Post-war migration flows and disparities in mortality from age 50 on: the case of Turin in Italy

Virginia Zarulli

key words:
internal migration, mortality inequality

Abstract

Compositional changes due to internal migration can modify the distribution of health outcomes, death rates and socioeconomic characteristics of a specific geographical area. Migration flows may affect patterns of socioeconomic inequalities in mortality as well. However, despite these inequalities are an important social and geopolitical feature of an area, there is still little empirical evidence on this effect. This paper contributes to deepen the knowledge about this phenomenon by investigating whether post-war internal migration in Italy affected the pattern of mortality inequality by socioeconomic status, from age 50 onwards, in Turin, one of the main industrial areas of the country, where many low educated individuals from the southern regions migrated to Turin with seeking jobs in the car factories. Migrants might be selected in terms of robustness because of the healthy migrant effect. However, lowly educated individuals are employed in heavier and riskier jobs. They thus undergo a faster health selection due to exposure to higher mortality risk that selects the most robust individuals. This paper hypothesized that the interplay of these mechanisms might have produced a homogenization process towards robustness of the population, by reducing the unobserved heterogeneity in survival chances, and that these processes affected men more than women, because women were likely to be more passive actors in the migratory decisions and less heavily involved in the industrialization process. The results show that women have higher level of heterogeneity in susceptibility to death and wider differentials mortality by education level than men, which both support the hypotheses.

1 Introduction

Migration is, together with fertility and mortality, one of the three major contributors to the demographic dynamics of a geographical area. Compared to the native population, migrants can have different reproductive behaviors (Mussino and Van Raalte 2013) especially shortly after their arrival (Andersson 2004; Mussino and Strozza 2012) or different mortality and morbidity profiles (Feinleib et al. 1982; Bennett 1993; McCredie et al. 1999) due to the influence of several factors like lifestyle, dietary habits, physical environment and hereditary predispositions of the population of origin.

Several studies have documented the so called “healthy migrant” effect, the phenomenon of migrants showing better health status than persons who are not migrants (Sharma et al. 1990; Singh and Siahpush 2001; Valkonen et al. 1992; Feinleib et al. 1982; Kington et al. 1998; Deboosere and Gadeyne 2005), as well as the opposite phenomenon, known as “salmon bias” (Palloni and Arias 2004), when weaker and sick migrants move back to their place of origin for receiving care from family, friends and institutions (Brimblecombe et al. 2000; Lanska and Peterson 1995).

Migratory movements affect the structure of health and mortality outcomes of both the areas of origin and destination. Compositional changes of the population due to migration can modify the distribution of health outcomes, cause specific or all cause death rates, life style factors and socioeconomic characteristics. The vast majority of studies have focused on long distance and international migrations while fewer analyses have investigated shorter distance migratory movements, such as within country or between adjacent countries migration. Among these studies, some have analyzed health inequalities between immigrants and natives (Wild and McKeigue 1997; Kolčić and Polašek 2009) or between migrants and non-migrant population in the region of origin (Westman et al. 2008). Others have focussed on the effect of migration on the geographical variation in health and mortality, investigating how internal migration affected regional differences in mortality (Brimblecombe et al. 2000, 1999; Luy and Caselli 2007), variation in area deprivation levels (Norman et al. 2005) and health neighborhood inequalities (van Lenthe et al. 2007). Finally, other studies have analyzed the impact of migration on all-cause mortality of the area of destination (Rasulo et al. 2012).

By contrast there is still little empirical evidence on the effect of internal migration flows on the patterns of socioeconomic inequality in mortality. However, as “the migration dimension cannot be understood independently of social class and gender" (Malmusi et al. 2010: 3), the reciprocal relation holds as well and socioeconomic differences in health and mortality need to be considered also in the light of migration. The study of soecioeconomic differences in health and mortality is an important publich health issue with crucial policy making implications. An extensive literature has showed that in many countries, despite medical improvement and the creation of welfare states, socioeconomic inequalities in health and mortality have been steadily widening in the last decades (Marmot 1986; Valkonen 1993; Hattersley 1997; Valkonen 2001; Kunst 2004; Mackenbach 2006; Strand 2010; Shkolnikov 2011; Zarulli et al. 2012). Analysing the trends and understanding the mechanisms behind this phenomenon is necessary to tackle it.

In the second half of the 20^{th} century, massive internal migration occurred from the poorer Southern regions of Italy to the Northern ones, which became center of the newly established industrial economy after WWII. These migratory movements accompanied the deepest social and economic transformation experienced in Italy (Ginsborg 1989). Migration reached its highest peak in the 60's, continued substantially also in the first half of the 70's and then gradually declined (Bonifazi and Heins 2000). Turin was among the major points of attraction, together with Milan and Genoa. The three Northwestern cities were also called the "industrial triangle". From 1955 to 1973 the in-migration rate to Turin was several times higher than the out-migration rate (Bonifazi and Heins 2000). Turin's economy was characterized by the car industry, which was growing quickly after WWII and mainly attracted lowly educated labor force seeking jobs in the factories.

The aim of this paper is to investigate whether the internal migration flows that characterized Italy in the second half of the 20th century affected the pattern of mortality inequality by socioeconomic status from age 50 on in the north-western industrial city of Turin.

With this respect, two possible mechanisms seem particularly relevant.

First, as several studies have documented, migrants might be selected in terms of robustness because of the healthy migrant effect (Sharma et al. 1990; Singh and Siahpush 2001; Valkonen et al. 1992; Feinleib et al. 1982; Kington et al. 1998; Deboosere and Gadeyne 2005). Therefore, a substantial number of individuals who were potentially healthier than the average population might have reached Turin, with possible consequences on the health and mortality profiles of the population of the city, especially among men. In the past women were more likely to be only passive actors in the migratory decision (Mincer 1978; Bielby and Bielby 1992; Cooke 2008) and it can be hypotesized that, if a migration related health selection occurred, men were more strongly affected than women.

Second, migrants to Turin during the post-war migration were mainly lowly educated and employed in less skilled, heavier and riskier jobs of the car factories of the city (Canteri 1964; Fofi 1964; Pellicciari and Albertelli 1970; D'onofrio 2010). Therefore, they were likely to be exposed to higher health and mortality hazard. This might have caused a faster health selection of the most robust individuals. According to the literature on unobserved heterogeneity of frailty, the higher the pressure of mortality, the stronger and faster the selection towards robustenss undergone by the surviving population, because the frailest individuals die earlier (Vaupel et al. 1979; Manton et al. 1981). As the migrants were characterized by a low level of education, such educational group might have benefited from a decreased average frailty, thus being left with a more robust population and, therefore, reducing the mortality differentials respect to other educational groups. Also in this case it is possible to imagine that this mechanism acted less strongly on women, as they were not as involved as men in the industrialization process.

The paper hypothesizes that the interplay of these two mechanisms might have produced a homogenization process with respect susceptibility to death (otherwise called unobserved frailty), especially of the male population. A decrease in the average frailty implies a decrease in the variance of individual differences in susceptibility to death, hence, a more homogeneous population.

This might have also had an impact on the differences in mortality risk by socioeconomic level. An inflow of lowly educated migrants, who are potentially healthier than non migrants, might have lowered non only the average relative level of frailty of the general population, but also of the group with low education compared to other educational groups. The fact that most of them were employed in heavier and riskier jobs and, therefore, were exposed to a higher mortality hazard and to its subsequent health selection, might have strengthened this process and reduced even further the health and mortality differentials by education.

Because most of the work related migrations happen at young working ages, if any selection due to exposure to heavier and riskier jobs occurs, it needs a certain time lapse to take place. Therefore, age 50 was selected as initial age of the mortality follow up.

2 Data and methods

2.1 Data

We used individual longitudinal data from the Turin Longitudinal Study. Individuals who were registered during at least one of the four censuses in 1971, 1981, 1991 and 2001 were selected. Their migration and vital status was followed up until the end of July 2007. The result is an observation window of 36 years (from October 24^{th} 1971, official date of the census in 1971, to the end of July 2007, end of the record linkage) during which the individuals were followed up until death, emigration from the city or end of observation period. Individuals started to be followed up from age 50. The study population contains 391,170 men and 456,216 women.

The available information were date of birth, date of exit from the study, cause of exit (death or emigration), sex, region of birth and education level, which, accordingly to the literature, was used as a proxy for the socioeconomic status (Mirowsky and Ross 2003; Doblhammer et al. 2009; Krieger et al. 1997). To facilitate the comparison between mortality levels of different education groups over a long follow up period and over the many different cohorts included in the study, we created three broad categories: high (high school or higher), medium (intermediate school (corresponding to 6^{th}, 7^{th} and 8^{th} grades)) and low (elementary school or lower).

The focus on people older than 50 makes this educational classification reasonable. The basic division into three levels of the Italian school system was set up in the 1920's[1]: a 5-year elementary school was followed by two types of 3-year intermediate school: one vocational type, which prepared the student for the labor market and did not allow accessing any higher levels of education, and one type that allowed accessing the technical high schools or the preparatory high schools for the university. The republican Constitution of 1948 kept the same system but introduced the eight schooling years obligation [2]. However, this obligation was not really implemented until the important reform which occurred in the 1960's[3]. The reform created a single compuslory 3-year intermediate level that did not preclude the access to the various high schools. Until that time, *de facto*, only elementary

[1] A series of royal legislative decrees between 1922 and 1923 (31 December 1922, n. 1679, 16 July 1923, n. 1753, 6 May 1923, n. 1054, 30 September 1923, n. 2102 and 1 October 1923, n. 2185) represent the so called Gentile Reform, named after the philosopher and minister of education Giovanni Gentile.

[2] Constitution of the Italian Republic, article 34.

[3] Law 31 December 1962 n.1859.

school was compulsory and even in that case many children, especially in the southern regions, still did not attend all the 5 years[4]. Only a small part of the population went on studying extra 3 years and individuals who attended high school then, especially women, represented an even more selected population who, almost certainly, would have continued their schooling career attending university, which is nowadays considered as a “high” level of education.

Data quality is very high. The share of linked records between census and population registry was higher than 96% (except for the 1971 when it was 84.7%). The unmatched census records are evenly distributed across the main socio-demographic categories, thus excluding bias probelms related to a selected population. Information about education was unknown or unavailable in only 0.4% of the records for both men and women.

2.2 Methods

Given the longitudinal and individual level nature of the data, the statistical technique of survival analysis regression for duration dependence models was used[5], adjusting for both right censoring and left truncation. Both type of censoring are present in the data, the first one due to emigration of the individuals before the end of the observation period or end of this period without having experience the event death; the second one due to the different ages at entry in the observation window of the individuals.

Before estimating the mortality differentials by education level and macro-region

[4]Despite the national and centrally managed school system, strong geographical differences in access to education and literacy rates have characterized Italy for long time, fading away only in the recent decades (Felice 2007).

[5]In this kind of analyses the duration of the process under scrutiny (time until death) together with the occurence of the event (death) represent the dependent variable. This allows controlling for the dynamic change over time of the population at risk, by taking into account, at any time unit of the observation period, the number of events already occurred and at which ages they have occurred, as well as the number of individuals that, instead, are still at risk of experiencing the event. This implies also taking into account the age structure of the surviving population over time. The aim is generally to model the age profile of the mortality risk of the standard individual (otherwise called baseline hazard), who is the individual with characteristics corresponding to the reference categories of the covariates of interest, and to estimate the effect of those covariates on the baseline hazard. It is common to report the exponentiated coefficients, which represent mortality rate ratios respect to the reference category, that is by how much the covariate increases or decreases the risk of death at each age. For the interested reader, a brief but complete and detailed introduction to survival analysis can be found in (Wienke 2010a).

of birth, a preliminary analysis for identyfing the best functional baseline mortality hazard was performed. The baseline was modeled parametrically with the Gompertz (Gompertz 1825) and the Makeham functions, known to describe adequately patterns of adult human mortality, and compared via AIC (Akaike Information Criterion) (Akaike 1974).

The results indicated that Gompertz was the best fit for the men and Makeham best for the women. Therefore, the rest of the analysis was carried on adopting Gompertz baseline for the men and Makeham for the women.

Neglecting unobserved sources of heterogeneity (often called unobserved frailty) in survival analysis models might lead to biased estimates of the hazard and of the coefficients of the covariates (Gail et al. 1984; Trussell and Rodriguez 1990; Rodriguez 1994; Chamberlain 1985; Aalen 1994; Aalen et al. 2008). Frailty models (Vaupel et al. 1979; Manton et al. 1981) allow controlling for this component. As individuals differ for many unobserved and unobservabel characteristics that influence their risk of death, we applied this approach and included a frailty term into the models. Frailty is a random effect term, assumed to follow a gamma distribution with mean 1 and variance σ^2 to be estimated, which indicates the level of heterogeneity of the population with respect to susceptibility to death (Wienke 2010b).

A first explorative analysis of the data revealed that from 1971 to 2007 a significant mortality improvement occurred, as it is shown in the mortality surface in figure 1. The different cohorts included in the study passed through the years of observation at different ages and benefit from the mortality improvements occurring during those years at different ages.

Figure 1 here.

To take this important phenomenon into account we included covariates for the calendar periods and split the individual survival experience into several spells, one for each period the individual passed through. This implied an organization of the data into clusters, where each cluster represents one individual that shares the same hidden frailty.

In these cases, to control for unobserved frailty, shared frailty models (Wienke 2010c) needed to be applied. Equation 1 describes the model:

$$\mu(x|u_{ij}, z_i) = z_i \mu_0(x) e^{u_{ij}\beta} \tag{1}$$

where u_{ij} is the covariate profile of the j-th observation in the i-th cluster, z_i is the hidden frailty shared by the i-th cluster and $\mu(0)$ is the baseline hazard.

Shared frailty models, especially with left truncated data, are very complex models (for the likelihood function please see (Van den Berg and Drepper 2011)) and, when applied to a large dataset, they can be computationally challenging[6]. For these reasons it was impossible to estimate this model with a calendar period classification finer than two periods, 1971-1990 and 1991-2007, and the optimization routine needed to be applied via random subsampling. We estimated the models repeatedly over 1% sample of the data, randomly drawn without replacement (Hartigan 1969, 1975; Politis and Romano 1994), stratified by the major characteristics (education and region of birth). As pointed out by Efron (1979: 24), random subsampling "is very similar to bootstrap". The basic idea is to approximate the parameter estimates based on the distribution of the repeated estimates. Given the complexity of the model and the large size of the data set, 250 estimate repetitions were performed[7]. It is generally advised to choose a sufficiently large number of repetitions (Hesterberg 2011), but in the application of very complex estimators, several studies have shown that between 100 and 500 can provide relatively small error margins (Efron and Tibshirani 1993; Manly 1997; Pattengale et al. 2010).

3 Results

3.1 Descriptive results

Table 1 shows the distribution of the population by education and macro-region of birth. The majority of individuals (especially women) have low education level. The share of medium education is similar for both sexes while the percentage of highly educated individuals among men is higher than among women.

As expected, the share of individuals born in the "North-West" region is the largest, as this is where Turin is located. We will consider these individuals as natives[8]. Among the individuals born in the other areas, 33% of men and 28% of

[6] Initially, the time from 1971 to 2007 had been divided into 12 periods of 3 years each. This means that the individual had up to maximum of 12 spells and the final size of the split dataset was several million lines. For computational reasons it was impossible to estimate this model so the number of calendar periods had to be reduced to two, 1971-1990 and 1991-2007, and the optimization routine needed to be applied via random subsampling.

[7] By ways of illustration, one repetition for the Gamma-Gompertz model with education and macro-region of birth covariates, applied to the 1% sample drawn from the men dataset, required between 12 and 13 hours on a 4CPUs - 4 GB computer.

[8] This is not completely correct because those born in the "North-West" region include also the short distance migrants from the areas around Turin. However, this does not affect our analysis

women were born in "South & Islands" region, around 10% for both men and women come from the "North-East" region and only small numbers from the "Center" and "Abroad" regions.

	Men	(%)	Women	(%)
Education				
Low	218,098	(55.75)	299,400	(65.62)
Medium	95972	(24.53)	100,926	(22.12)
High	77100	(19.71)	55890	(12.25)
Tot	391,170	(100)	456,216	(100)
Macro-Region of birth				
North-West	192,660	(49.25)	244,613	(53.61)
North-East	39,189	(10.02)	49,680	(10.89)
Center	13,198	(3.37)	19,900	(3.05)
South & Islands	128,461	(32.84)	126,041	(27.63)
Abroad	17,662	(4.51)	21,982	(4.82)

Table 1: Distribution of the Turin population included in the follow-up by education and by macro-region of birth.

The data confirm that migrants from the South were often individuals with low education level. Figure 2 **??** shows that among men with low education, slightly less than 40% were natives and slightly more than 40% were born in the South, while for the other educational groups the share of individuals from the South is smaller than the share of natives. The higher the level of education the lower the share of individuals born in the South: almost 70% of the highly educated men were born in the "North-West" area (therefore they are natives) and only around 20% was born in the "South & Islands" area; almost 60 % of the individuals with medium education comes from "North-West" and only slightly more than 25 % from "South & Islands". Women present a similar distribution.

Figure 2 here.

Interestingly, these distributions change across different cohorts, as shown by the Figures 3, 4 and 5.

The share of the population with low education was the largest among the oldest cohorts (between 70% and 80%) and it gradually decreased from older to younger

because we are focused on the South-North migration flow.

cohorts. Meanwhile, the proportion of individuals in the medium and high education categories increased. This is consistent with the gradual extension of access to high education that took place worldwide over the twentieth century, especially in economically developed countries, and sharply accelerated after 1960 (Schofer and Meyer 2005). In Italy and more specifically in the North-Western region, the rate of high school and university enrollment increased from 6.51 % in 1951 to 64.26 % in 2001 (Felice 2007: 385). This expansion process was facilitated by the economic growth and the education system reform in 1969[9] that liberalized the access to a mainly publicly funded university, which was made possible for anyone with a 5-year high school degree and not only for those coming from the preparatory high school.

Among younger cohorts it is also possible to see that the shares of medium and high education are similarly distributed among men and women, whereas women belonging to previous cohorts had always been the minority among the highly educated group. Concerning those in the low education group, the gender gap still persists. Although the difference tends to lessen, even in the youngest cohorts the proportion of women with low education is bigger than the proportion of men.

Figure 3 here.

The Italian history of post war migration flows can be read in Figure 4. The share of migrants from the Southern regions constantly grew from older to younger cohorts. It is particularly high among the immediate postwar generations, reflecting the fact that young adults with many small children (those were the years of the post war baby boom) migrated from the South to the industrial centres in the North. It is striking that among the Turin population born between 1947-1957 almost 50% were born in the Southern region and only slightly more than 40% were natives of the North-western one. Quantitatively, the major migration flows in those decades have been from the South to the North, of which more than 70 % was directed to the North-West (Salvatore 1980); the main determinants were regional differences in unemployment rates and earnings levels (Salvatore 1977).

The “North-East” area, on the contrary, was an important source of migrants to the city among older cohorts, while its contribution decreases among younger cohorts, reflecting the economic history of this area in the post war decades (Caselli and Egidi 1981; Caselli and Reale 1999). The proportions coming from the “Center” and “Abroad” areas are negligible.

[9]Presidential Decree 31 October 1969, n. 1236.

The recent economic history of Italy helps to understand these dynamics. Until WWII, due to a less advanced stage of economic development, the North-Eastern regions experieced significant emigration flows both towards the North-West and foreigner countries, mostly European ones, while individuals from Southern regions emigrated mainly towards transoceanic destinations (Del Boca and Venturini 2005). After the war, the North-East went through some years of demographic depression but, at the same time, together with the Central regions, started a slow (and for a long time unrecongnized) process of economic development. This became visible in the 1970's, when a "Third Italy" (Bagnasco 1977) was discovered. This refers to the alternative industrial model characterized by small manifacturing firms, that were operating in the North-East and the Center, beyond the poor South and the modern North-West with its large scale factories. Between the 1980's and the 1990's firm birth rates were the highest in the regions of the "Third Italy" (Garofoli 1994) and from the 1990's, Veneto, one of the most important North-Eastern regions, has become a major labor importer (Anastasia and Corò 1996; Del Boca and Venturini 2005).

Figure 4 here.

Finally, figure 5 shows how the pattern of persons with different levels of education and coming from different regions of origin changed cohort by cohort.

From older to younger cohorts all the regions of origin share the same pattern of gradual decrease of lowly educated individuals and increase of medium and highly educated ones. This is true among both sexes. Cohort over cohort the gender educational differential, which saw the women overrepresented among the lowly educated, narrows down. Among the youngest cohort the distribution of education levels is virtually the same for men and women, with the exception of Southern regions where the share of low education is still bigger among women than among men. Meanwhile, men still have bigger shares of medium and high educational levels than women.

Figure 5 here.

3.2 Regression analysis results

Table 3.2 reports the results of the regression models for men and women: the baseline parameters (a, b of the Gompertz function used for the men; a, b and c

of the Makeham function used for the women), the variance of frailty σ^2 and the mortality rate ratios for the education level, the region of birth and the calendar period.

The estimated variance of frailty for the women is 0.29 while for men is 0.26. Therefore, women appear to be more heterogeneous than men with respect to their hidden susceptibility to death.

Another gender difference appears in the pace of mortality improvement over time, which is faster among men than among women. With respect to the period 1971-1990, the mortality rate ratios for the period 1991-2007 are 0.728 and 0.888 repectively. This is consistent with the recent narrowing of the gender mortality gap between Italian men and women.

The differences in mortality risks between education groups are found to be wider among women than among men. The relative risk for women with medium and low education, compared to women with high education, are 1.25 and 1.47 respectively. On the contrary, no clear gradient is visible among men: although medium and lowly educated men have an higher mortality risk compared to highly educated men, the rate ratios of medium and low education groups are very similar, 1.27 and 1.26 respectively.

Concerning the geographical differentials in mortality, the model did not identify any clear and significant pattern.

4 Discussion

After the destructions caused by WWII, Italy underwent a process of reconstruction and transformation into an industrial society. Key of this economic success was, among other factors, also the significant internal South-North migrations that accompanied the development of the industrial centres in the North of the country. In the 60's and in the 70's millions of Italians migrated from the poorer southern regions to the more diveloped northern areas.

The city of Turin, together with Milan and Genoa, was one of the three main centers of attraction and the area between these three cities was called the "industrial triangle". Turin became one of the most important industrial cities of the country, thanks to the production of FIAT, the main Italian car producing company, that still characterizes the economy of the city today.

The majority of the migrants to Turin were lowly educated individuals from the South of Italy seeking jobs in the car factories. The aim of this paper was

	Women		Men	
	mean estim.	0.025-0.975 quant.	mean estim.	0.025-0.975 quant.
a	0.008	(0.000-0.016)	0.004	(0.000-0.010)
b	0.084	(0.073-0.106)	0.069	(0.061-0.163)
c	0.000	(0.000-0.000)	-	-
σ^2	0.292	(0.174-0.367)	0.269	(0.026-0.367)
Education Level				
High	1	(-)	1	(-)
Medium	1.256	(1.053-1.347)	1.277	(1.054-1.349)
Low	1.475	(1.103-1.641)	1.268	(1.074-1.591)
Region of Birth				
N-West	1	(-)	1	(-)
N-East	1.122	(0.888-1.217)	1.075	(0.855-1.220)
Center	1.102	(0.864-1.218)	1.081	(0.854-1.217)
South	1.130	(0.904-1.220)	1.037	(0.854-1.216)
Abroad	1.082	(0.847-1.215)	1.082	(0.864-1.218)
Calendar year				
1971-1990	1	(-)	1	(-)
1991-2007	0.888	(0.671-1.035)	0.728	(0.613-0.985)

Table 2: Results of the regression models with period covariates. The model reports the mean value and the 0.025-0.975 quantiles of the empirical distribution of parameters obtained from the repeated estimates via random subsampling.

to investigate whether these migration flows had an impact on the disparities in susceptibility to death at late adult-old ages in Turin.

Our hypothesis was that the combined effect of the supposedly healthier status of migrants (Sharma et al. 1990; Singh and Siahpush 2001; Valkonen et al. 1992; Feinleib et al. 1982; Kington et al. 1998; Deboosere and Gadeyne 2005) and a faster selection due to higher mortality of the lowly educated ones involved in heavier and risker jobs could have caused a process of "homogenization" and reduction of differences in susceptibility to death. Moreover, this effect should be more pronounced among men because women were likely to be more passive actors in the migratory decision (Mincer 1978; Bielby and Bielby 1992) and less heavily involved in the industrialization process and its aftermath.

By applying a survival analysis model with unobserved heterogeneity component to a 36 year mortality follow up period, the results of the regression showed that the female population is somewhat more heterogeneous than the male population (the estimated variance of the random term was higher for women than for men).

Moreover, the male mortality gradient by education level was less defined than the women's gradient. In particular, compared to the reference category (high education), the rate ratios for women with low education were higher than the the rate ratio of women with medium education. Among men, by contrast, although medium and low education groups showed higher mortality with respect to the high education group, the rate ratios for the medium and low categories were very similar. This suggests that there is no clear mortality difference between the two educational groups and seems to confirm the hypothesis that a process of homogenization took place, which made the male population more homogeneous with respect to susceptibility to death and reduced the differences in mortality by education, at least between the medium and the low education groups. This is furthemore strengthened by the estimated lower variance of frailty compared to the women.

Another mechanism might have played a role in somehow reducing the differences in surivival, especially between educational groups among men. The so called "salmon bias" (Palloni and Arias 2004) (when weaker and sick migrants move back to their place of origin for receiving care (Brimblecombe et al. 2000; Lanska and Peterson 1995)) might have drained part of the weaker individuals out of the lowly educated population, thus contributing to its healthier and less frail status, reflected in narrower mortality differential compared to the other educational groups. Unfortunately the available data don't allow taking this mechanism into account, as it is not possible to follow the individuals up after they leave the city of Turin.

The analysis included also a variable for the macro-region of birth for which the model did not identify any clear pattern. At first this is quite surprising because Italy is characterized by significant geographical differences in the survival chances. For most of the 20^{th} century male mortality in the South was significantly lower than in the North (Caselli and Egidi 1981, 1980; Caselli and Reale 1999; Barbi and Caselli 2003) and only cohorts born after WWII show a reversing trend (Caselli and Reale 1999; Biggeri et al. 2011).

On the one hand, one must be aware that these differences refer to mortality by region of death and not by the region of birth, because the available data on regional deaths do not allow tracking of the migrants. The observation that cohorts born after WWII experience a reversed pattern supports the idea that the level of mortality was determined by the level of industrialization and of exposure to its most negative effects. For older cohorts, who were involved in the first phase of industrialization, this coincided with an exposure to its dangers that raised mortality in the industrialized regions of the North and spared the less industrialized South. As industrialization proceeded, it also improved wealth which started to compensate the mortality disadvantage with better health services, higher incomes and pensions and so on. As a consequence, for more recent cohorts, living in the economically depressed South is not an advantage anymore.

On the other hand, it is possible that the model failed to identify any pattern because the number of bootstrapping repetitions (250) might not have been sufficient, although such a number is considered by the literature adequate for very complex models (Manly 1997; Pattengale et al. 2010). A more likely explanation seems to be that the surivival advantange of the male migrants from the South (either due to the migration related health selection or due to a persisting geographical component) was nearly entirely captured by the low education variable, as the majority of male migrants from the South had a low education level. This would explain why low and medium education groups did not show clear differences in mortality, instead of presenting the expected gradient high-medium-low. It also lends support to the hypothesis that the migratory process contributed to reduce male mortality disparities and heterogeneity in susceptibility to death.

References

Aalen, O. 1994. "Effects of frailty in survival analysis." *Statistical Methods in Medical Research* 3(3):227.

Aalen, O., Ø. Borgan and H. Gjessing. 2008. *Survival and Event History Analysis*, chap. 6: Unobserved heterogeneity: The odd effects of frailty, pp. 231–270. Springer.

Akaike, H. 1974. "A new look at the statistical model identification." *Automatic Control, IEEE Transactions on* 19(6):716–723.

Anastasia, B. and G. Corò. 1996. *Evoluzione di un'economia regionale: il Nordest dopo il successo [The evolution of a regional economy: the North-East after the success.].* Nuova Dimensione Edizioni.

Andersson, G. 2004. "Childbearing after Migration: Fertility Patterns of Foreign-born Women in Sweden1." *International Migration Review* 38(2):747–774.

Bagnasco, A. 1977. *Tre Italie: la problematica territoriale dello sviluppo italiano [Three Italies: the territorial problem of the Italian development.].* Il Mulino.

Barbi, E. and G. Caselli. 2003. "Selection effects on regional differences in survivorship in Italy." *Genus* 59(2):37–61.

Bennett, S. 1993. "Inequalities in risk factors and cardiovascular mortality among Australia's immigrants." *Australian journal of public health* 17(3):251–261.

Bielby, W. and D. Bielby. 1992. "I will follow him: Family ties, gender-role beliefs, and reluctance to relocate for a better job." *American Journal of Sociology* pp. 1241–1267.

Biggeri, A., G. Accetta and V. Egidi. 2011. "Evoluzione del profilo di mortalita 30-74 anni per le coorti di nascita dal 1889 al 1968 nelle regioni italiane [Mortality Time Trends 30-74 years by Birth Cohorts 1889-1968 in the Italian Regions]." *Epidemiologia & Prevenzione* 35(5-6):50–67.

Bonifazi, C. and F. Heins. 2000. "Long-term trends of internal migration in Italy." *International Journal of Population Geography* 6(2):111–131.

Brimblecombe, N., D. Dorling and M. Shaw. 1999. "Mortality and migration in Britain, first results from the British Household Panel Survey." *Social Science & Medicine* 49(7):981–988.

———. 2000. "Migration and geographical inequalities in health in Britain." *Social Science & Medicine* 50(6):861–878.

Canteri, C. 1964. *Immigrati a Torino*. Edizioni l'Avanti!

Caselli, G. and V. Egidi. 1980. *Le differenze territoriali di mortalitá in Italia: tavole di mortalita provinciali:(1971-72)*. Istituto di Demografia della Sapienza Universitá di Roma, Rome.

———. 1981. "L'analyse des données multidimensionnelles dan l'étude des relations entre mortalité et variable socio-économiques d' environment et de comportement individuel [Multivariate methods in the analysis of the relations between mortality and socio-economic, environmental and behavioural variables]." *Genus* 37(3/4):57–91.

Caselli, G. and A. Reale. 1999. "Does cohort analysis contribute to the study of the geography of mortality?" *Genus* 55(1/2):27–59.

Chamberlain, G. 1985. "Heterogeneity, omitted variable bias, and duration dependence." In *Longitudinal Analysis of Labor Market Data*, edited by J. Heckman and B. Singer, pp. 3–38, Cambridge University Press.

Cooke, T. J. 2008. "Gender role beliefs and family migration." *Population, Space and Place* 14(3):163–175.

Deboosere, P. and S. Gadeyne. 2005. "Adult migrant mortality advantage in Belgium: evidence using census and register data." *Population (english edition)* 60(5):655–698.

Del Boca, D. and A. Venturini. 2005. "Italian Migration." In *European Migration: What Do We Know?: What Do We Know?*, edited by K. F. Zimmermann, pp. 303–336, Oxford University Press.

Doblhammer, G., R. Hoffmann, E. Muth, C. Westphal and A. Kruse. 2009. "A systematic literature review of studies analyzing the effect of sex, age, education, marital status, obesity, and smoking on health transitions." *Demographic Research* 20(5):37–64.

D'onofrio, G. 2010. "Un caso di studio: Torino." In *L'Italia delle migrazioni interne negli anni del miracolo economico*, edited by G. D'onofrio, pp. 61–90, Universitá degli Studi di Bologna.

Efron, B. 1979. "Bootstrap methods: another look at the jackknife." *The Annals of Statistics* 7(1):1–26.

Efron, B. and R. Tibshirani. 1993. *An introduction to the bootstrap*, vol. 57. Chapman & Hall/CRC.

Feinleib, M., P. Lambert, T. Zeiner-Henriksen, E. Rogot, B. Hunt and L. Ingster-Moore. 1982. "The British-Norwegian migrant study–analysis of parameters of mortality differentials associated with angina." *Biometrics* 38:55–71.

Felice, E. 2007. "I divari regionali in Italia sulla base degli indicatori sociali (1871-2001) [Regional differences in Italy based on social indicators (1871-2001)]." *Rivista di Politica Economica* 97(3/4):359–405.

Fofi, G. 1964. *L'immigrazione meridionale a Torino*. Feltrinelli.

Gail, M., S. Wieand and S. Piantadosi. 1984. "Biased estimates of treatment effect in randomized experiments with nonlinear regressions and omitted covariates." *Biometrika* 71(3):431.

Garofoli, G. 1994. "New firm formation and regional development: the Italian case." *Regional studies* 28(4):381–393.

Ginsborg, P. 1989. *Storia d'Italia dal dopoguerra a oggi: Dal" miracolo economico" agli anni'80*. Einaudi.

Gompertz, B. 1825. "On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies." *Philosophical transactions of the Royal Society of London* 115:513–583.

Hartigan, J. 1975. "Necessary and sufficient conditions for asymptotic joint normality of a statistic and its subsample values." *The Annals of Statistics* 3(3):573–580.

Hartigan, J. A. 1969. "Using subsample values as typical values." *Journal of the American Statistical Association* 64(328):1303–1317.

Hattersley, L. 1997. *Expectation of life by social class*. London: Office for National Statistics.

Hesterberg, T. 2011. "Bootstrap." *Wiley Interdisciplinary Reviews: Computational Statistics* 3(6):497–526.

Kington, R., D. Carlisle, D. McCaffrey, H. Myers and W. Allen. 1998. "Racial differences in functional status among elderly US migrants from the south." *Social Science & Medicine* 47(6):831–840.

Kolčić, I. and O. Polašek. 2009. "Healthy migrant effect within Croatia." *Collegium antropologicum* 33(Suppl.):141–145.

Krieger, N., D. Williams and N. Moss. 1997. "Measuring social class in US public health research: concepts, methodologies, and guidelines." *Annual Review of Public Health* 18(1):341–378.

Kunst, A. 2004. "Monitoring of trends in socioeconomic inequalities in mortality: Experiences from a European project." *Demographic Research* Special Collection 2(9):229–254.

Lanska, D. and P. Peterson. 1995. "Effects of interstate migration on the geographic distribution of stroke mortality in the United States." *Stroke* 26(4):554–561.

Luy, M. and G. Caselli. 2007. "The impact of a migration-caused selection effect on regional mortality differences in Italy and Germany." *Genus* 63(1/2):33–64.

Mackenbach, J. 2006. *Health inequalities: Europe in profile.* Tech. Rep.

Malmusi, D., C. Borrell and J. Benach. 2010. "Migration-related health inequalities: Showing the complex interactions between gender, social class and place of origin." *Social Science & Medicine* 71(9):1610–1619.

Manly, B. 1997. *Randomization, bootstrap and Monte Carlo methods in biology.* Chapman & Hall/CRC.

Manton, K., E. Stallard and J. Vaupel. 1981. "Methods for comparing the mortality experience of heterogeneous populations." *Demography* 18(3):389–410.

Marmot, M. 1986. "Mortality decline and widening social inequalities." *The Lancet* 328(8501):274–276.

McCredie, M., S. Williams and M. Coates. 1999. "Cancer mortality in migrants from the British Isles and continental Europe to New South Wales, Australia, 1975–1995." *International journal of cancer* 83(2):179–185.

Mincer, J. 1978. "Family Migration Decisions." *Journal of Political Economy* 86:749–775.

Mirowsky, J. and C. Ross. 2003. *Education, social status, and health.* Aldine de Gruyter.

Mussino, E. and S. Strozza. 2012. "The fertility of immigrants after arrival: The Italian case." *Demographic Research* 26(4):99–130.

Mussino, E. and A. Van Raalte. 2013. "Immigrant Fertility: A Comparative Study between Italy and Russia." *International Migration* 51(5):148–164.

Norman, P., P. Boyle and P. Rees. 2005. "Selective migration, health and deprivation: a longitudinal analysis." *Social Science & Medicine* 60(12):2755–2771.

Palloni, A. and E. Arias. 2004. "Paradox lost: explaining the Hispanic adult mortality advantage." *Demography* 41(3):385–415.

Pattengale, N., M. Alipour, O. Bininda-Emonds, B. Moret and A. Stamatakis. 2010. "How many bootstrap replicates are necessary?" *Journal of Computational Biology* 17(3):337–354.

Pellicciari, G. and G. Albertelli. 1970. *L'immigrazione nel triangolo industirale*. Franco Angeli Editore.

Politis, D. and J. Romano. 1994. "Large sample confidence regions based on subsamples under minimal assumptions." *The Annals of Statistics* 22(4):2031–2050.

Rasulo, D., T. Spadea, R. Onorati and G. Costa. 2012. "The impact of migration in all-cause mortality: The Turin Longitudinal Study, 1971–2005." *Social Science & Medicine* 74(6):897–906.

Rodriguez, G. 1994. "Statistical issues in the analysis of reproductive histories using hazard models." *Annals of the New York Academy of Sciences* 709(1):266–279.

Salvatore, D. 1977. "An econometric analysis of internal migration in Italy." *Journal of Regional Science* 17(3):395–408.

———. 1980. "A simultaneous equations model of internal migration with dynamic policy simulations and forecasting: Italy, 1952–1976." *Journal of development economics* 7(2):231–246.

Schofer, E. and J. W. Meyer. 2005. "The worldwide expansion of higher education in the twentieth century." *American sociological review* 70(6):898–920.

Sharma, R., M. Michalowski and R. Verma. 1990. "Mortality Differentials among Immigrant Populations in Canada." *International Migration* 28(4):443–450.

Shkolnikov, V. 2011. "Increasing absolute mortality disparities by education in Finland, Norway and Sweden, 1971-2000." *Journal of Epidemiology and Community Health* 66(4):372–378.

Singh, G. and M. Siahpush. 2001. "All-cause and cause-specific mortality of immigrants and native born in the United States." *American Journal of Public Health* 91(3):392.

Strand, B. 2010. "Educational inequalities in mortality over four decades in Norway: Prospective study of middle aged men and women followed for cause specific mortality, 1960-2000." *BMJ: British Medical Journal* 340:1–10.

Trussell, J. and G. Rodriguez. 1990. "Heterogeneity in demographic research." In *Convergent issues in genetics and demography*, Oxford University Press, USA.

Valkonen, T. 1993. *Socioeconomic mortality differences in Finland 1981-1990.* Tech. Rep.

———. 2001. *Trends in differential mortality in European countries*, pp. 185–322. Strasbourg: Council of Europe Publishing.

Valkonen, T., A. Brancker and M. Reijo. 1992. "Mortality differentials between three populations–residents of Scandinavia, Scandinavian immigrants to Canada and Canadian-born residents of Canada, 1979-1985." *Health reports/Statistics Canada, Canadian Centre for Health Information= Rapports sur la santé/Statistique Canada, Centre canadien d'information sur la santé* 4(2):137.

Van den Berg, G. J. and B. Drepper. 2011. "Inference for Shared-Frailty Survival Models with Left-Truncated Data." *IZA Discussion Paper No. 6031* .

van Lenthe, F., P. Martikainen and J. Mackenbach. 2007. "Neighbourhood inequalities in health and health-related behaviour: Results of selective migration?" *Health & Place* 13(1):123–137.

Vaupel, J., K. Manton and E. Stallard. 1979. "The impact of heterogeneity in individual frailty on the dynamics of mortality." *Demography* 16(3):439–454.

Westman, J., T. Martelin, T. Härkänen, S. Koskinen and K. Sundquist. 2008. "Migration and self-rated health: a comparison between Finns living in Sweden and Finns living in Finland." *Scandinavian Journal of Public Health* 36(7):698–705.

Wienke, A. 2010a. *Frailty models in survival analysis*, chap. 2: Survival analysis, pp. 15–54.

———. 2010b. *Frailty models in survival analysis*, chap. 3: Univariate frailty models, pp. 15–54.

———. 2010c. *Frailty models in survival analysis*, chap. 4: Shared Frailty Models, pp. 131–160.

Wild, S. and P. McKeigue. 1997. "Cross sectional analysis of mortality by country of birth in England and Wales, 1970-92." *BMJ* 314(7082):705.

Zarulli, V., D. Jasilionis and D. A. Jdanov. 2012. "Changes in educational differentials in old-age mortality in Finland and Sweden between 1971-1975 and 1996-2000." *Demographic Research* 26(19):489–510.