live in areas that were located east of the Oder-Neisse rivers such as Silesia, East Prussia or Pomerania. Additionally, individuals of ethnic German origin (*Volksdeutsche*) that lived as minorities in foreign countries made up 3.5 million migrants, who mainly came from Sudetenland, which was located in Czechoslovakia in close proximity to the German border (Connor, 2007). The map also includes territories that were already lost after World War I but included individuals who were displaced after World War II and thus belong to the primary study group within this paper.

Expulsion of Germans from their homelands took place in three phases starting in October 1944 when advancing Soviet troops entered through the eastern border of the German Reich. The cruelty of the approaching Red Army caused Germans to flee westwards. These so-called

[6] The historical details regarding this mass migration have been documented thoroughly. Douglas (2012) provides a detailed overview about the historical background of the mass migration while Connor (2007) focuses on the integration of forced migrants into post-war Germany. An alternative source from the Eastern European perspective is Eberhardt (2011), who reports the number of Germans that were displaced from Poland and territories incorporated into Poland. Finally, Lüttinger (1986) and Reichling (1958) provide summaries of important aggregate statistics on forced migrants in Germany that were drawn from censuses of the post-war period.

'treks' were very risky, as forced migrants were exposed to extreme weather conditions (e.g. strong winter), malnutrition, as well as air strikes by the allied troops. Many did not survive this exhausting flight. By the end of the war in May 1945 and the unconditional surrender of Nazi Germany, the second phase of displacement, so-called 'wild' expulsions conducted by Polish and Czechoslovakian authorities, started. Those affected by these 'wild' expulsions were forced to leave their homes and personal belongings behind and were put into internment camps, where they had to stay before being transferred to Germany. The Potsdam Treaty from August 1945 started the third phase of expulsion. As a result of this treaty, the eastern border of Germany was shifted westwards to the Oder-Neisse rivers and Germany was divided into four zones of occupation. Furthermore, the treaty legalized resettlement of the remaining German population in Poland and Czechoslovakia, which lasted until 1950.

Until 1950, and especially during the early phases of expulsion, self-selection of the displaced was arguably a minor issue. By 1946, about 6.2 million displaced persons were counted in West Germany, which amounted to almost 80% of the 7.9 million who eventually arrived until 1950 (Reichling, 1958, p. 15).[7] This indicates that the majority of forced migrants was displaced in the early phases under 'wild' expulsions. For this group displacement was almost universal and largely involuntarily.[8]

Displaced persons who finally arrived in West Germany had to start their new lives often without any financial means or a social network, since larger groups of forced migrants from the same home region were often distributed geographically to prevent the formation of 'a state in a state' (Connor, 2007). They entered into a country that was earmarked by vast destruction of the war accompanied by severe food and housing shortage. Tensions between

[7] A small share of displaced persons who were initially located in the Soviet occupation zone eventually arrived in West Germany. This type of east-west sorting amounted to a total of about 500,000 individuals who moved from East Germany (Soviet occupation zone including Berlin) to West Germany from 1950 to 1955 (Reichling, 1958, p. 357 - 358).

[8] The expulsion was not entirely universal regarding some ethnic Germans who remained in Eastern Europe for several decades (Aussiedler/Spätaussiedler), also including German minorities, e.g. in Upper Silesia or Transilvania.

the non-displaced and the newcomers arose quickly, especially in rural areas, where non-displaced were less open to outsiders, experienced fewer damage by the war and possessed fewer knowledge on the actual circumstances of the expulsion. In contrast, displaced persons in larger cities were soon recognized as precious economic resource for war reconstruction. For forced migrants, who had to build up their life from scratch, this often meant working as unskilled laborer rather than as self-employed farmers, which many did before displacement. Further tensions appeared from different religions of non-displaced and newcomers[9], fear of foreign influence by the non-displaced and local politicians that did little to foster integration (Connor, 2007).

The post-war West German government, however, implemented several measures to support the integration of newly arrived Germans from the former Eastern territories of the German Reich and other Eastern European countries. The 'Law of Equalization of Burden' (*Lastenausgleichgesetz*) from 1952 compensated forced migrants as well as the indigenous German population for the loss of property or savings. The idea behind this law was that the ones with high losses from the war were compensated by those who suffered from fewer losses. Another important law was the 'Law for Foreign Pensions', which acknowledges specific periods of work for the pension claims of forced migrants and therewith reduces the burden for forced migrants. This law is crucial for the following empirical analysis, since it allows us to identify displaced individuals in our data.

[9] For example, some catholic Sudeten Germans were placed in protestant North Hesse and Franconia, while many protestant migrants were settled to catholic areas in Lower Bavaria (Connor, 2007).

3 Data and Empirical Strategy

3.1 Data and Variables

The empirical analysis is based on large-scale administrative data from German pension insurance accounts. Specifically, we use individual records that document the termination of individual pensions due to death (*Rentenwegfall*) for the universe of deceased individuals who formerly contributed to the German public pension system. The pension shortfall is available for the entire German population covering the years 1994 - 2013[10], accessible via remote-computing provided by the German federal pension insurance (*Deutsche Rentenversicherung*).[11]

Using these social security records is advantageous in various respects for answering the questions at hand. First, all data points are generated within the administrative process and document the exact date of birth and death. Second, pension shortfall records provide an almost universal picture on mortality in Germany, because they include everyone who has ever been registered for an insurance account and received a pension at some point in time. Since we study a sample of the deceased population, from age 68 onwards, it is almost certain that eligible individuals receive pension benefits. In fact, the data source covers 82% of the total number of death cases according to official mortality statistics, including 96% among men and 75% among women.[12] People who never became actively registered within the German public pension system do not have a pension insurance account and thus do

[10]Precisely, the covered period is from December 1993 to November 2013. Due to the nature of the administrative process, each annual wave of pension shortfall records samples all cases of death from January to November of a given calendar year adding those cases documented for December of the preceding year. For example, the wave of 2012, includes all deaths that were documented from December 2011 - November 2012.

[11]A detailed description of the sampling design (in German) and a code book is provided by (Forschungsdatenzentrum FDZ-RV, 2017).

[12]These shares are obtained from relating the total number of deaths reported in the official mortality statistics for Germany (Federal Statistical Office, 2016) to the number of deaths in the pension shortfall records that we use as primary data source.

not appear in the data. Men exhibit high labor force participation rates and thus most of them do have pension insurance accounts.[13] Among women, the coverage rate is smaller because some of them never worked. In 1960, when women in our sample were at prime working age, labor force participation amounted to only 47% for women in contrast to 90% among men (Federal Statistical Office, 2018). Lower labor force participation rates among women translate into lower register rates in the public pension system. For women who never worked, the only way of becoming registered in the public pension system would be the accumulation of periods of child raising.

To account for differences in regional birth and death rates we supplement the analysis with historical mortality data from statistical yearbooks of the German Reich.[14] Systematic differences in mortality rates between the country of origin (former Eastern Territories of the German Reich) and the destination country (West-Germany) would compromise our results. Variation in regional mortality patterns is possible, for example, due to the effects of maternal nutrition and health status on mortality (see Barker, 1990, 1995). Health and mortality outcomes could also be influenced by general living standards and regional-specific industries that may involve physically demanding work. In fact, while Silesia and Bavaria had quite low life expectancies, regions in the North-West of Germany reported higher life expectancies (Kibele et al., 2015). Finally, differential fertility rates across regions could affect cohort sizes of the displaced population differently than for the non-displaced. Depending on the region of origin, this could play a role when comparing mortality outcomes. To encounter these challenges, we use birth and death rates for each region of the German Reich, allowing us to control for ex ante differentials in mortality across regions. Both the annual birth and

[13] Only 4% of male cases of death documented in official mortality statistics are not covered by the pension shortfall. Probably the most plausible explanation is that a considerable share of these men worked as civil servants from the beginning of their employment biography. Pensions of civil servants are tax-financed and handled separately from the public (pay-as-you-go) system.

[14] We use an electronically preprocessed version of the original print by Besser (2008), using information only for those birth cohorts that are sampled in our primary data from pension shortfall records, i.e. persons born between 1885 and 1925.

death rates are calculated by relating the number of births and deaths respectively to the number of residents within each region in a given year. The corresponding variation across region and time as shown in Figure 2 is used to account for regional disparities in mortality patterns, covering the relevant pre-treatment period.

3.1.1 Separating the Displaced from the Non-Displaced

We identify displaced persons based on the legal framework of the 'Law for Foreign Pensions' (LFP) (*Fremdrentengesetz*) enacted in 1960. The goal of the LFP was to acknowledge periods of work or work-related periods (e.g. unemployment, illness, pregnancy and child-raising) accumulated previous to expulsion from Eastern Europe for the pension claims of forced migrants in the West-German public pension system. The value of an acknowledged period was based on the level of education and the occupational group of the displaced person, referring to average earnings of these groups in Germany. For each person with LFP-related pension claims, the data document the region of origin, the amount of acknowledged earnings and the length of the acknowledged period in months.

Based on this information, we define forced migrants as those persons who entered West-Germany between 1944 and 1950, encompassing the period of mass expulsion, who obtained pension claims referring to the LFP and who were expelled from the former Eastern Territories or one of the Eastern European countries that were affected by the mass expulsion. Unique identification of forced migrants via LFP pension claims requires the existence of such claims. The LFP provided a generous program that aimed at treating forced migrants similar to the West German population in terms of pension claims. Since the LFP accounted not only for working times but also for other labor market related periods, it is unlikely that displaced persons were not covered or did not call on these claims, because otherwise their old age incomes would deteriorate. Non-displaced persons are defined to be individuals with German citizenship without recent migration history, living in West-Germany. We drop all observations that neither satisfy the definition of the displaced nor the one of the non-displaced.

3.1.2 Definition of Mortality

Throughout the paper, we define mortality as the age at death. From the observed monthly date of birth and date of death (the latter is documented from pension shortfall due to death), we calculate the age at death for each individual in the data. This mortality measure is then used, in the main analysis, to model the probability of dying conditional on having survived until the respective age.

3.1.3 Approximating Lifetime Earnings

To examine mortality differentials across the entire earnings distribution, we approximate pre-retirement lifetime earnings using a variable that includes the individual sum of so-called earnings points (EP). This measure documents pension claims that predominantly consist of plain labor earnings. One EP from labor earnings of individual i in year t is defined as $EP_{it} = \frac{y_{it}}{\overline{y}_t}$, where y_{it} are labor earnings of individual i in year t and $\overline{y}_t$ are the average labor earnings of all contributors of the public pension system in that year. Intuitively, one EP reflects the relative earnings position of each individual in a given year. For example, an employee receives exactly one EP per year if she contributes at average earnings, two EP if she contributes at twice the average earnings. For each individual, our data include the sum of annual earnings points that are given by $EP_i = \sum_{t=1}^{T} EP_{it}$, where T is the last year of gainful employment before an individual retires. Although EP are not a perfect measure of earnings because they also include creditable periods, e.g., from education or child-raising (see online appendix A for a detailed description of its limitations), we argue that they are a fairly reasonable proxy of pre-retirement lifetime earnings.

The actual distribution of earnings is shown in Figure 3, reflecting the typical labor force patterns of men and women within the observed birth cohorts: while men overwhelmingly work in full-time jobs and often have working biographies of 40 years or more (sample mean: 44 EP), women either work in part-time jobs or not at all (sample mean: 20 EP). The patterns in Figure 3 summarize these labor market histories by showing a remarkable spike for women at low EP-values (between 8-12), that are due to creditable periods of child-raising (max. 2 EP per child until 1992) and another spike between 20 - 30 EP, reflecting gainful employment that is dominated by part-time work. Among men, the patterns are much more homogenous with the highest densities between 40 - 60 EP, reflecting typical full-time work biographies. Table 1 further documents substantial gender differences in mortality patterns and labor market histories. Women strongly outlive their male counterparts and were less active in the labor market, having collected a lower number of earnings points and obtained fewer contribution months to the pension insurance.

3.1.4 Observational Plan and Estimation Sample

The final sample documents cases of death between 1994 and 2013 covering the West-German population. The observational plan is depicted in a lexis diagram (Figure 4), spanning the space of sampled birth cohorts (baseline: 1885 - 1925) over age and years. Based on this observational plan, the final estimation sample consists of 4.98 million observations (2,985,403 observations for females and 1,992,963 observations for males).

The particular cohort choice is motivated by identifying forced migrants according to their pre-migration pension claims. For this purpose, we need to ensure that individuals have reached working age once they migrated. Taking into account the beginning of mass expulsion in 1944, we set the minimum age at migration to 18 by restricting the sample to birth cohorts that were born in 1925 or earlier. Choosing younger cohorts (born after 1925), while possible, increases the likelihood that displaced individuals are mistakenly sampled as non-displaced Germans because they did not have the opportunity to collect creditable pension claims previous to migration. The lower margin of the cohort distribution (i.e. older persons) is unrestricted and only constitutes that individuals born in 1885 are the oldest persons to reach the observation period (starting in 1994) in terms of survival.[15] Our baseline cohort choice balances the sample at two margins: it allows to cover a large number of birth years and ensures to uniquely distinguish the displaced from the non-displaced. Later, we relax this restriction by also including younger cohorts up to 1929 to show the sensitivity of our baseline estimates against this choice.

3.2 Selection

3.2.1 Sampling of Deceased Individuals and Age Selection

The sample consists of deceased individuals only, which is a direct consequence of how the administrative records document pension shortfall due to death. The fact that individuals enter the sample only if they die within the observation period introduces a specific type of selection. The population at risk of dying, however, includes everyone who receives any type of pension.[16]

Two margins require attention in this regard: the observation period (calendar years 1994 - 2013) and the cohort selection (birth years 1885 - 1925), implying that all persons in the sample have at least celebrated their 68th birthday. Thus the sample is positively selected in age: all results and interpretations refer to the elderly population at the upper margin of the age distribution. This means that we study a population that is predominantly retired (more than 99% have claimed a pension by the age of 68), offering the opportunity to observe the entire pre-retirement earnings history in terms of completed labor market

[15]The cohort distribution is depicted in Figure B.1 (online appendix), showing the percentage of each cohort in the sample.

[16]For example, the population at risk also includes persons who out-migrate. They only need to have accumulated pension claims in Germany at some time, for example by employment that is subject to social security contributions or by periods of child raising. In this case a persons' death is documented in the shortfall records because pension payments are terminated.

biographies. Comparing earnings biographies is more feasible if they are completed by the date of retirement, which holds for our sample.[17] This allows us to estimate the mortality effect of displacement across the lifetime earnings distribution to reveal heterogeneity in connection to labor market integration.

By the end of the observation period in 2013, the youngest cohort (1925) has reached age 88, implying that we only face minor selection issues concerning the unknown population of those who are not sampled because they are still alive.

3.2.2 Positive Selection of Forced Migrants

An important question regarding the influx of forced migrants into West Germany is who eventually arrived in West Germany after the expulsion. This comprises concerns not only about the prevalence of death during the flight but also about self-selection into East and West Germany. The crucial point about all of these sources of selection is that they potentially induce a positively selected pool of displaced persons. Some of the selection mechanisms and pathways are not well documented[18] or even unobservable. However, based on aggregate statistics we substantiate that the displaced population was arguably a positively selected group upon arrival in West Germany, implying that our estimates constitute a lower bound of the effect of forced migration on mortality.

First, in 1950 the displaced were over-represented at younger ages (up to age 40) and under-represented in older age groups (Reichling, 1958, p. 54 - 56). One explanation for the different age composition of displaced persons is that missing individuals among the displaced perished in the process of occupation and expulsion (Reichling, 1958). Even if one were to argue that low- and medium-aged persons should be subject to war-specific losses, this should apply to the non-displaced German population similarly and therefore would

[17] Comparing earnings biographies is more feasible if they are completed by the date of retirement because this rules out false comparisons of earnings biographies that only differ due to age differentials.

[18] One exception is the crossing of the inner-German border before the iron curtain (i.e. the wall) was built.

not confound our comparisons of the displaced to the non-displaced. Thus, for those who eventually arrived in Germany, the over-representation of displaced individuals at younger ages suggests a type of positive selection, especially when viewing younger people to be more healthy and more resistant to the strenuous flight.

Second, sorting from East to West Germany amounted to a total of 500,000 displaced individuals who migrated from East to West between 1950 to 1955 (Reichling, 1958), accounting for about 5% of the displaced persons in West Germany. At the same time, only 50,000 displaced persons moved into the opposite direction from West to East Germany. This asymmetric movement towards West Germany can be explained historically by a much poorer economic development in East Germany already towards the mid-1950s. To prevent further out-migration the wall was eventually built in 1961. The dominance of East-to-West movements speaks in favor of positive selection if those with higher motivation or abilities are actually attracted to invest in moving towards a more promising economic environment in West Germany. In addition to this, changes in the initial distribution of displaced individuals across Germany were rather unlikely. After being assigned to a specific locality in West-Germany, strict moving restrictions prevented displaced persons from moving (see Müller and Simon, 1959, for details). These restrictions were relaxed in 1949 but remained rather strict thereafter.

3.2.3 Attrition

The share of forced migrants in our sample of about 0.7% seems low when having in mind that, just after World War II, forced migrants represented no less than 16% of the West-German population. However, due to the birth cohort restriction, the predominant part of forced migrants in our sample is between age 25 to 45 at the time of entering the country (see Figure B.1 in online appendix). Only 30% of the total inflow of the displaced in West Germany belonged to this age group (Reichling, 1958). Moreover, our sample only contains those migrants that have at least celebrated their 68th birthday, i.e. have not died before 1994. Finally, having collected any earnings points that count for pension claims is a prerequisite to enter the sample. Despite these requirements and the age restriction, our final sample still includes more than 33,000 displaced persons (12,072 men and 21,031 women).

The nature of the sampling and the cohort-age restrictions partly explain the small number of displaced persons in the final sample. While this reduces concerns of endogenous self-selection, different mortality patterns between displaced and non-displaced, before they reach age 68, remain relevant for the unbiased estimation of the displacement effect. A quantification of these different mortality rates is possible based on census statistics from 1950 (Reichling, 1958). To scale down the initial population share, we use the coverage rate of the pension records relative to official mortality statistics of 82% of the total population.[19] The attrition rate calculated from 1950 to 1993 is 81% among the displaced and 58% among the non-displaced.[20] A higher attrition rate among the displaced before age 68 coincides to a more positively selected pool at the beginning of the observation period in 1994. In terms of the measured mortality outcomes, this is plausible to the extent that healthier people tend to grow older. As for the positive selection in age, this would imply that we estimate a lower bound of the true displacement effect.

3.3 Empirical Strategy

We use a discrete time version of the Cox proportional hazards model (see, for example, van den Berg et al., 2006; Palme and Sandgren, 2008; Kalwij et al., 2013) since age is measured in years and thus the event death can only occur at the level of age in years.

[19]This is necessary to make these population counts comparable to the pension shortfall records that only document participants of the public pension system.

[20]The mortality rates are higher among men (displaced: 95%, non-displaced: 75%) compared to women (displaced: 74%, non-displaced: 50%). The calculation assumes that everyone has deceased in the observed birth cohorts by the end of the observation period. The calculations are available from the authors upon request.

Integrating the continuous time hazard function $\lambda_{1i}(a) = \lambda_0(t)\exp(\alpha D_i + \beta' x_{ia})$ over the interval $[a, a+1)$ yields the discrete time hazard function

$$\gamma(a) = 1 - \exp(\exp(\alpha D_i + \beta' x_i + \delta_a)) \tag{1}$$

with extreme-value distributed survival spells (complementary log-log model) for i individuals. The survival time a reflects age measured in years. The indicator D_i takes the value one for displaced individuals from the former Eastern territories and zero for non-displaced West-German individuals. Our primary interest lies in the displacement effect, measured by the parameter α. We control for time-invariant observable characteristics that may affect the mortality hazard in vector x_i, which includes, in all specifications, birth cohort dummies to balance out differences in birth cohort representation between displaced and non-displaced persons. In some specifications, x_i further includes earnings points (EP) as a proxy of lifetime earnings and regional-specific birth and death rates to account for initial regional disparities in fertility and mortality patterns that may relate to mortality outcomes of the displaced and the non-displaced later in life.

The sampling structure implies that individuals are at risk beginning at age 68. For any subsequent age, we include duration indicator variables into the model, estimating the parameters δ_a.[21] This procedure is possible due to the large data set and is particularly advantageous because it leaves the baseline hazard in its most flexible version. All parameters in Equation (1) are estimated separately for men and women.

[21] Due to the small number of observations at the upper margin of the age-at-death distribution, we use 40 duration indicator variables for the ages $a = 68, ..., 107$. Estimating δ_a for ages above 107 is difficult because we observe only very few persons to survive this age (e.g. the maximum age reached by one single person in the sample is 111).

4 Results

4.1 Basic Results

Comparing the age at death distribution at the sample mean yields a first indication that non-displaced West-Germans outlive their displaced counterparts considerably (Table 1). The gap is statistically significant and amounts to 0.9 years among men and 1.6 years among women. How this translates into differences in the mortality distribution between displaced and non-displaced persons is illustrated in Figure 5. Two things are noteworthy: first, non-displaced Germans live considerably longer than displaced persons. Second, the mortality differential is larger among women. For example, observing individuals at risk beginning at age 68, Figure 5 documents an 11 percentage point gap between displaced and non-displaced women at age 88, i.e., at this age 73% of displaced women have deceased in contrast to only 62% of non-displaced women. Among men, the difference in the mortality distribution is smaller but still considerable in size. At age 88, we measure a 6 percentage point difference in mortality, based on 80% displaced men and only 74% non-displaced men who have deceased by that age.

To facilitate the interpretation of our estimates from Equation 1, we present estimated mortality hazard ratios (HR) instead of coefficients or marginal effects. Our estimates of the mortality hazard ratio of displaced to non-displaced men range from 1.15 to 1.27 (Table 2), indicating that the mortality risk of displaced men is 15% to 27% higher than for non-displaced Germans. Similarly, for women the estimated mortality hazards translate into a 4% to 11% higher mortality risk of displaced women compared to non-displaced women.[22]

[22]To further support that inter-regional variation reflected by the displacement indicator does matter, we have also implemented a multilevel mixed-effects linear regression with regions of origin (Eastern Europe vs. West Germany) at the higher level and individuals at the lower level. From this exercise we infer that the share of variation in mortality explained at the regional level ranges from 2.7% to 7.7% (men) and 0% to 7% (women); detailed results are available from the authors upon request. Although the random effects parameters indicate only a small variance contribution at the regional level for women once including birth cohort dummies (close to zero), the overall results indicate that inter-regional variation does matter. A large

The smallest estimates of the effect of displacement on the mortality risk are obtained from specifications that only include the displacement indicator, a set of duration dummies and birth cohort dummies (columns (1) and (5) of Table 2). To make the two populations under consideration as homogeneous as possible, we also control for lifetime earnings[23] and pre-war disparities in regional birth and death rates. While including lifetime earnings alone changes the results only very little (columns (2) and (6) of Table 2), the difference in estimating the displacement effect on mortality increases when controlling for regional birth and death rates (columns (3) and (7) of Table 2). Both of these variables covary positively with the mortality rate and the displacement indicator, since pre-war birth and death rates were considerably higher in Eastern Europe than in West Germany (see Figure 2).

4.2 Differential Mortality Across the Earnings Distribution

The baseline estimates suggest that pre-retirement lifetime earnings (EP) do not influence differential mortality risks between displaced and non-displaced persons. So far, however, these estimates are measured at the mean of the earnings distribution.

Estimating the displacement effect separately by earnings quintile reveals effect heterogeneity across different margins of the distribution (see Table 3). While the displacement effect is large and positive within the first and second quintile, it declines when moving towards the upper margin of the earnings distribution (for a graphical presentation compare to Figure B.2 in online appendix). Within the fourth quintile, the measured effect is signifi-

fraction of variation is explained at the individual level but given that mortality is determined by various – presumably individually driven – factors, the estimated variance shares at the regional level can still be considered as substantial.

[23]It should be noted that lifetime earnings may represent a "bad control" (see Angrist and Pischke (2009) for a discussion), since they may themselves be an outcome and thus be affected by displacement. In particular, comparing displaced to non-displaced individuals at a given value of lifetime earnings may differ by some unobserved characteristics that compensate the initial disadvantage in the earnings potential of the displaced. This problem is arguably a minor one since including earnings into the baseline specifications does not considerably change the estimated hazard ratios (see Table 2, comparing column (1) and (2) for men and column (4) and (5) for women).

cantly smaller compared to the third quintile. In the top quintile, the displacement effect is negative and significantly differs from zero. These principal patterns hold for both men and women and are robust across different specifications.

These results indicate that the adverse displacement effect, as previously measured at the sample mean, appears to be driven by individuals in the lower parts of the earnings distribution while displaced individuals with exceptionally successful labor market biographies manage to overcome the long-lasting negative consequences of flight and expulsion. However, analyzing the displacement effect across the entire earnings distribution may introduce another layer of selection. In particular, if a displaced and a non-displaced person have identical lifetime earnings, they are also likely to differ in some other unobservables. If, for example, forced migration reduced lifetime earnings on average, then a displaced person with given lifetime earnings should have more favorable unobserved characteristics compared to a non-displaced person with identical lifetime earnings.[24] Despite these concerns about the bad control variable and connected issues of selection, heterogeneity of the displacement effect across the earnings distribution is particularly informative about the long-run transmission channel of the effect that is arguably mediated through poor integration.

4.3 Sensitivity Analysis

In a first sensitivity check, we consider only the non-displaced population living in the federal state of Bavaria in order to make the non-displaced population more comparable to forced migrants for two reasons. First, regional-specific birth and mortality rates in Bavaria, measured at time of birth, are much more similar to those of the former Eastern Territories than for overall West-Germany (Figure 2). Second, Bavaria exhibited by far the largest inflow of displaced persons in the aftermath of World-War II due to its geographical location

[24] Arguably, this would be the case since the displaced person ended up earning the same despite of being displaced.

in the south-east of Germany.[25] The spatial distribution of displaced persons also did not change much over time, a fact that has been documented by Schumann (2014), who finds that population shocks in the aftermath of World War II are highly persistent. Table 3 shows that our main estimation results change only very little for the full earnings point distribution as well as for the different quintiles if we compare only non-displaced Bavarians to the displaced rather than using the full West German sample of non-displaced.

To identify displaced persons in the data, our core estimates are based on a strong age restriction that only allows individuals to enter the sample if they were born until 1925. Columns (2)-(3) and (6)-(7) of Table 4 show the estimation results when relaxing this cohort restriction by including also younger cohorts. The adverse displacement effect appears to slightly increase when relaxing the baseline restriction (born until 1925) by further including the cohorts up to 1927 or 1929.[26] As discussed above, adding younger cohorts to our sample increases the likelihood of mis-classifying displaced as non-displaced. One would expect the estimated mortality differential to become smaller as an increasing number of displaced with worse mortality expectations erroneously moves to the comparison group of non-displaced. However, this is not entirely clear due to changes of the relative representation of the displaced across birth cohorts. Nevertheless, the overall changes are fairly small indicating that our baseline estimates seem to produce the most conservative results of the displacement effect on mortality.

A final concern is that heterogeneity between displaced and non-displaced may arise from the length of individual working biographies. For example, individuals who start their working careers early may work in manual jobs rather than obtaining a university degree and following an academic track. Since this type of selection may correlate with health impair-

[25] A total of 1,937,297 displaced persons were registered in Bavaria by 1950. In relative terms, this amounted to 21% of the Bavarian population at that time (Reichling, 1958).

[26] We also conducted sensitivity checks by moving the cohort restriction into the opposite direction, thus shrinking the sample to older cohorts (youngest cohort is 1923 or 1921). Again, the results only change little and are available from the authors upon request.

ments accumulated over time, we take a sub-sample of individuals that have contributed to the German public pension system at least for 40 years and thus must have started their working biography at young ages. While the results for males stay unchanged, the displacement effect strongly changes and becomes negative for females (columns (4) and (8) in Table 4). Contributing to the pension scheme for at least 40 years is, however, a tough restriction that is only fulfilled by 60% of the men and 10% of the women. This indicates that the female sample in particular is positively selected in terms of labor force participation. In contrast, the results for the restricted male sample (at least 40 years of contribution) are almost unchanged if compared to the baseline estimates. This is not surprising as for men it is much more common to contribute 40 years or longer and thus the sample composition remains rather similar. The results for long contribution periods are in line with the findings for the top earnings quintile (Table 3), suggesting that the adverse displacement effect disappears or even becomes positive among displaced individuals who are successfully integrated into the labor market.

5 Conclusion

This paper analyzes the long-run consequences of forced migration on mortality. In the aftermath of World War II, almost 8 million Germans that were expelled from the former Eastern territories of the German Reich and other Eastern European countries arrived in West-Germany. We use this natural experiment to identify the displacement effect on differential mortality patterns of displaced and comparable non-displaced persons.

Our results show that the mortality risk of displaced individuals is substantially higher than among comparable non-displaced Germans. These estimates are robust across several specifications but they are also much larger than the ones documented for a similar forced migration episode in Finland, where some areas were ceded to the Soviet Union in the aftermath of World War II. However, this seems highly plausible when taking into account that post-war displacement in Finland was rather well-organized and generous in terms of compensations (see e.g. Haukka et al., 2017), whereas the displacement from Eastern Europe to post-war Germany took place under arguably harsher conditions especially during the early phases of expulsion.

While controlling for pre-retirement lifetime earnings at the mean seems to have only little influence, estimating the displacement effect within each quintile of the earnings distribution reveals that the adverse displacement effect is driven by displaced individuals in the lower parts of the earnings distribution. At the upper margin of the distribution, the adverse displacement effect not only becomes smaller in magnitude but turns positive. This effect heterogeneity, equally measurable among men and women, provides evidence that displaced individuals with exceptionally successful labor market biographies manage to overcome the long-lasting negative consequences of flight and expulsion.

The documented mortality gap between displaced and non-displaced persons may operate through two distinct channels. The first one refers to the direct and long-lasting link of the traumatizing event of displacement on health outcomes later in life and its correlation with mortality. The second channel reflects indirect effects of poor integration of forced migrants into the labor market and the society. Although the German legislation adopted a number of different laws that aimed at reducing the burden of war related financial losses (e.g. *Lastenausgleichsgesetz*), the integration of forced migrants remained a societal challenge. Despite an enormous labor demand in post-war Germany, unemployment rates among the displaced remained higher compared to the non-displaced even in the long-run (Reichling, 1958; Lüttinger, 1986; Bauer et al., 2013). Resentments of the non-displaced Germans towards the newcomers, which already started arising during the post-war period that was characterized by severe food and housing shortages due to major destruction, also impeded their integration.

The *cumulative disadvantage* over life plausibly describes that the consequences of expulsion reached far beyond local displacement and the loss of material possessions. The atrocities experienced by forced migrants during flight and expulsion seem to have caused long-lasting wounds that did not heal by the end of the war. For some, this was the beginning of poor integration into society and the labor market that cumulatively contributed to long-run disadvantages. The cumulative disadvantage represents a meaningful explanation for the transmission mechanism over life, thus allowing to attribute differential mortality to forced migration even decades after the expulsion took place.

Appendix Intended for Online Publication

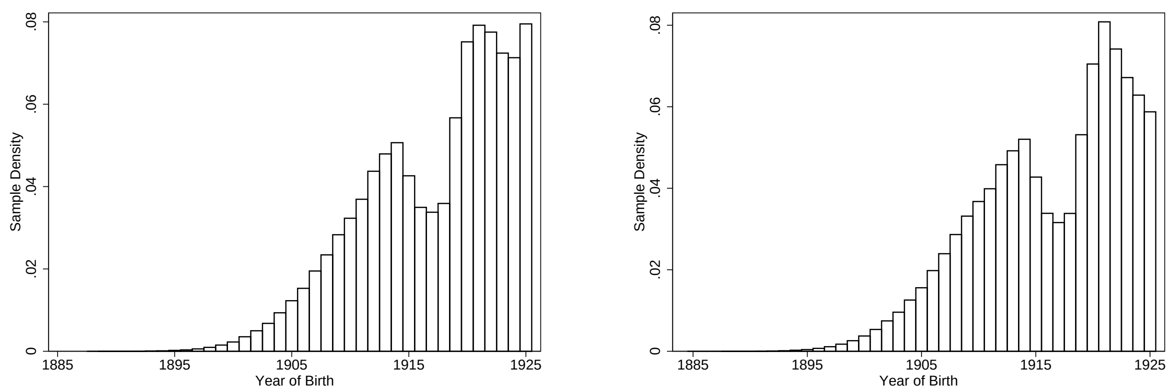
A Approximating Pre-Retirement Lifetime Earnings by Earnings Points: Limitations

One important limitation of our data is that the EP measure also contains information other than plain labor earnings, including all labor market related information as needed by German federal pension insurance to calculate monthly pension benefits. This information comprises the entire labor market history (pre-retirement lifetime earnings, creditable periods of unemployment), education (creditable periods of vocational training or higher education), family background (creditable periods of child raising) and health-related aspects (creditable periods of illness). While not reflecting plain labor earnings, it is apparent that EP are a particularly measure of a wide range of observable socio-economic characteristics. Effectively, EP are completed earnings biographies that capture large parts of variation that may be part of the process that also determines mortality outcomes. The reason why we still refer to EP as "earnings" or "pre-retirement lifetime earnings" is that it predominantly consists of total labor earnings previous to retirement. Although we cannot show the share of labor earnings within individual EP, EP strongly correlate to labor earnings. The amount of EP from other sources than earnings are strongly limited because there exist defined maxima (e.g. EP are limited to a maximum of 3 for education and a maximum of 2 per child for child-raising for children born until 1992). In this regard, a person who never worked will have low values of EP or, put differently, it is virtually impossible to reach the top 20% of the earnings point distribution without high labor earnings and/or long periods of gainful employment.

Another limitation is that labor earnings are top-coded since they are measured in terms of contributions to the German public pension system that are subject to a contribution ceiling. This threshold is annually fixed and introduces censoring at the upper margin of the earnings distribution. The contribution ceiling is adjusted every year. For example, it was fixed at 85200 Euro in the most recent observation year 2013 and, evaluated at average earnings in this year (33659 Euro), the maximum amount documented in the data is $85200/33659 = 2.5$ earnings points. However, top-coding is not a drawback in our empirical setting because we mainly use the earnings distribution, stratified by quintiles, to obtain separate estimates of the mortality hazard. Furthermore, Figure 3 shows that there are no significant anomalies such as bunching of individuals in upper regions of the distribution.

B Additional Figures

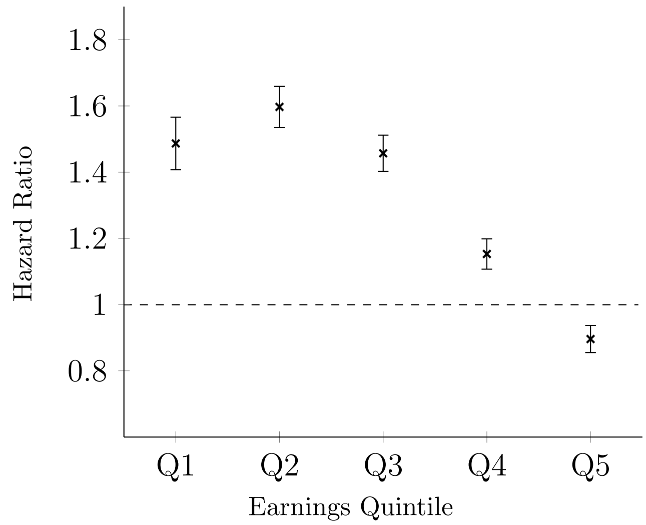
Figure B.1: Birth Cohort Distribution



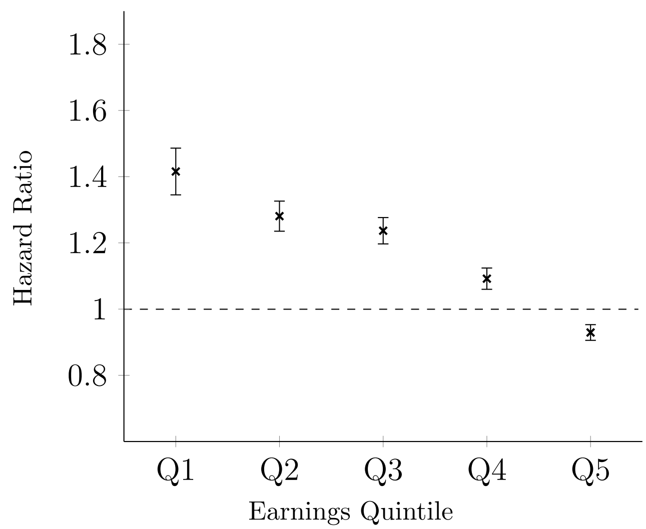
(a) Men (b) Women

Source: Own calculations based on pension shortfall records 1994 - 2013 (FDZ-RV–RTWF94-13DemoRWI). *Note:* The upper margin of the cohort distribution is limited by the age restriction imposed to identify displaced persons. The lower margin is unrestricted, only limiting the sample to individuals who reach the observation period (starting in 1994) in terms of survival. Older cohorts are thus represented to a lesser extent. The figure also shows that fewer individuals who were born during World War I (1914 - 1918) are sampled due to lower birth rates at that time.

Figure B.2: Mortality Hazard Ratio of Displacement across the Earnings Distribution



(a) Men



(b) Women

Source: Own calculations based on pension shortfall records 1994 - 2013 (FDZ-RV–RTWF94-13DemoRWI).
Note: Plotted values are hazard ratios and corresponding 95% confidence intervals from the Cox proportional hazards model (vertical axis). Estimates are obtained separately within each earnings quintile (horizontal axis).

References

Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Barker, D. J. (1990). The Fetal and Infant Origins of Adult Disease. *British Medical Journal 301*(6761), 1111–1111.

Barker, D. J. (1995). Fetal Origins of Coronary Heart Disease. *British Medical Journal 311*(6998), 171–174.

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

Besser, C. (2008). Grunddaten zur Bevölkerungsstatistik Deutschlands von 1871 bis 1939. *GESIS: Köln. ZA8295 Datenfile Version 1*(0).

Borjas, G. J. (1985). Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants. *Journal of Labor Economics 3*(4), 463–489.

Borjas, G. J. (1987). Self-Selection and the Earnings of Immigrants. *American Economic Review 77*(4), 531–553.

Borjas, G. J. (1999). The Economic Analysis of Immigration. *Handbook of Labor Economics 3*, 1697–1760.

Braun, S. T. and N. Dwenger (2017). The Local Environment Shapes Refugee Integration: Evidence from Post-war Germany. *School of Economics and Finance Discussion Paper No. 1711, University of St. Andrews.*

Case, A. and A. Deaton (2015). Rising Morbidity and Mortality in Midlife among White

Non-Hispanic Americans in the 21st Century. *Proceedings of the National Academy of Sciences 112*(49), 15078–15083.

Case, A. and A. Deaton (2017). Mortality and Morbidity in the 21st Century. *Brookings Papers on Economic Activity, Spring 2017*.

Chetty, R., M. Stepner, S. Abraham, S. Lin, B. Scuderi, N. Turner, A. Bergeron, and D. Cutler (2016). The Association Between Income and Life Expectancy in the United States, 2001-2014. *JAMA 315*(16), 1750–1766.

Connor, I. (2007). *Refugees and Expellees in Post-War Germany.* Manchester University Press.

Deaton, A. (2003). Health, Inequality, and Economic Development. *Journal of Economic Literature 41*(1), 113–158.

Douglas, R. M. (2012). *Orderly and Humane: The Expulsion of the Germans after the Second World War.* New Haven: Yale University Press.

Eberhardt, P. (2011). Political migrations on polish territories (1939-1950). *Monografie-Polska Akademia Nauk. Instytut Geografii i Przestrzennego Zagospodarowania im. Stanisława Leszczyckiego*.

Falck, O., S. Heblich, and S. Link (2012). Forced Migration and the Effects of an Integration Policy in Post-WWII Germany. *The BE Journal of Economic Analysis & Policy 12*(1), 1–29.

Federal Statistical Office (2016). Statistik der Sterbefälle, Deutschland 1970 - 2014. Technical report, Statistisches Bundesamt, Wiesbaden.

Federal Statistical Office (2018). Labor Force Participation Statistics. Retrieved 7 June 2018, from `https://www-genesis.destatis.de/genesis/online/link/tabelleErgebnis/12211-0001`.

Fiala, N. (2015). Economic Consequences of Forced Displacement. *The Journal of Development Studies 51*(10), 1275–1293.

Forschungsdatenzentrum FDZ-RV (2017). Codeplan SUFDemografie Rentenbestand & Rentenwegfall 1993 – 2014. Technical report, Deutsche Rentenversicherung Bund, Berlin/Würzburg.

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

Kalwij, A. S., R. J. M. Alessie, and M. G. Knoef (2013). The Association Between Individual Income and Remaining Life Expectancy at the Age of 65 in the Netherlands. *Demography 50*(1), 181–206.

Kibele, E., S. Klüsener, and R. Scholz (2015). Drastic Changes in Regional Life Expectancy Disparities in Germany: in Search of the Determinants. Technical report, Max Planck Institute for Demographic Research.

Lubotsky, D. (2007). Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings. *Journal of Political Economy 115*(5), 820–867.

Lüttinger, P. (1986). Der Mythos der schnellen Integration: Eine empirische Untersuchung zur Integration der Vertriebenen und Flüchtlinge in der Bundesrepublik Deutschland bis 1971. *Zeitschrift für Soziologie 15*(1), 20–36.

Müller, W. and H. Simon (1959). *Die Vertriebenen in Westdeutschland. Ihre Eingliederung und ihr Einfluss auf Gesellschaft, Wirtschaft, Politik und Geistesleben*. Ferdinand Hirt.

Palme, M. and S. Sandgren (2008). Parental Income, Lifetime Income, and Mortality. *Journal of the European Economic Association 6*(4), 890–911.

Reichling, G. (1958). *Die Heimatvertriebenen im Spiegel der Statistik*. Duncker & Humblot.

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

Ruiz, I. and C. Vargas-Silva (2015). The Labor Market Impacts of Forced Migration. *The American Economic Review, Papers and Proceedings 105*(5), 581–586.

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

Saarela, J. M. and I. T. Elo (2016). Forced Migration in Childhood: Are There Long-Term Health Effects? *SSM - Population Health 2*, 813–823.

Sarvimäki, M., R. Uusitalo, and M. Jantti (2009). Long-Term Effects of Forced Migration. *IZA Discussion Paper 4003*.

Schumann, A. (2014). Persistence of Population Shocks: Evidence from the Occupation of West Germany after World War II. *American Economic Journal: Applied Economics 6*(3), 189–205.

Schwandt, H. and T. von Wachter (2017). Unlucky Cohorts: Earnings, Income, and Mortality Effects from Entering the Labor Market in a Recession. *Mimeo (April 2017)*.

Smith, J. P. (1999). Healthy Bodies and Thick Wallets: The Dual Relation between Health and Economic Status. *Journal of Economic Perspectives 13*(2), 145–166.

UNHCR (2016). Global Trends - Forced Displacement in 2015. Technical report, United Nations High Commissioner for Refugees.

van den Berg, G. J., M. Lindeboom, and F. Portrait (2006). Economic Conditions Early in Life and Individual Mortality. *The American Economic Review 96*(1), 290–302.

Tables

Table 1: Descriptive Statistics

	Men			
	Displaced	Non-Displaced	Difference	t-Stat.(p-value)
Age at Death	83.3	84.2	0.9	15.7(0.000)
Birth Year	1917.2	1917.1	-0.1	-1.9(0.000)
Earnings Points	45.9	44.2	-1.7	-10.5(0.000)
Pension Insurance Months	471.8	449.4	-22.4	-16.2(0.000)
Regional Birth Rate	0.027	0.024	-0.003	-66.0(0.000)
Regional Death Rate	0.019	0.017	-0.002	-46.7(0.000)
N	12 072	1 980 891		
Share Displaced (%)	0.6			

	Women			
	Displaced	Non-Displaced	Difference	t-Stat.(p-value)
Age at Death	84.8	86.4	1.6	37.4(0.000)
Birth Year	1918.5	1916.3	-2.2	-50.5(0.000)
Earnings Points	20.8	16.3	-4.5	-51.1(0.000)
Pension Insurance Months	303.7	262.5	-41.2	-41.6(0.000)
Regional Birth Rate	0.027	0.025	-0.003	-79.0(0.000)
Regional Death Rate	0.018	0.016	-0.002	-110.0(0.000)
N	21 031	2 964 372		
Share Displaced (%)	0.7			

Source: Own calculations based on pension shortfall records 1994 - 2013 (FDZ-RV–RTWF94-13DemoRWI).
Note: Reported values are means for the displaced and the non-displaced and their regarding differences.

Table 2: Baseline Estimates of the Mortality Hazard

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Displaced (HR)	1.15	1.16	1.24	1.27	1.04	1.05	1.10	1.11
	(1.13,1.17)	(1.14,1.18)	(1.22,1.26)	(1.24,1.29)	(1.02,1.05)	(1.03,1.06)	(1.08,1.11)	(1.09,1.12)
Earnings Points		-0.0001		-0.0001		-0.0001		-0.0001
		(0.0000)		(0.0000)		(0.0000)		(0.0000)
Death Rate			0.0060	0.0060			0.0128	0.0129
			(0.0007)	(0.0007)			(0.0006)	(0.0006)
Birth Rate			-0.0150	-0.0168			-0.0145	-0.0144
			(0.0007)	(0.0007)			(0.005)	(0.0006)
∅ Mortality Rate (%)	6.0				5.3			
N Displaced	12 072				21 031			
N Non-Displaced	1 980 891				2 964 372			

Source: Own calculations based on pension shortfall records 1994 - 2013 (FDZ-RV-RTWF94-13DemoRWI).
Note: Reported values on displacement indicator are hazard ratios (HR) from a discrete time Cox proportional hazards model (95% confidence intervals in parentheses). Reported values on continuous regressors (earnings, birth and death rates) are average marginal effects (standard errors in parentheses). All specifications include duration dummies for each year at risk (reference: first year at risk) and birth cohort dummies (reference: oldest birth category, including births from 1885 - 1900). Estimated coefficients are available from the authors upon request. Birth and death rates are measured by region of origin at the year of birth for each individual.

Table 3: Estimates of the Mortality Hazard across the Earnings Distribution

Specification	Men		Women	
	All Non-Displaced	Bavarian Non-Displaced	All Non-Displaced	Bavarian Non-Displaced
Full Earnings Distribution				
Displaced (HR)	1.27	1.25	1.11	1.10
	(1.24,1.29)	(1.22,1.27)	(1.09,1.12)	(1.08,1.12)
∅ Mortality Rate (%)	6.0	5.9	5.3	5.3
N Displaced	12 072	12 072	21 031	21 031
N Non-Displaced	1 980 891	334 369	2 964 372	531 188
Quintile 1				
Displaced (HR)	1.49	1.52	1.42	1.43
	(1.41,1.57)	(1.41,1.64)	(1.35,1.49)	(1.35,1.51)
∅ Mortality Rate (%)	5.9	5.9	5.2	5.2
N Displaced	1 422	909	1 641	1 907
N Non-Displaced	397 197	68 386	595 445	108 537
Quintile 2				
Displaced (HR)	1.60	1.57	1.28	1.26
	(1.54,1.66)	(1.48,1.66)	(1.24,1.33)	(1.21,1.31)
∅ Mortality Rate (%)	5.2	6.0	6.0	5.2
N Displaced	2 785	1 951	3 390	3 941
N Non-Displaced	395 837	67 341	593 775	106 507
Quintile 3				
Displaced (HR)	1.46	1.54	1.24	1.19
	(1.40,1.51)	(1.47,1.62)	(1.20,1.28)	(1.15,1.24)
∅ Mortality Rate (%)	6.1	6.1	5.3	5.2
N Displaced	3 052	2 791	4 135	4 167
N Non-Displaced	395 664	66 512	592 921	106 273
Quintile 4				
Displaced (HR)	1.15	1.21	1.09	1.09
	(1.11,1.20)	(1.16,1.27)	(1.06,1.12)	(1.06,1.13)
∅ Mortality Rate (%)	5.3	6.0	5.6	5.3
N Displaced	2 768	3 485	5 042	4 873
N Non-Displaced	395 756	65 776	592 035	105 572
Quintile 5				
Displaced (HR)	0.90	0.93	0.93	0.93
	(0.86,0.94)	(0.89,0.97)	(0.91,0.95)	(0.90,0.95)
∅ Mortality Rate (%)	6.2	5.5	5.4	5.3
N Displaced	2 045	2 936	6 823	6 143
N Non-Displaced	396 437	66 348	590 195	104 299

Source: Own calculations based on pension shortfall records 1994 - 2013 (FDZ-RV-RTWF94-13DemoRWI).
Note: Reported values on displacement indicator are hazard ratios (HR) from a discrete time Cox proportional hazards model (95% confidence intervals in parentheses). All specifications include duration dummies for each year at risk (reference: first year at risk), birth cohort dummies (reference: oldest birth category, including births from 1885 - 1900) and the full set of covariates as in columns (4) and (8) of Table 2. Estimated coefficients are available from the authors upon request.

Table 4: Sensitivity Checks on Birth Cohort Restriction and Contribution Period

	Men			
	Baseline	Cohort Restriction		Contribution ≥ 40 Yrs
		1927	1929	
	(1)	(2)	(3)	(4)
Without Covariates				
Displaced (HR)	1.15	1.16	1.16	1.13
	(1.13,1.17)	(1.14,1.18)	(1.15,1.18)	(1.11,1.16)
With Covariates				
Displaced (HR)	1.27	1.28	1.29	1.28
	(1.24,1.29)	(1.26,1.31)	(1.27,1.31)	(1.25,1.31)
∅ Mortality Rate (%)	6.0	6.3	6.6	6.0
N Displaced	12 072	13 993	17 055	8 067
N Non-Displaced	1 980 891	2 291 845	2 596 879	1 189 894

	Women			
	Baseline	Cohort Restriction		Contribution ≥ 40 Yrs
		1927	1929	
	(5)	(6)	(7)	(8)
Without Covariates				
Displaced (HR)	1.04	1.05	1.06	0.89
	(1.02,1.05)	(1.04,1.07)	(1.05,1.08)	(0.86,0.92)
With Covariates				
Displaced (HR)	1.11	1.12	1.13	0.94
	(1.09,1.12)	(1.11,1.14)	(1.12,1.15)	(0.90,0.97)
∅ Mortality Rate (%)	5.3	5.5	5.6	5.4
N Displaced	21 031	24 508	27 117	3 131
N Non-Displaced	2 964 372	3 248 183	3 475 929	290 478

Source: Own calculations based on pension shortfall records 1994 - 2013 (FDZ-RV-RTWF94-13DemoRWI). *Note:* Reported values on displacement indicator are hazard ratios (HR) from a discrete time Cox proportional hazards model (95% confidence intervals in parentheses). All specifications include duration dummies for each year at risk (reference: first year at risk), birth cohort dummies (reference: oldest birth category, including births from 1885 - 1900) and the full set of covariates as in columns (4) and (8) of Table 2. Estimated coefficients are available from the authors upon request.

Figures

Figure 1: Regions of Origin of Displaced Individuals

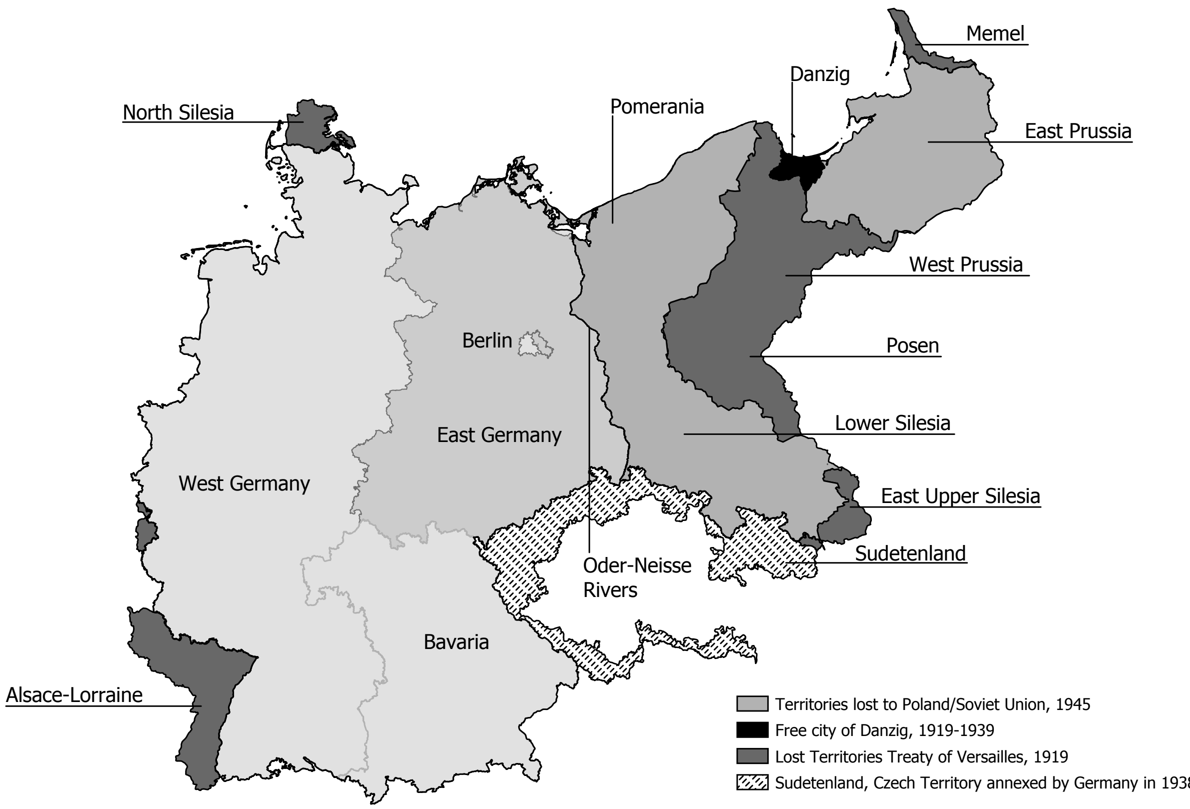
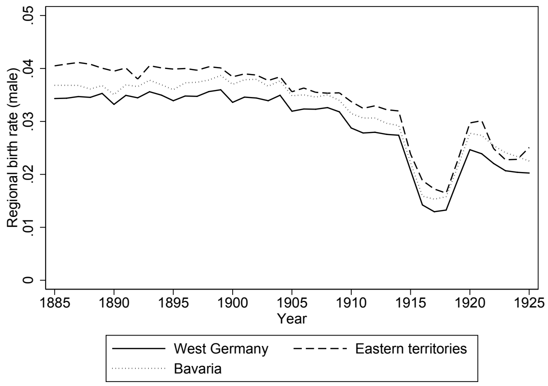
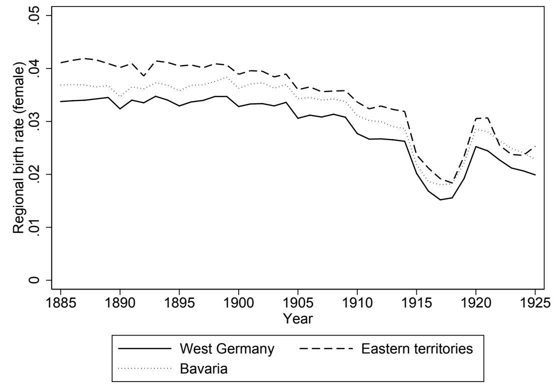


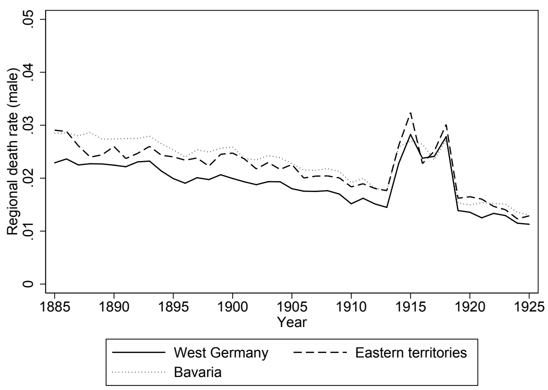
Figure 2: Regional Birth and Death Rates by Calendar Year



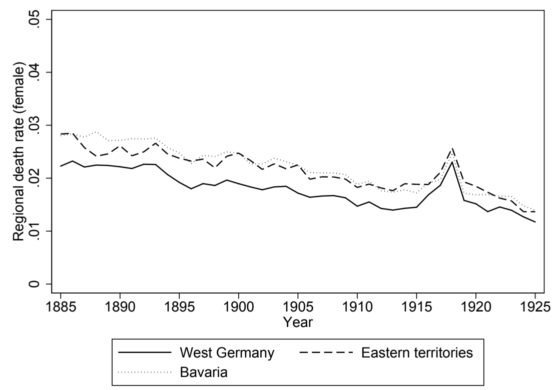
(a) Birth rate - Men



(b) Birth rate - Women



(c) Death rate - Men

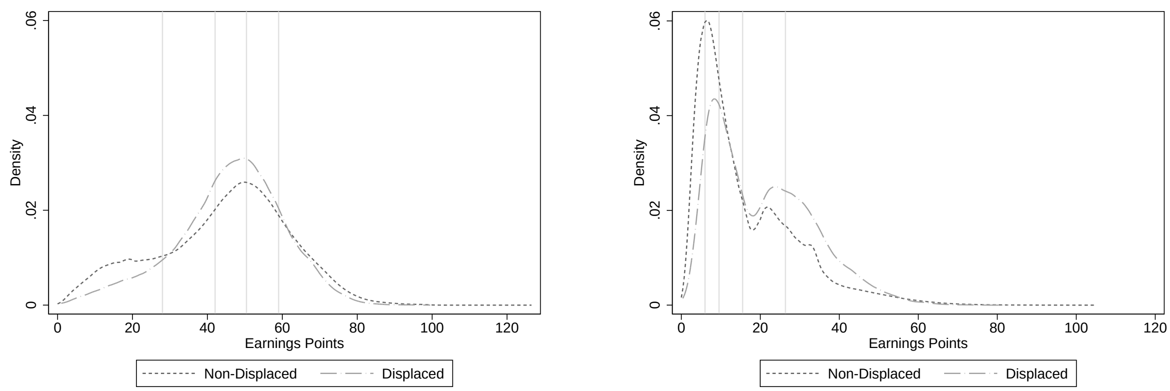


(d) Death rate - Women

Source: Own calculations based on Besser (2008).

Note: Annual birth rates are calculated as the number of births relative to the number of residents within a region for each year. Annual death rates are calculated as the number of deaths relative to the number of residents within a region for each year.

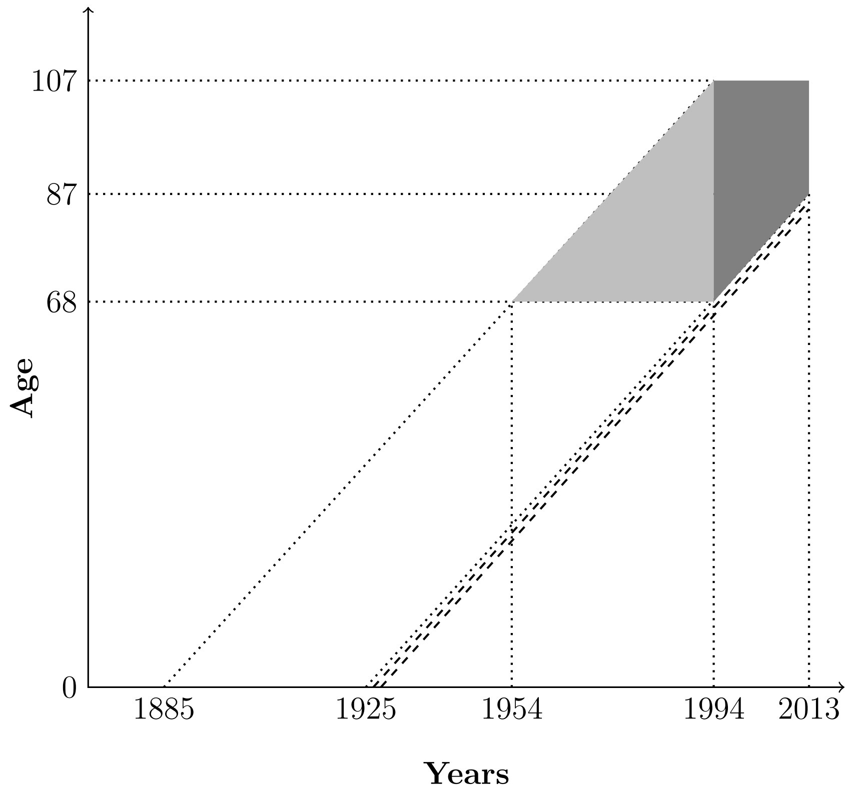
Figure 3: Pre-Retirement Lifetime Earnings Distribution



(a) Men (b) Women

Source: Own calculations based on pension shortfall records 1994 - 2013 (FDZ-RV–RTWF94-13DemoRWI).
Note: To approximate pre-retirement lifetime earnings, we use the total sum of earnings points that document pension claims. This variable predominantly includes labor earnings which are the primary source of public pension claims. In addition, the measure also includes earnings points that are acquired from indirect labor market related activities that are creditable for monthly pension benefits (e.g. periods of unemployment, education, child raising and creditable periods of illness). Vertical lines indicate the quintile cutoffs.

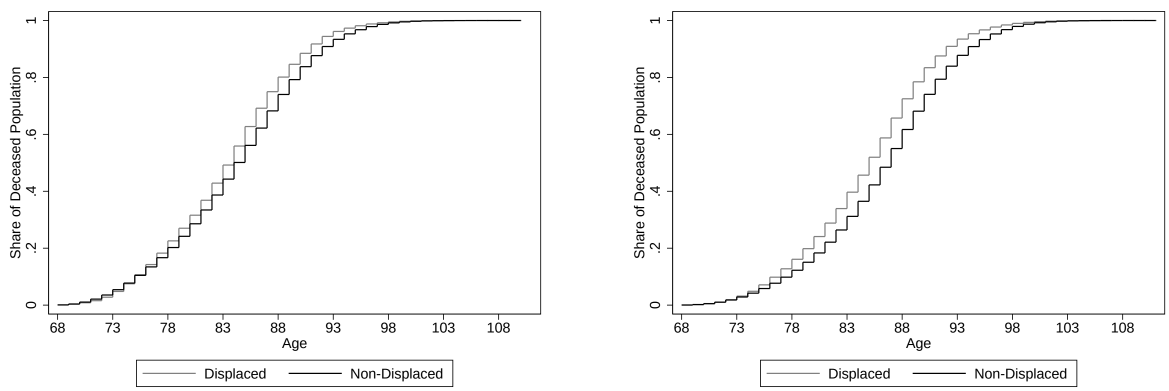
Figure 4: Lexis Diagram of the Observational Plan



Source: Own graph.

Note: The gray area spans the space of sampled individuals over the observation period (1994 - 2013). Individuals who are still alive at the first observed date (i.e. 1994) but are above age 68 are followed back to age 68 (light-gray area). The corresponding birth cohorts are from 1885 (no younger than 68 at the beginning of 1954) to 1925 (no younger than 68 at the beginning of 1994). The two dashed lines indicate the inclusion of additional younger cohorts (1927 and 1929, respectively) within the sensitivity analysis.

Figure 5: Cumulative Distribution of the Deceased Population



(a) Men

(b) Women

Source: Own calculations based on pension shortfall records 1994 - 2013 (FDZ-RV–RTWF94-13DemoRWI).

Appendix Not Intended for Publication

Table 5: Multi-Level Model Results (Dependent Variable: Mortality)

	Men		Women	
	BC	BC Omitted	BC	BC Omitted
Random Effects Parameters:				
Baseline				
s.d.(Fixed Effect)	0.0063	0.0073	0.00000000211	0.0058
	(0.0041)	(0.0045)	(0.000000297)	(0.0035)
s.d.(Residual)	0.2244	0.2269	0.2113	0.2133
	(0.0003)	(0.0003)	(0.0002)	(0.0002)
100*[s.d. FE / (s.d. FE + s.d. Resid.)]	2.7%	3.1%	0.0000001%	2.6%
Random Effects Parameters:				
Baseline + Covariates				
s.d.(Fixed Effect)	0.0094	0.0189	0.00000000584	0.0159
	(0.0054)	(0.0097)	(0.0000000929)	(0.0081)
s.d.(Residual)	0.2244	0.2250	0.2113	0.2120
	(0.0003)	(0.0003)	(0.0002)	(0.0002)
100*[s.d. FE / (s.d. FE + s.d. Resid.)]	4.0%	7.7%	0.000003%	7.0%
Observations	398592		597080	
Observations per group minimum	2358		3854	

Source: Own calculations based on pension shortfall records 1994 - 2013 (FDZ-RV–RTWF94-13DemoRWI). *Note:* Reported values are estimates of the standard deviation (s.d.) from a multilevel mixed-effects linear regression, with regions at the higher level and individuals at the lower level (standard errors in parentheses). Average deviation of mortality explained at the regional level is noted as the standard deviation from the fixed effect (s.d.(Fixed Effect)) and the remaining variation is at the individual level (s.d.(Residual)). All specifications are estimated including birth cohort dummies (BC) and omitting the birth cohort dummies (BC Omitted). To obtain model convergence, the estimates are based on a randomly drawn 20% sub-sample.

Table 6: Baseline Estimates with Interactions Between Displacement and Age (Duration)

	Men		Women	
	Hazard Ratio (s.e.)	p-value (95% CI)	Hazard Ratio (s.e.)	p-value (95% CI)
Displaced	.79 (.20)	0.388 (.48,1.32)	1.18 (.24)	0.411 (.78,1.79)
age69	2.86 (.05)	0.000 (2.75,2.98)	2.84 (.07)	0.000 (2.70,2.98)
age70	5.02 (.09)	0.000 (4.83,5.22)	5.18 (.12)	0.000 (4.95,5.42)
age71	7.71 (.14)	0.000 (7.43,8.00)	7.98 (.17)	0.000 (7.63,8.34)
age72	10.57 (.19)	0.000 (10.19,10.96)	11.68 (.25)	0.000 (11.18,12.20)
age73	14.01 (.25)	0.000 (13.51,14.52)	15.81 (.34)	0.000 (15.14,16.50)
age74	17.77 (.32)	0.000 (17.13,18.42)	20.74 (.45)	0.000 (19.87,21.64)
age75	21.13 (.38)	0.000 (20.38,21.90)	24.82 (.53)	0.000 (23.79,25.90)
age76	23.98 (.43)	0.000 (23.14,24.86)	29.56 (.63)	0.000 (28.33,30.84)
age77	28.32 (.51)	0.000 (27.33,29.36)	35.19 (.75)	0.000 (33.74,36.71)
age78	32.96 (.60)	0.000 (31.80,34.16)	42.29 (.91)	0.000 (40.54,44.11)
age79	39.97 (.72)	0.000 (38.57,41.42)	51.90 (1.11)	0.000 (49.76,54.13)
age80	48.03 (.87)	0.000 (46.35,49.77)	63.88 (1.37)	0.000 (61.25,66.63)
age81	58.06 (1.05)	0.000 (56.04,60.16)	79.30 (1.69)	0.000 (76.04,82.70)
age82	69.46 (1.25)	0.000 (67.04,71.98)	95.82 (2.04)	0.000 (91.88,99.92)
age83	83.40 (1.51)	0.000 (80.50,86.42)	118.62 (2.53)	0.000 (113.76,123.70)
age84	100.08 (1.81)	0.000 (96.59,103.69)	146.01 (3.11)	0.000 (140.02,152.25)
age85	121.13 (2.19)	0.000 (116.91,125.50)	182.28 (3.89)	0.000 (174.81,190.07)
age86	149.33 (2.70)	0.000 (144.12,154.72)	230.30 (4.91)	0.000 (220.87,240.14)
age87	187.38 (3.39)	0.000 (180.85,194.15)	298.87 (6.37)	0.000 (286.63,311.63)
age88	231.15 (4.18)	0.000 (223.08,239.51)	381.81 (8.14)	0.000 (366.17,398.12)
age89	278.78 (5.06)	0.000 (269.03,288.88)	471.68 (10.07)	0.000 (452.35,491.84)
age90	334.70 (6.09)	0.000 (322.96,346.86)	581.73 (12.43)	0.000 (557.87,606.61)
age91	406.32 (7.42)	0.000 (392.02,421.15)	712.00 (15.23)	0.000 (682.77,742.49)
age92	479.45 (8.81)	0.000 (462.49,497.04)	846.78 (18.14)	0.000 (811.96,883.10)
age93	554.21 (10.26)	0.000 (534.45,574.71)	978.11 (21.00)	0.000 (937.80,1020.15)
age94	626.52 (11.73)	0.000 (603.93,649.95)	1118.75 (24.08)	0.000 (1072.53,1166.97)
age95	734.35 (13.94)	0.000 (707.52,762.20)	1294.56 (27.96)	0.000 (1240.89,1350.54)
age96	854.72 (16.56)	0.000 (822.85,887.82)	1513.09 (32.83)	0.000 (1450.09,1578.84)
age97	982.41 (19.67)	0.000 (944.60,1021.74)	1797.35 (39.26)	0.000 (1722.02,1875.97)
age98	1169.82 (24.55)	0.000 (1122.67,1218.96)	2104.00 (46.47)	0.000 (2014.86,2197.08)
age99	1316.41 (29.95)	0.000 (1258.99,1376.45)	2409.11 (54.20)	0.000 (2305.18,2517.71)
age100	1608.88 (40.62)	0.000 (1531.20,1690.5)	2768.72 (64.13)	0.000 (2645.82,2897.32)
age101	1868.88 (55.71)	0.000 (1762.82,1981.33)	3221.64 (78.14)	0.000 (3072.06,3378.50)
age102	2079.98 (78.01)	0.000 (1932.56,2238.65)	3673.95 (96.19)	0.000 (3490.16,3867.42)
age103	2541.81 (122.93)	0.000 (2311.93,2794.54)	4349.30 (127.46)	0.000 (4106.50,4606.45)
age104	2647.75 (181.26)	0.000 (2315.28,3027.97)	4876.12 (171.09)	0.000 (4552.05,5223.27)
age105	3601.26 (336.53)	0.000 (2998.54,4325.13)	5853.78 (257.91)	0.000 (5369.49,6381.74)
age106	3626.61 (543.69)	0.000 (2703.27,4865.32)	6572.09 (391.85)	0.000 (5847.25,7386.79)
age107	3752.86 (848.70)	0.000 (2409.15,5846.04)	6958.25 (595.38)	0.000 (5883.92,8228.74)
diplaceXage69	1.12 (.33)	0.694 (.62,2.00)	.76 (.19)	0.297 (.46,1.26)
diplaceXage70	1.02 (.29)	0.931 (.58,1.78)	.71 (.17)	0.164 (.45,1.14)
diplaceXage71	.88 (.24)	0.654 (.51,1.52)	.72 (.16)	0.153 (.46,1.12)
diplaceXage72	1.29 (.34)	0.332 (.76,2.19)	.85 (.18)	0.480 (.55,1.31)
diplaceXage73	1.53 (.40)	0.106 (.91,2.58)	.91 (.19)	0.682 (.59,1.40)
diplaceXage74	1.53 (.40)	0.104 (.91,2.57)	.90 (.19)	0.649 (.59,1.38)
diplaceXage75	1.54 (.40)	0.098 (.92,2.59)	1.09 (.23)	0.672 (.71,1.66)
diplaceXage76	1.65 (.43)	0.055 (.98,2.77)	1.00 (.21)	0.988 (.65,1.52)

Continued on next page

Table 6 – *Continued from previous page*

	Men		Women	
	Hazard Ratio (s.e.)	p-value (95% CI)	Hazard Ratio (s.e.)	p-value (95% CI)
diplaceXage77	1.64 (.43)	0.059 (.98,2.74)	1.00 (.21)	0.983 (.66,1.52)
diplaceXage78	1.61 (.42)	0.067 (.96,2.70)	.99 (.21)	0.968 (.65,1.50)
diplaceXage79	1.44 (.37)	0.164 (.86,2.41)	.95 (.20)	0.818 (.62,1.44)
diplaceXage80	1.38 (.36)	0.213 (.82,2.31)	.97 (.20)	0.895 (.64,1.47)
diplaceXage81	1.54 (.40)	0.094 (.92,2.58)	.89 (.18)	0.592 (.58,1.35)
diplaceXage82	1.57 (.41)	0.083 (.94,2.62)	.89 (.18)	0.600 (.59,1.35)
diplaceXage83	1.54 (.40)	0.095 (.92,2.58)	.92 (.19)	0.709 (.61,1.39)
diplaceXage84	1.61 (.42)	0.066 (.96,2.69)	.79 (.16)	0.286 (.52,1.20)
diplaceXage85	1.58 (.41)	0.077 (.95,2.64)	.86 (.18)	0.507 (.57,1.31)
diplaceXage86	1.43 (.37)	0.166 (.86,2.39)	.83 (.17)	0.407 (.55,1.26)
diplaceXage87	1.33 (.34)	0.269 (.79,2.23)	.87 (.18)	0.531 (.57,1.32)
diplaceXage88	1.22 (.32)	0.441 (.73,2.04)	.83 (.17)	0.409 (.55,1.27)
diplaceXage89	1.26 (.33)	0.363 (.75,2.12)	.83 (.17)	0.401 (.55,1.26)
diplaceXage90	1.21 (.32)	0.455 (.72,2.04)	.78 (.16)	0.242 (.51,1.18)
diplaceXage91	1.40 (.37)	0.195 (.83,2.36)	.84 (.17)	0.426 (.55,1.28)
diplaceXage92	1.29 (.34)	0.325 (.77,2.19)	.81 (.17)	0.332 (.53,1.23)
diplaceXage93	1.16 (.31)	0.566 (.68,1.98)	.86 (.18)	0.483 (.56,1.31)
diplaceXage94	1.11 (.30)	0.683 (.65,1.91)	.79 (.17)	0.278 (.51,1.20)
diplaceXage95	1.14 (.32)	0.621 (.66,1.98)	.80 (.17)	0.332 (.52,1.24)
diplaceXage96	1.08 (.31)	0.775 (.61,1.91)	.73 (.16)	0.170 (.47,1.14)
diplaceXage97	1.06 (.32)	0.824 (.59,1.93)	.80 (.18)	0.326 (.51,1.24)
diplaceXage98	1.16 (.36)	0.638 (.62,2.15)	.70 (.16)	0.139 (.44,1.11)
diplaceXage99	1.19 (.41)	0.598 (.60,2.36)	.78 (.19)	0.318 (.48,1.26)
diplaceXage100	1.03 (.41)	0.936 (.46,2.27)	.74 (.19)	0.266 (.44,1.25)
diplaceXage101	.91 (.44)	0.853 (.34,2.38)	1.01 (.27)	0.971 (.58,1.73)
diplaceXage102	.77 (.50)	0.699 (.22,2.75)	.56 (.21)	0.125 (.27,1.17)
diplaceXage103	.87 (.67)	0.867 (.19,3.98)	.61 (.25)	0.240 (.27,1.38)
diplaceXage104	–		1.07 (.45)	0.858 (.47,2.46)
diplaceXage105	–		.18 (.19)	0.104 (.02,1.40)
diplaceXage106	–		.96 (.62)	0.957 (.27,3.45)
diplaceXage107	–		–	
Intercept	.0001586 (.00)	0.000 (.00,.00)	.00 (.00)	0.000 (.00,.00)
Observations	1980852		2964355	

Source: Own calculations based on pension shortfall records 1994 - 2013 (FDZ-RV–RTWF94-13DemoRWI).

Note: Reported values on displacement indicator are hazard ratios (HR) from a discrete time Cox proportional hazards model (95% confidence intervals in parentheses), separately for men and women. Both specifications include age/duration dummies for each age-year at risk (reference: first year at risk) and corresponding interactions for each age indicator with the displacement indicator. Note that, due to small numbers of observations at very high ages (above age 100), some of the interaction terms between the age and the displacement indicator cannot be estimated. This also explains why the sample sizes are slightly smaller for both men and women as compared to the baseline specification that are similar but do not include the interaction terms (table 2, colums (1) and (5)). The specifications further include birth cohort dummies (reference: oldest birth category, including births from 1885 - 1900) but these are not reported for brevity (available from the authors upon request).