Economics
Working Papers

2019-4

Intergenerational Health Mobility: Evidence from Danish Registers
Carsten Andersen

Abstract:
To what extent status depends on family background has been of great interest in the social sciences and the general public for centuries. The transmission of income, earnings and educational attainment is often studied, while equality of opportunity with respect to health outcomes has received far less attention. This paper is the first to investigate intergenerational health mobility using high-quality administrative data from Denmark. The attractiveness of this approach lies in objective health measures and large sample sizes allowing twin analyses. I operationalise health mobility by a variety of statistics: rank-rank slopes, intergenerational correlations and sibling and identical twin correlations. Mobility in health is found to be relatively high for men, both when compared to similar US-based studies, and when contrasted with outcomes such as educational attainment and income. For Danish women, health-related dependence on family background is on par with similar statistics for income and earnings for other Scandinavian countries. Mobility is thus, perhaps somewhat nonintuitively, higher in health than in income. Contrasting sibling and identical twin correlations with parent-child associations confirm earlier findings in the literature on equality of opportunity, namely that sibling correlations capture far more variation than traditional intergenerational correlations. 14-38 percent of the variation in individual health outcomes can be ascribed to family background and genes, factors which the individual cannot be held accountable for. Only a negligible share of this variation can be explained by parental health, which suggests that other family-specific characteristics may play an important role for health mobility.

DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

Intergenerational Health Mobility: Evidence from Danish Registers

Carsten Andersen

March 25, 2019

Abstract

To what extent status depends on family background has been of great interest in the social sciences and the general public for centuries. The transmission of income, earnings and educational attainment is often studied, while equality of opportunity with respect to health outcomes has received far less attention. This paper is the first to investigate intergenerational health mobility using high-quality administrative data from Denmark. The attractiveness of this approach lies in objective health measures and large sample sizes allowing twin analyses.

I operationalise health mobility by a variety of statistics: rank-rank slopes, intergenerational correlations and sibling and identical twin correlations. Mobility in health is found to be relatively high for men, both when compared to similar US-based studies, and when contrasted with outcomes such as educational attainment and income. For Danish women, health-related dependence on family background is on par with similar statistics for income and earnings for other Scandinavian countries. Mobility is thus, perhaps somewhat nonintuitively, higher in health than in income. Contrasting sibling and identical twin correlations with parent-child associations confirm earlier findings in the literature on equality of opportunity, namely that sibling correlations capture far more variation than traditional intergenerational correlations. 14-38 percent of the variation in individual health outcomes can be ascribed to family background and genes, factors which the individual cannot be held accountable for. Only a negligible share of this variation can be explained by parental health, which suggests that other family-specific characteristics may play an important role for health mobility.

1 Introduction

Inequality, and whether it is perpetuated in the family, is at the forefront of public and academic discourse. Most would agree that society should aim for a situation in which individuals' prospects in life are not heavily, if at all, constrained by their family background. In other words, policy makers should pursue *equality of opportunity*. In this paper I estimate to what extent *health* is passed down through generations and reproduced within the family in a Danish context. Exploiting rich register data I construct an intergenerational sample that enables me to identify not only parent-child links but also siblings and twins.

Economists have shown a long standing interest in the transmission of status across generations. Research on intergenerational mobility in traditional outcomes such as income or earnings (Solon (1992); Zimmerman (1992); Jantti et al. (2006); Björklund and Jäntti (2012); Chetty et al. (2014); Landersø and Heckman (2017), and wealth (Boserup et al. (2016)) has received a lot of attention. Sociologists have been more concerned with transmission of class and occupation status (Erikson and Goldthorpe (1992); Torche (2015)). However, health, a key ingredient in human existence remains severely understudied. This is unfortunate for several reasons. In itself, health is arguably central to human well-being: Grossman (1972) develop a model in which health enters the utility function directly as something individuals derive pleasure from. Jones and Klenow (2016) find that mortality is the most important predictor of welfare across countries and time and the World Happiness Report of 2016 concludes that healthy life expectancy is as important as GDP per income in predicting happiness levels across countries (Helliwell et al. (2016)). Moreover, health impacts how individuals fare on the labour market. Health shocks induce retirement (McClellan (1998); Dwyer and Mitchell (1999); Datta Gupta et al. (2015)) and lead to unemployment and lower income (García-Gómez et al. (2013)). Thus, knowing to what extent children inherit the health status of their parents should be important for policy makers and for normative discussions of fairness. Substituting earnings or income with health as the primary outcome presents an interesting case, over and above the arguments about its importance to human welfare: it is an area where inequality differences in mortality seems to be lower, if at all present, between Scandinavian countries and the US (Avendano and Galama (2011)). Extrapolating from the case of income we may then suspect that equality of opportunity with respect to health status is similar. If this is the case, we may conclude that most of health transmission is purely genetic, or that institutions either are ineffective or not significantly different in fostering mobility. Similarly, if mobility is indeed higher in countries with universal health insurance it may be seen as indirect evidence that such policies succeed in lowering the importance of the birth lottery relative to countries with a higher fraction of employer based private health insurance. Other societal conditions less explicitly related to health such as income inequality may also play an important role for transmission of health (Marmot et al. (1991)).

Increased public and academic interest in mobility and inequality should be seen in the light of growing inequality in OECD countries. Lack of mobility simply becomes more consequential in

absolute terms when inequality is high. While most economies have been growing prior to the financial crisis, the lower part of the income distribution has experienced very low levels of real income growth (OECD (2011)). The OECD goes further to note that "Rising income inequality creates economic, social and political challenges. It can stifle upward social mobility, making it harder for talented and hard-working people to get the rewards they deserve." ((OECD, 2011, p. 40)). Inequality in health, at least when measured as disparity in life expectancy, is also on the rise. In the US, gains in life expectancy in the period 1980-2010 is concentrated almost exclusively among the upper three quintiles of the income distribution (Schanzenbach et al. (2016)). In Denmark in the same period, longevity growth is also concentrated on the best educated and the most affluent (Baadsgaard and Brønnum-Hansen (2011)).

Expanding the focus of intergenerational mobility from traditional labour market outcomes to health outcomes is challenging for several reasons. First and foremost, it is less obvious how to measure health compared to income[1]. The literature analysing parent-child health links has seen different approaches, where examples include body mass index (Dolton and Xiao (2017)), birth weight (Currie and Moretti (2007)), longevity (Parman et al. (2010)), principal components from self-reported ailments (Halliday and Mazumder (2017)) and self-reported health (Halliday et al. (2018)). As a contribution to this literature I measure health by taking the first principal component of a battery of objective health utilisation measures relating to hospital and doctor visits. This measure, as all of the above, has its strengths and weaknesses. I argue that it is reasonable to rely on health care utilisation in a context where access is free and universal and proceed to show that the measure meaningfully correlates with observables such as mortality and transition to disability pension. One significant advantage of developing a measure based on register data compared to, say, self-reported health, is the large sample sizes then being available for analysis. Moreover, the register data covers an entire population of Danes. Finally, although measures as self-reported health have been found to predict mortality quite well, the nature of such health assessments represent an inherent subjectivity which administrative data on health care utilisation obviously are not prone to[2].

Methodologically, I supplement traditional parent-child associations with sibling correlations and identical twin correlations, for the latter using a method developed in Nicoletti and Rabe (2013). The rationale for this exercise is promoted in Björklund and Jäntti (2009, 2012) but rarely used in the literature, namely that sibling correlations and identical twin correlations constitute plausible lower

---

[1]As noted in Currie and Madrian (1999): "The concept of "health" is similar to the concept of "ability" in that while everyone has some idea of what is meant by the term, it is remarkably difficult to measure." While health is difficult to measure it does not imply that income is trivial to measure. The choice of income measure can in itself affect results: Landersø and Heckman (2017) show how cross-country comparisons in mobility are sensitive to e.g., whether pre-tax or post tax income is used.

[2]The opposing view, that what matters is how people *perceive* their health status as opposed to how it may appear based on 'objective' criteria, may be argued with similar justification. Ideally, researchers should implement both approaches and compare results as Halliday et al. (2018) who in fact find similar results in health mobility for self-reported health and - less subjective - self-reported ailments.

and upper bounds, respectively, on the total importance of family background[3]. Many factors complicate the implications of mobility estimates for normative fairness considerations. Most prominently, separating effort and circumstances is notoriously difficult[4]. Hence, sizeable sibling correlations may not logically imply lack of equal opportunities although it may be seen as an indication thereof. In this paper I adopt the pragmatic view that your parents, and more generally your family, are not self-chosen[5]. On the contrary, it is a circumstance from the point of view of the offspring generation. Therefore, sibling and twin correlations are useful ways of indicating how important the birth lottery is, or in other words, whether the playing field is somewhat level for the offspring generation.

I find that health status is dependent on parental health and family background in Denmark. Thus, Denmark (and probably Scandinavia as a whole) can not be said to have equality of opportunity in terms of health status[6]. However, in terms of the magnitudes the evidence seems to suggest that this comparably egalitarian country with universal health insurance is relatively mobile. I find rank-rank slopes in the range 0.11-0.15 which is around half of what is found for US data (Halliday et al. (2018). I find sibling correlations in health in the range 0.14-0.20 and identical twin correlations between 0.34-0.38. Family background seems to play a larger role for men in terms of income relative to health while for women the dependence is more similar. Policy implications and further discussion will be provided at the end of the paper.

The rest of the paper is organized as follows. Section 2 provides some background on related studies on intergenerational mobility. Section 3 details the econometric approach used to evaluate how health is reproduced in families. Section 4 describes the data and measurement of health. Section 5 presents results. Finally, Section 6 concludes and discusses implications.

2 Background

In the following, selected results from the literature on intergenerational mobility are summarized, both health and other outcomes. I will primarily emphasize studies that measure health in adulthood

---

[3]This is possibly due to the heavy data requirements. Absent a twin database this methodology requires large sample sizes as twins only comprise 1-2 % of the general population.

[4]For one thing, effort may be a function of circumstances. As Roemer (2009) also notes, whether people should be held accountable for their outcomes can ultimately hinge on whether one believes in free will or not. However, the reader may rest assured that the question of free will is beyond the scope of this paper.

[5]As James Heckman puts it, picking the wrong parents may be the "biggest market failure of all" (Reeves (2018)).

[6]Admittedly, some would disagree with the logical implication from "positive intergenerational correlations and sibling correlations" to "lack of equality of opportunity" that I adopt here. Jencks and Tach (2006) is particularly critical of the mapping from equality of opportunity to intergenerational correlations (and sibling correlations, I assume). Again, the crucial, yet problematic, distinction between *circumstances* and *personal choices* is invoked to evaluate whether some inequality is "justifiable". A sensible and pragmatic approach is taken in Miles Corak's summary of some attempts in the literature to group factors into whether the individual is in control of them or not: "[they] find that indices of inequality of opportunity are in fact strongly correlated with indicators of generational mobility, be it in earnings or in education. It is in this sense that the Great Gatsby Curve can be understood to be signaling—rather than directly measuring—a negative relationship between inequality and equality of opportunity." (Corak, 2013, p. 6).

as opposed to e.g., childhood ailments, transmission of birth weight or the correlations between mother health and caregiver reported child health[7]. The vastness of the general literature on social mobility makes it impossible to provide a comprehensive summary here, but excellent surveys are available elsewhere (Solon (1999); Black and Devereux (2010); Torche (2015)). The literature studying traditional labour market outcomes such as income and education can and should not be viewed as separate from the studies on mobility in health but for clarity they are summarized separately below.

2.1 Health Outcomes

The literature on the relation between health outcomes and family background at least goes back to the turn of the nineteenth century. Galton (1886) studied heredity of human height and found that the average height of individuals in the second generation is given as two thirds of the parents' deviation from the population average. From this simple starting point recent research has diverged considerably in terms of methodology and how to quantify health. The study most comparable to mine is probably Halliday et al. (2018) who proxy health by self-reported health and, in a robustness check, by objective health measures. They do not undertake sibling correlations but the similarity lies in measuring health in adulthood, reporting statistics as the rank-rank slopes and applying the rigour of the vast literature on income mobility, such as dealing with attenuation bias and life cycle bias. They find rank-rank slopes in the range 0.21-0.29 where the strongest link is found between mother and daughter.

On US twin data Kessler et al. (2004) find monozygotic twin correlations in health perception of 0.26-0.38 and 0.20-0.38 in self-rated physical health, in both instances female twin pairs correlating more than males. The corresponding figures for non-twin sibling are lower, at 0.11-0.15 and 0.10-0.38, respectively. However, a weakness of the study is the failure to average outcomes over several years and the low response rate of the twin part of the survey (MIDUS). Bügelmayer and Schnitzlein (2018) find sibling correlations in physical health of 0.23 using the German Socio-Economic Panel (SOEP) and Roos et al. (2014) report comparable magnitudes of 0.26-0.32 when proxying health by hospital and physician costs in a Canadian context. Both these studies, however, are based on individuals in their youth or adolescence[8].

Many other studies analyse transmission of anthropometric outcomes such as height (as Galton), weight and body mass index. For Sweden, Björklund and Jäntti (2012) find a brother correlation of

---

[7]For example, Halliday and Mazumder (2017) study sibling correlations among children aged 0-18 in caregiver reported health and diagnoses. Coneus and Spiess (2012) investigate the relation between self-reported health of the mother and the health of her offspring aged 0-3. In terms of understanding the origin of health inequalities and the strength at different points of the lifecycle, such studies are highly interesting, but due to the very different ages of measurement less meaningful to benchmark my estimates to.

[8]In the current paper, sibling correlations are computed based on outcomes in the age interval 36-50. It is not clear which correlation *ceteris paribus* one would expect to be highest. On the one hand, sibling correlations should be higher if individuals live together, and thus share an environment and influence each other. On the other hand, the correlation may be higher later in life when (health) outcomes to a larger extent are realised.

adult height of 0.53 which is virtually identical to what Mazumder (2008) reports for the US. For body mass index and weight, Mazumder (2008) estimates sibling correlations around 0.27-0.33, not too different from 0.35 for body mass index found for German adolescents (Bügelmayer and Schnitzlein (2018)). Also, when gauging mobility in body mass index by parent-child associations, Dolton and Xiao (2017) find a striking resemblance across six very different countries[9], all of them displaying intergenerational elasticities of around 0.20. The advantage of these anthropometric measures is that they are readily available in many data sets and probably prone to only negligible measurement error. The downside is that it is less clear whether they should be thought of as "health" variables or, say, as risk factors. Also, while the outcomes are continuous it can complicate interpretation that they are not monotone, i.e., it may be considered unhealthy to be obese as well as being underweight.

Finally, some papers study how longevity depends on family background. Parman et al. (2010), using death certificates from North Carolina, finds an elasticity of 0.21-0.28 between parent and child life span. Piraino et al. (2014) assemble a fascinating data set with longevity of four generations of settlers in South Africa's Cape Colony, allowing them to investigate parent-child associations and sibling, cousin and second cousin correlations. They find relatively low sibling correlations of 0.08-0.15 depending on specification. This contrasts with a parent-child elasticity of 0.05-0.1, showing that parental longevity explains a very small fraction of the overall sibling longevity resemblance[10]. Another relevant study of reference, both by virtue of its Danish sample and its focus on identical twins, is Herskind et al. (1996) who find identical twin correlations in longevity of 0.18-0.33 and fraternal twin correlations of 0.02-0.08. These estimates should, however, also be seen in the light of low sample sizes and precision. While longevity is obviously appealing as an intuitive metric of health, it involves the downside of yielding very retrospective insights if one wants to analyse two consecutive generations with fully resolved life spans. For instance, all of the above mentioned studies sample individuals born before 1900, which may be an interesting study population in its own right but not necessarily representative for the experience of a contemporary population.

At this point, only few conclusions can be drawn from the literature on reproduction of health in families. In that sense, the literature is not as evolved as the corresponding literature on income or educational mobility. The health literature is simply too sparse and scattered in terms of geographical area of study, time period, health measurement and empirical methodology. For example, while the income literature is relatively unified in the conclusion that Nordic countries exhibit higher mobility than, say, Canada, which again is more mobile than the US, no such conclusion can, at this point, be drawn with health as the outcome domain. However, at least three conclusions stand out as somewhat robust. First, the estimated magnitudes of mobility coefficients differ widely depending on health measure (for example, height versus longevity). Second, the health of females depend more on family background than that of males. Third, the bulk of the estimates suggest that health as

---

[9]Countries are sampled to be at very different stages of economic development: China, Indonesia, England, Spain, Mexico and the US.

[10]A more elaborate unfolding of this argument can be found in the Methodology section of this paper, specifically see equation 3.

a domain exhibits higher mobility than more often studied labour market outcomes[11]. This paper contributes to the literature by bringing a relatively novel methodology to high-quality health to an institutional setting different from places as Germany, Canada and the US.

2.2 Income and Education

2.2.1 Income

The literature on intergenerational mobility in income and earnings has taken big steps forward over the past decades. Work by Solon (1992) and Zimmerman (1992) showed that using single-year measurements of earnings led to a downward bias in the estimated intergenerational elasticities, hence overstating mobility. For the US Solon (1992) reported an intergenerational elasticity of 0.41, considerably higher than the conclusion of Becker and Tomes that "The point estimates for most of the studies indicate that a 10% increase in father's earnings (or income) raises son's earnings by less than 2%." (Becker and Tomes (1986)). A decade after, as two-generation datasets became longer and it became even more apparent that intergenerational persistence turned out to be higher than initially believed and that mobility actually significantly differed between countries[12] (Corak (2004)). Jantti et al. (2006) find, in terms of intergenerational correlations based on six selected countries, that earnings mobility is higher in the Nordic countries than in the US (0.09 versus 0.36), with the UK in between (0.20). A similar picture emerges when measuring mobility by brother correlations in long-run earnings, where this correlation is in the range 0.14-0.26 for the Nordic countries and 0.40 in the US (Björklund et al. (2002)). In general, as will be repeated throughout this paper, sibling, and specifically also twin correlations, often provide a fuller picture of equality of opportunity than simple parent-child associations. For example, Björklund and Jäntti (2012), contrasting sibling and twin correlations with parent-child associations, find that family background explains between 20 (16) and 77 (40) percent of the variation in earnings for Swedish brothers (sisters), and that no more than

---

[11]At first sight this may seem counter-intuitive as many probably conjecture that health is somehow very "genetically determined" as opposed to income or education which has more often been the explicit focus of equality of opportunity enhancing policies. Halliday et al. (2018), finding that the transmission of health is substantially weaker than the income transmission, speculate that policies and institutions may be much more effective at breaking intergenerational health linkages, through for example provision of clean air and water and adequate nutrition. In contrast, income persistence is still relatively high, possibly due to increasing returns to schooling.

[12]In Piketty (2000) it was noted that: "Overall, the relative consensus at the end of the twentieth century seems to be that commonality and inertia are the main characteristics of intergenerational mobility: mobility rates just do not seem to vary very much [between countries] [...] it is by no means impossible that, as better data sets become available and more detailed comparative studies develop, we become able to identify interesting cross-country variations". Having nearly twenty years more of evidence it is now clear that better data and methods have demonstrated that persistence in various status measures are higher than believed thirty years ago, and that there exist significant differences between countries (for one illustration of this, see Figure 2 in Corak (2012). This curve, showing negative correlation between inequality and intergenerational mobility, is known as 'The Great Gatsby Curve'.). Almost universally, it is found that of the OECD countries, the US exhibit the highest level of persistence in status, followed by the UK and many continental European countries, while Scandinavia and Canada enjoy the highest levels of mobility.

ten percent of this sibling resemblance can be explained by parental earnings.

Landersø and Heckman (2017) nuanced the discussion on cross-country comparisons, and inter-generational mobility in general, by pointing out that conclusions may be very sensitive to the income measure used. In particular, Denmark and the US exhibit more similar mobility levels before government transfers are taken into account, hence the seemingly higher mobility levels in Scandinavia are primarily achieved through income redistribution policies[13]. However, it remains an open question which income measure best captures what researchers are actually interested in measuring, regardless of whether the focus is a cross-country comparison or not.

Mobility variation *within* countries have also been demonstrated. Chetty et al. (2014) used 40 million observations of federal tax records to show that the level of mobility *within* the US varies extensively: some local US areas are as mobile as Denmark and Canada, while others are less mobile than any developed country for which data are available. Specifically, when restricting attention to the largest commuting zones in the US, the authors find a range of rank-rank slopes from 0.23 (Los Angeles, California) to 0.43 (Cincinatti, Ohio). Another effect of this paper was the popularization of the rank-rank measure. In this paper I supplement intergenerational correlations with results on (health) rank measures as it eases comparison with other domains such as income and earnings. As for the attention given to the subject of inequality and status perpetuation the year of 2014 also saw publication of *The Capital* by Thomas Piketty, where it is set out how contemporary society, due to the slow output growth relative to the return on capital, are headed for a situation with increasing inequality and decreasing mobility[14].

2.2.2 Education

The literature on educational mobility faces different challenges than than studies of income or health mobility. In terms of life cycle issues schooling is more straightforward to handle as people tend to have completed education by their late twenties. On the other hand, as pointed out by Black and Devereux (2010) it is not entirely clear whether education meaningfully can be seen as a continuous variable (e.g., years of education) or whether the discrete nature should be respected (e.g., high school graduation versus college degree). Conceptually, education is closely linked to outcomes such as health and income. Economists have emphasized human capital as a key driver of reproduction of (economic) inequality through generations (e.g., Becker and Tomes (1986)), and within health economics education plays a prominent role in the Grossman model (Grossman (1972)) in that higher educated agents more efficiently transform health inputs into health capital.

In terms of cross-country analyses the study of Hertz et al. (2007) probably represent the most

---

[13]The authors emphasize different mechanisms explaining these similar pre-transfer mobility levels, one of them being remarkably similar educational mobility, which however has been challenged (Andrade and Thomsen (2018))

[14]If someone on New Year's Eve of 2013 had predicted that a 700 page book by a professor in Economics would top the Amazon bestseller list next year this person might have been ridiculed. Perhaps a more credible prediction would have been that such a book would top the list over the most abandoned books on the Kindle, which (also) turned out to be the case (Wall Street Journal (2014))

comprehensive comparison of transmission of educational inequality. To the extent possible, they use comparable sample and variable definitions and found the highest intergenerational correlations in years of education of 0.60 in South America, 0.46 in the US, 0.40 in Western Europe and 0.30-0.40 in the Nordic countries. Once again, sibling correlations turn out to capture far more intergenerational transmission than parent-child associations. Björklund and Jäntti (2012), using Swedish data, that parental schooling only accounts for a third of the sibling resemblance in schooling. Sibling correlations in years of education are around 0.4, identical twin correlations around 0.75. Using Danish data, Bredtmann and Smith (2018) finds the same magnitude in years of education but when breaking the education variable up in different categories it seems that the correlation in obtaining a tertiary degree is higher than the sibling correlation in acquiring upper secondary education.

In conclusion, across every outcome, geographic area and time period considered there seems to be strong evidence that outcomes are linked to the family background of the individual. Here, "outcomes" are to be understood in a broad sense, including various measures of income, education or health. Magnitudes often differ both by country and across domains in the sense that educational attainment seems to be the domain with the strongest transmission, followed by income, and then health outcomes. While these figures are useful benchmarks, and while it is instructive to design variables that are readily comparable across domains, it does not indicate a hierarchy of importance. For example, the above ranking does not indicate that intergenerational health mobility is somehow less pressing for policy makers to address. In terms of human welfare, it may potentially be more detrimental to "fall behind" in some domains than others.

Another issue is the interplay of the variables considered above. In reality, the mobility in any outcome is likely a function of the transmission of many other outcomes in ways that are extraordinarily complex to disentangle. In the descriptive part of the literature, which this paper and most of the cited references fall into, it is often a matter of convention which variable is presented as the outcome of interest and which outcomes are hypothesized to be plausible 'mechanisms'. One of the aims of this paper is simply to acknowledge the difficulty in separating out the intricate causal patterns but instead provide descriptive findings on health mobility as an outcome that is interesting in its own right.

3   Methodology

This section falls into two parts. The first part is concerned with traditional parent-child regression models while the second part deals with sibling correlations in the offspring generation. The subsections reflect the various methods that will be used to measure transmission of health status. Later it will be argued that sibling correlations are a much broader and richer way of measurement than the traditional parent-child associations that constitute the first subsection. However, parent-child associations are included because they are necessary to demonstrate the above point empirically and they represent a natural benchmark to related literature.

3.1 Parent-Child Associations: The Workhorse model of intergenerational mobility

The departure of most of the literature on intergenerational mobility is a simple linear regression model relating outcomes of consecutive generations as follows.

$$Y_C = \alpha + \beta Y_P + \delta \mathbf{X} + \epsilon, \tag{1}$$

where Y_c and Y_p denotes status of the child and the parent, respectively. 'Status' is to be understood broadly and may encompass absolute or relative measures of income status, education status, health status and so on. The most frequent implementation of the above is probably the case where both Y-variables are log of income or earnings. The model then answers what percentage child income is expected to increase if parent income increases by one percent. This statistic is referred to as the intergenerational elasticity.

$\mathbf{X}$ is a vector variables and may contain potential mechanisms such as education and cognitive ability. ϵ is the error term and thus represents all unobserved factors affecting child status. Interpretation of this term depends on the content of $\mathbf{X}$. It will in most applications contain factors such as genetic endowment and random shocks. β is the parameter of interest and shows the strength of the association between parent and child outcomes. If positive and numerically large, mobility is low. If close to zero, intergenerational mobility is high; that is, child status is not particularly dependent on parent status. In the current context, X will be empty as investigating mechanisms is beyond the scope of this paper. Also, all outcome variables are residualised on cohort effects before estimation.

A natural companion statistic to the intergenerational elasticity is the intergenerational correlation. Simple algebra performed on the expression for β gives the intergenerational correlation, γ, as

$$\gamma = \frac{\sigma_P}{\sigma_C}\beta, \tag{2}$$

where notation follows equation (1) and σ_C and σ_P are the associated standard deviations of child and parent distributions, respectively. It is seen that the two mobility statistics are identical only when $\sigma_C = \sigma_P$, i.e., when child and parent outcomes are equally dispersed. In the modal outcome variable in the literature, income, this will often lead to elasticities being higher than correlations because the outcomes of the child generation are more dispersed (that is, $\sigma_C > \sigma_P$). Thus, γ can be thought of as a statistic that accounts for the fact that inequality may differ between generations. A third statistic, the rank-rank slope, has gained prominence in the literature recently (e.g., Chetty et al. (2014); Halliday et al. (2018)). Mathematically, the rank-rank slope is the Spearman correlation coefficient between Y_C and Y_P. This will be referred to as β_R. This measure solely uses information on the relative position between individuals; hence, it makes no assumptions about the distribution of the data apart from monotonicity. It has also been found to be more robust to measurement issues such as life-cycle bias (discussed in the Appendix) and inclusion of zero incomes (Mazumder (2015)). Another appealing feature of the rank-rank slope in the current setting is that it readily

compares across domains (e.g., comparing rank-rank slopes with rank-rank slopes in health). However, empirically it turns out that for health rank-rank slopes and intergenerational correlations are virtually identical (see the Results section). The empirical part of this paper will be principally concerned with intergenerational correlations and rank-rank slopes as opposed to elasticities or 'raw' interpretation of β.

Whether β, or, by implication, all the above statistics that are related to β, should be interpreted causally, depends on the research design and which assumptions the researcher wants to impose. Simply computing the raw associations between generations obviously does not reveal whether an increase in parent status will affect child status or whether e.g., immutable genes alone are responsible for the statistical relationship. The literature on intergenerational associations falls in two categories in this respect. One part aims at estimating "causal effects" by exploiting plausibly exogenous shifts in parent status or exploiting sibling or twin fixed effects estimators (Black and Devereux (2010)). In the current setting of health, one could explore whether increasing parental health by a certain intervention would lead to increased offspring health. Another strand of the literature maintains that the more descriptive mobility estimates are interesting in and of themselves, sometimes combined with analyses of which background variables attenuate the relationship (e.g., Mazumder (2008); Halliday et al. (2018). This paper falls in the second category. That is, the analysis here is deliberately descriptive, and going forward it is useful to note that e.g., "sibling correlations" and "sibling fixed effects" are distinct methods, where only the former will be employed here. Similarly, the aim is not to find an instrumental variable affecting parental health to somehow back out the "causal effect" of parental health on child health. Such studies are (also) interesting but the ambition of this paper is simply to *describe* the extent of transmission and reproduction of health within families. Conceptually, the two above mentioned strands of previous literature are quite different. Whereas the latter, descriptive strand estimate only a few parameters, the former, more causally oriented is virtually unlimited in scope. One could imagine a multitude of interventions taking place, at different ages in different doses and forms, answering often locally defined questions in a concise way. The descriptive part of the literature instead estimates (few) parameters, that are themselves the result of institutional settings, genetic endowments and infinite feedback loops between an essentially infinite number of variables.

3.2 Sibling and Twin Correlations: Bounds on the Total Importance of Family Background

3.2.1 Sibling Correlations

As noted by many researchers the literature on intergenerational mobility may overlook important components of family background by focusing on a *single* parental resource[15] (Solon (1999); Mazumder (2008); Björklund and Jäntti (2009); Bredtmann and Smith (2018)). Ultimately, what is, or, what *should*, be relevant for gauging equality of opportunity is whether health status is reproduced according

---

[15]i.e., Y_P.

to general family background, not the correlation of child health with one specific parental variable. In addition to the parental resource, a sibling correlation takes into account all unobserved factors that are shared by siblings and that are uncorrelated with the parental resource. This can be written informally as

$$\text{Sibling correlation} = \gamma^2 + \text{other shared factors} \tag{3}$$
$$\text{that are uncorrelated with the parental variable,}$$

where γ is the intergenerational correlation from equation (2). See Solon (1999) for a derivation. The above equation shows how much of the importance of family background that potentially is missed by only estimating models of the kind in equation (1). The conceptual framework of the sibling correlation is simple. The outcome of the child, Y (dropping subscript from equation (1) for simplicity), can be decomposed as

$$Y_{ij} = a_i + b_{ij}, \tag{4}$$

where a_i is shared by siblings in a family and b_{ij} is the individual-level deviation from a_i for individual j in family i. These two components are then orthogonal by construction. Then, the variance of the outcome is given as the sum of the variances of the two individual components

$$\sigma_Y^2 = \sigma_a^2 + \sigma_b^2 \tag{5}$$

Thus, the fraction of the variance in the outcome, which is attributable to family background, is given as

$$\rho = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_b^2} \tag{6}$$

The sibling correlation, ρ, has the interpretation of being the correlation being two randomly picked siblings. It is also be referred to as the *intraclass correlation*. The numerator holds the between-familiy variation and the denominator holds the sum of this component and the within-family variation. Intuitively, if the numerator is large relative to the denominator, that is, if the variation between families is large relative to the individual variation, it must imply that what happens in the family, is crucial. In other words, siblings should then resemble each other to a high degree. If ρ is zero, family background has no bearing on health outcomes of children, while a positive ρ indicates that siblings to some degree share health outcomes, thus pointing to the importance of family background. In all the estimations that follow, the variance components of equation (5) are estimated by Restricted Maximum Likelihood (REML), following Mazumder (2008); Björklund and Jäntti (2012); Bredtmann and Smith (2018). Similarly in line with most literature, I define siblings as individuals who share

a biological mother and a biological father[16]. The outcome variables for all sibling correlations are the 'raw' outcomes, i.e., predicted health and log of income[17]. All standard errors are obtained by bootstrapping with 100 replications.

While the sibling correlation arguably is a broader measure of intergenerational mobility than regressions of the kind in equation (1), it is important to note there are aspects it does not capture. Examples include the fifty percent genes the siblings do not share, differential treatment by parents, and (other) time-varying factors in the upbringing. Still, the sibling correlation is a useful way to gauge whether conventional intergenerational mobility estimates miss important aspects of family background. The advantage of the sibling correlation is that it reflects everything that is time constant within the family, observed as well as unobserved factors. Therefore, variables difficult to measure such as family values, attitudes and preferences, to the extent that they do not vary over time, are also captured by this broad 'omnibus' measure[18]. For example, Solon (1999) notes that only forty percent of US sibling correlations in long-run earnings can be explained by parental earnings. Björklund and Jäntti (2012) find that importance of background is even more severely underestimated in a Swedish setting for a wide range of measures, including earnings, height and IQ, when estimating equations of the form in equation (1) compared to estimating sibling correlations.

3.2.2 Identical Twin Correlations

As described above, the sibling correlation can be interpreted as a lower bound on the total importance of family background. Björklund and Salvanes (2011) make the point that identical twin correlations can be seen as a corresponding upper bound. On top of what non-twin siblings share, identical twins fully share genetic endowment and shocks to the family[19]. Plausibly, they also interact more extensively than the general population, potentially further increasing the correlation between them. One the one hand, this can be seen as "excessive" interaction that is misleading to generalize to the general population but on the other hand it may lead to greater trust in the identical twin correlation being a genuine upper bound.

For estimation of the identical twin correlations I employ a technique from Nicoletti and Rabe (2013) and Björklund and Jäntti (2012). The method is a way to identify identical twin correlations in situations where the data do not have information on zygosity as is the case here. This method is described in detail in the above references, so an explicit presentation will not be given here[20]. The

---

[16]This is based on the registers which is of high quality in terms of parent-child linkage. However, in very few cases, the definition will instead only capture the social father.

[17]This is most comparable to other literature on sibling correlations. However, for future research it would be interesting to contrast rank sibling correlations with sibling correlations in the 'raw' variables.

[18]Whereas parent-child associations capture how far the apple falls from the tree, sibling correlations measure how far the apples fall from each other.

[19]Regarding genetics, even identical twins have different epigenetic profiles at age five (Mill et al. (2006). However, the bounding argument made here will for practical purposes disregard such epigenetic drift.

[20]Specifically, the reader is referred to Appendix A of Björklund and Jäntti (2012).

key insight of the method is that twins of different sexes *must* be dizygotic[21]. If one then assumes that the difference between same-sex twins and mixed-sex twins is identical to the difference between same sex non-twin siblings and mixed-sex non-twin siblings the correlation between monozygotic twins is identified.

4 Data

4.1 Data Sources and Sample Selection

I use Danish register data provided by Statistics Denmark. All Danish registers contain a unique personal identifier allowing me to link individuals to parents and across different registers. In addition to standard socio-economic information the data has information on general practitioner visits and in- and outpatient hospital care. Key registers used for investigating health outcomes, with starting year in parenthesis, are the The National Patient Register (1977) and The National Health Insurance Service Register (1990). The National Patient Register has information on all patients discharged from nonpsychiatric hospitals since 1977, and for each patient contact a diagnosis is recorded according to the International Classification of Diseases[22]. The National Health Insurance Service Register contains information about the activities of health professionals contracted with the tax-funded public healthcare system and thus documents activities in primary health care[23]. For outcomes with respect to income I draw on the Income Statistics Register (1980), and the Cause of Death Register (1986) is used for restricting the sample in terms of longevity.

For constructing a health measure the aim is to obtain health information on two consecutive generations, and by implication, information on sibling and twin links. On the one hand, this speaks for having as early generations as possible as health information is revealed progressively over the life cycle. Also, to overcome measurement error (attenuation bias) it is likely misleading to include only, say, one year of health measurement[24]. On the other hand it is not desirable to include too early cohorts as they to a large extent will be deceased before they enter the health registers. Further, the experience of individuals of very early cohorts may not be representative of the current day environment.

The Danish registers help balance these requirement as they imply some concrete restrictions. Linkage percentages to parents rapidly decline moving down birth cohorts in the 1950s. And currently, parents (and children) need to be alive in 1986 for the analyst to observe the link. Moreover, my subset of the data ends in 1963 (i.e., with birth cohort 1963). In the main specification I select children born between 1960-1963 with parents born 1930-1940 to balance these requirements. With regards to

---

[21]As opposed to being confronted with same-sex twins where only around half will be identical twins.

[22]For information about this data source, see Schmidt et al. (2015).

[23]For more on this registry see Sahl Andersen et al. (2011).

[24]In the income literature it is suggested to use at least 9 years of measurement to overcome this bias (Mazumder (2005)). For health, many studies fail to consider these issues. One exception is Halliday et al. (2018) who find that estimates tend to flatten after 10 years of included health measurements.

health measurement the oldest birth cohort, 1930, is 60 years old when they start being observed in the National Health Insurance Service Register. With respect to income the sample selection implies that some cohorts are relatively old before a multi-year average of income can be observed. The oldest cohort, 1930, will have income measured in 1980, thus in the first year of measurement they are 50, which is past the prime age of proxying lifetime income (Mazumder (2015)). If anything, I then underestimate mobility in income due to measurement error in the independent variable. However, I can compare my income mobility findings to others who have investigated this using Danish data (e.g., Landersø and Heckman (2017)). The income measure employed is total income before deductions and taxes.

The cohort selection and utilisation of data sources is illustrated in Figure 1, which is scaled to show the years 1918-2018, where the text above the line denotes the cohort selection and the text below the line denotes the data sources used. For health, I measure parental health in the age interval 60-70 and child health in age 36-50[25]. I proxy lifetime income by averaging over the ages 50-55 for parents and 36-50 for children. The most illuminating comparison is to the results with the father being the parent as changing labour market norms imply that the income of women born in the 1930s is a somewhat misleading indicator of their social status. Measuring income at ages 50-55 is on the high side compared to recommendations in the literature but reflects that the data were not specifically set up for this analysis. It turns out that the resulting mobility statistics serve as a useful benchmark nonetheless. The primary estimation sample is then the subset of children in birth cohorts 1960-1963 whose parents are neither dead nor emigrated by age 70 and where the fathers have non-missing income information. The children themselves need to be alive by age 50 and have non-missing income information. In total, this leaves 36,178 sons from 32,674 families and 35,282 daughters from 31,983 families. Thus, the total sample size of individuals in the main sample is 71,460. The main sample contains 444 male twins, 446 female twins and 430 mixed sex twins.

4.2 Health Measure and Descriptive Statistics

As alluded to earlier the literature offers many examples of how to measure health. Approaches include BMI, longevity and self-reported health. The aim here is to extract a measure of health that is approximately continuous lending itself to comparison with other (percentile) rank-based measures[26]. Hence, mortality as an outcome measure is ruled out at the outset as not even the parent generation will provide fully uncensored outcomes (see Figure 1). In the following I will construct a health measure by taking the first principal component of a battery of indicator variables from health utilisation measures of hospital and general practitioner visits (the latter also referred to as "doctor visits"). On the face of it this may be seen as a conflation of two distinct concepts, namely *health* and *health care*. Those concepts are indeed distinguished between in the Grossman model (Grossman (1972)), where

---

[25]It may seem suboptimal to measure outcomes at different points in the life cycle. However, results turn out to be insensitive to both particular cohort and age restrictions.

[26]For example, the rank-rank correlation as popularized in Chetty et al. (2014).

individuals demand health care as a means to "producing" health which is treated as a durable capital good. Also, there is no denying that individuals may seek different levels of care given the same levels of underlying "true" health status.

However, the claim I will make here is simple: When health care is essentially free, which it is in a Danish context, individuals seek care when they need it. Although there are hassle costs associated with seeking health care the most significant impediment to care, monetary costs and lack of insurance, is virtually nonexisting in the current setting of universal health insurance[27]. Thus, health care utilisation is a reasonable proxy for underlying latent health status[28]. Below it is shown that the extracted health measure meaningfully correlates with relatively objective health events such as transition to mortality and disability pension. The health measure consists of both hospital stays and doctor visits, where the latter may seem the most suspect in terms of also capturing "behaviour", e.g., risk aversion. However, it turns out that objective health outcomes also exhibit uniform correlation with health care utilisation when the hospital dimension is left out. In short: individuals utilising more health care are on average less healthy, whether they only consume primary health care or whether health care utilisation includes both primary care and hospital visits. Equating health care with health status is not a novel approach in the literature. It is listed as one of several summary measures of health (Currie and Madrian, 1999, p. 3314). The staunch critic unpersuaded by the above rationale for proxying health by health care utilization may instead find it helpful to view this outcome as interesting in itself. In other words, from a perspective of worldwide rapidly growing health expenditures it is equally interesting to analyse whether the tendency to consume health services is passed on from parents to children and reproduced within the family.

Table 1 with descriptive statistics shows the variables going into the health measure. The first set of variables are the hospital variables. The categorization of these diagnoses are made on the basis of a 23-category aggregation of International Classification of Diseases (ICD) used by Statistics Denmark (Statistics Denmark (2019)). These variables are all from the National Patient Register. This register contains individual level inpatient and outpatient hospital contacts and the duration of the stay. From this information, I construct the set of binary variables in Table 1 which indicates the fraction who have ever been registered with the given ailment (hence, disregarding information about the duration of the stay). The intuition is simple: having had hospital visits within many diagnosis categories is associated with worse health. When looking at the included diagnoses in the Table 1 it may seem that they do not all capture "important" health events. For instance, contacts within the cardiovascular category are more intuitively related to a concept of bad health than, say, diagnoses related to eye-sight or infections. Ultimately, I included all of the diagnosis categories apart from

---

[27]Also, the health personnel has to deem them in need of care for a large part of the health care variables I employ. In general, an individual can visit her physician almost infinitely often but most hospital services will require some kind of referral. That is, the general practictioner acts as a gate keeper in the Danish health care system.

[28]Of course, this approach would be very questionable in a setting where it is common to experience financial barriers to vital health care. Figure 1 in Devaux and De Looper (2012) shows how number of doctor consultations in the US correlates positively with income, where the opposite is the case for Denmark.

sterilisations, poisonings and concussions. I also excluded the explicitly gender related categories such as pregnancy complications and gynecological conditions[29]. Further excluding some categories is arguably problematic. For instance, while impaired eye sight may sometimes be random it can also be the consequence of diabetes (Shah and Gardner (2017)). Hence, I have refrained from further selecting certain kinds of diagnoses but have instead opted for a 'catch-all' approach. It is certainly interesting to distinguish between different kinds of diagnoses: for genetic reasons alone some are probably to a higher degree transmitted within the family than others. Furthermore, behaviour and the interaction between environmental and genetic influences is obviously a monumental research question in general. However, such an exercise is beyond the scope of this paper as the primary motivation lies in obtaining a continuous one-dimensional health measure; that is, a measure that is easily comparable to other outcome variables within the general literature on intergenerational mobility and equal opportunity. Specifically, the aim is to measure aggregate health in a meaningful way. Hence, for the purposes of this paper it is second order nature whether health is a result of "genetics", "behaviour" or the interaction.

In Table 1 it is seen that for most categories women are more often diagnosed relative to men. Also, unsurprisingly, the parent generation suffers more from a category like cardiovascular diseases. 29 percent of fathers have been in contact with a hospital on this account some time during his sixties. The next variables, the five quintile variables of physician contacts are made based on the National Health Insurance Service Register. This includes services as home visits and telephone consultations, but dentist services have been excluded as they are not fully subsidized. Again, the intuition is that more physician contacts are associated with worse health on average. The register contains some (incomplete) distinctions between whether a contact is of 'preventive' nature or not but excluding such visits does not affect results. Moving to the lower part of Table 1 shows the mean number of doctor visits by gender and and type-specific quintile. Large differences are seen here. For example, fathers in the lowest quintile on average only have 20 doctor contacts over the course of 10 years, while this number is 255 for the highest quintile. In general, the medical literature find that women consume more health services than males (Bertakis et al. (2000); Redondo-Sendino et al. (2006)). It is therefore unsurprising to see from Table 1 that women have more doctor visits than men. It makes little sense to comprehensively compare the evolution of ailments and general health care usage across generations. Secular trends in hospital diagnoses, health and labour markets complicate interpretation of Table 1 as telling us something about differences in absolute health. However, this should pose no problem as my specifications standardize intergenerational associations by either the intergenerational correlation or the rank-rank slope. In addition, all outcome variables are net of cohort effects. Even if it should be a case of concern it presents yet another argument for relying on sibling correlations, from a pure measurement perspective, over conventional approaches regressing parent status on child status. Siblings are often closely spaced in time compared to parents and offspring. Hence, secular

---

[29]Such conditions may reflect health levels as well as fertility decision. Results for women are insensitive to inclusion of these.

trends broadly defined will largely be shared by siblings.

All the variables of Table 1 form the ingredients in condensing multidimensional health register information into a one-dimensional health measure via principal components[30]. For rank-rank specifications the individuals are then assigned a percentile rank within each cohort and gender according to their predicted value of the first principal component. The ambition is not to construct an elaborate factor model of a latent phenomenon, but it is rather an exercise in pragmatic dimensionality reduction. Figure 2 shows the scree plots from the estimation procedure by generation and sex. The plots show the eigenvalues on the y-axis as a function of the included factors on the x-axis. The four plots are seen to exhibit a similar pattern. According to two rules of thumb, the *Kaiser rule*, and the *scree test* (Brown (2014)), the number of factors to include should be guided by the number of factors with eigenvalue above 1 and where on the graph the curve seems to be genuinely levelling off[31]. In the light of these it seems reasonable to retain only the first component although this does not strictly adhere to the first rule of thumb. However, in all four graphs the first component is seen to, by far, contain the most information, where the added gain from each additional component is very marginal[32]. Also, the whole purpose of this section is to come up with a one-dimensional health measure, which is comparable to e.g., income measures, so in a sense the scree plots can be seen as confirmatory rather than exploratory. Figure 3 shows the component loadings for the various types. The interpretation of the component loadings is not straightforward as all the input variables are binary. However, the loadings can be evaluated according to their sign and the relative magnitudes. First, it is seen that the loadings are relatively similar for sons, daughters, fathers and mothers. Also, some ailments tend to have high loadings, such as cardiovascular diseases and endocrine diseases (e.g., diabetes). Less salient is the loading of varicose veins. When thinking about health in terms of lifestyle diseases and behaviour such as lack of exercise and unhealthy diet, these patterns are meaningful. Unsurprisingly, it is also seen that being in the upper two fifths of the doctor visits distribution has loadings similar to the hospital ailments while the loadings on having less than average doctor visits are negative (i.e., signs of being healthy).

With a health measure in hand for all individuals it can now be tested how this variable relates to events that arguably signal objective health problems. It would probably be a source of concern if the measure had no relation to, say, mortality. For parents I show how the health measure predicts mortality. For children the health measure is shown to predict uptake of disability pension and mortality. Figure 4 shows the 4-year mortality rate for parents across the health distribution split by parent gender. So, this is the observed mortality for four years after the health observation window, i.e., it will be the mortality rate in the years where they are between 70 and 74 years old. As the youngest parent birth cohort is 1940 and the data ends in 2014 this is the maximum window in which

---

[30]'pca' in Stata 14.

[31]Loosely referred to as 'the elbow point'.

[32]The first component explains 13-14 percent of the overall variability depending on the type of individual, with all component loadings being significant at conventional levels (results available upon request). This explained variability is probably on the low side relative to other applications of this method.

they can be followed. It is seen that men at the very bottom of the health distribution have a mortality rate of 35 percent, while this figure drops to around 7 percent for men in the very top. For women, there is the same tendency although, as expected, the absolute rates are lower. Thus, the measure does indeed seem to capture some health signal, which is reassuring. Figure 5 shows the corresponding graph for the child generation. Here, the health measure is computed based on the years 36-46 as opposed to 36-50 in the main estimation. This is a necessary implementation as the years 46-50 are needed for observing mortality and disability pension outcomes. Figure 5 shows predicted mortality to be a somewhat decreasing function of health rank, although the picture is less clear than for parents. Also, in particular for males, it seems that the signal is strongest in the bottom fifth. As mortality rates for these birth cohorts, 1960-1963, are rather low it is a demanding outcome to predict. Figure 6 shows the corresponding figure for children for another outcome, disability pension. This is an outcome with a higher incidence and also plausibly serves the purpose of validating the health measure; that is, the health measure gets more credibility if it can be clearly anchored to health-related "objective" events. Figure 6 shows a clear downward trend in disability rate all the way down the health percentiles. For example, women in the bottom of the distribution have a 15 percent chance of taking up disability pension at some point in this four year window, for women at the median it is around 2 percent while it is essentially zero for the most healthy individuals.

Figure 7, Figure 8 and Figure 9 disregard hospital stays and show the relation between objective health events and quintiles of doctor visits. These figures are included to demonstrate that while doctor visits may be theoretically conjectured to capture other aspects than health status (as opposed to hospital stays) it turns out empirically that the least healthy also display the most doctor visits and vice versa. In Figure 7 it is seen that those fathers with the fewest doctor visits (i.e., those in the first quintile) have a 4-year mortality rate of 7 percent in contrast to 17 percent for those with the most visits (i.e., those in quintile five). For mothers, there is the same gradient, albeit less steep. The figure can be contrasted to Figure 4 who showed the relation between the health rank extracted from principal components, which included hospital visits *and* doctor visits. Figure 4 qualitatively showed the same gradient, although the predictive power in relation to mortality was stronger. Hence, utilising doctor visits and hospital visits in unison gives a more powerful health measure than doctor visits alone, which however exhibits the same relation to objective health outcomes on its own, although to a lesser extent. Figure 8 shows that for individuals in the offspring generation, uptake of disability pension correlates positively with quintile of doctor visits. For sons the probability of transitioning to disability is virtually zero for the first quintile and nearly five percent for the fifth quintile. In this case, the gradient is even clearer for women (i.e., daughters). This can be contrasted with Figure 6 which showed similar, but stronger relation to disability when the health measure consisted of both types of health care utilisation. Figure 9 shows the relation to child mortality and quintile of doctor visits and is primarily included for completeness. As commented upon in relation to Figure 5, mortality is demanding to predict for individuals around age 50 as it has a very low prevalence. Hence, Figure 9 shows only modest differences in mortality according to utilisation of doctor visits, most of them

statistically insignificant. However, at at the very least, being in the quintile with the highest tendency to visit the doctor is associated with a markedly higher mortality (for sons, around 1.5 percent versus less than 1 percent for the other groups).

Overall, as expected, being in the lower health percentiles (i.e., being less healthy) is associated with markedly increased probability of being deceased, or on disability pension, over the next four years. The same holds when doing the analysis exclusively for doctor visits. Thus, it is certainly not the case that the main health measure picks up health the wrong way due to high-SES types constantly seeking doctor advice, and generally, *investing* more in their health this way. At the very least, if this is happening to some degree it is not driving the results.

5   Results

This section presents results. The first subsection is concerned with the parent-child models while the second subsection presents results from the sibling models. All results are net of cohort fixed effects and results are carried out separately by gender. I have experimented with specifications also controlling for birth order and the age difference between offspring and parents but found no substantial differences.

5.1   Parent-Child Associations

5.1.1   Main Results

Table 2 shows parent-child associations in health. Panel A contains the rank-rank slopes, i.e., the specification where each individual is assigned a percentile rank from the predicted health from the principal components procedure. Panel B shows the intergenerational correlations, where individuals are not ranked but where the estimate is standardized according to equation (2). Recall that the sample is based on individuals born 1960-1963 with identifiable parents from birth cohorts 1930-1940. Lastly, note that panel A of the tables, qua the rank-rank correlations, are directly comparable to recent work in the literature, Chetty et al. (2014); Landersø and Heckman (2017) for income and earnings, and Halliday et al. (2018) for health).

Table 2 shows that rank-rank slopes for health are in the range 0.11-0.15. Thus, for an increase in parent health of 10 percentiles, child health is predicted to increase by 1.1-1.5 percentiles depending on the particular combination of parent and child. It is clear that the strongest transmission seems to be the combinations where the parent is the mother and in particular between mother and daughter. Panel B shows very similar estimates so in this case rank-rank slopes and intergenerational correlations yield identical results. The four combinations exhibit the pattern that links to the mother are stronger than links to the father, the difference being largest for daughters. This is consistent with studies considering specific health outcomes. Goode et al. (2008) find that a history of cardiovascular disease and unhealthy eating correlates only among mothers and daughters. Loureiro et al. (2010) find that smoking behaviour is more transmitted in same-sex parent-child links. Also, it is obvious that e.g,

smoking behaviour of fathers cannot have the same intrauterine effects on the child as smoking by the mothers (Smith (2009)). Moreover, it is still the case that women perform the most household work, including the preparation of food (Bond and Sales (2001)). Taken together, these studies imply the partly opposing tendencies that same-sex child-parent links display stronger transmission and that the health of the mother matters the most, uniformly across sons and daughters. The relative strength of these will likely matter across diseases so the results in Table 2 can be rationalized by both tendencies operating while the latter dominates when all health information is aggregated into a common measure, which is essentially the approach adopted here. That is, the "mother effect" exceeds the "same-sex effect" in the aggregate. Figure 10 shows binned scatter plots of the relation between parent and child rank. All parent-child configurations exhibit approximate linearity which is convenient since only two parameters, the slope from Table 2, and the intercept (not reported here), are necessary to characterize the association between generations. Figure 10 shows that a daughter with a mother in the bottom of the health distribution, i.e., percentile zero, is expected to end up at around percentile 43 herself. Had the mother instead been the healthiest in her generation, i.e., percentile 100, the daughter's predicted health rank would be around 57. This may seem negligible: it may seem like the importance of the birth lottery is minor. While this is true in some sense, Figure 11 gets into more detail across the distribution. Figure 11 is a heat map, showing the density of individuals from the child generation present at each location of the grid. Specifically, it shows the heat map for the father-son combination[33] which was associated with an intergenerational correlation and a rank-rank slope of around 0.11. Hence, the figure can be thought of as a continuous transition matrix. If the relation between child and parent health was entirely random the map would be uniformly coloured. This is obviously not the case: unsurprisingly, there is a "heated" diagonal. Also, there is increasing density in the south west and north east corner[34]. For example, having a father with maximum health rank leads to a probability of around 0.6 percent of the son having the lowest possible health rank, which can be contrasted with a probability of the son being maximally healthy of around 1.4 percent (the thought experiment involves going from the south east corner to the north east corner of Figure 11). Similarly, going from the south west corner to the north west corner reveals that a child of the most unhealthy father will have a probability of 1.6 percent of ending up as the most unhealthy child himself, while the chance of moving to the absolute top is 0.6 percent.

Nevertheless, benchmarking the estimates of Table 2 to the vast literature on intergenerational mobility paints a picture of comparably high mobility. Thus, the health status of parents is a relatively weak predictor of own health status. The most obvious comparison is Halliday et al. (2018) who

---

[33]For brevity, a heat map is only showed for this particular combination of parent and child. The primary purpose of the heat maps is to show how health seems to have a more uniform transmission across the distribution than income. All heat maps for health looked qualitatively similar.

[34]It is tempting to think of behavioural theories of how the transmission of status then must be higher in the tails of the distribution than in the middle. While this may be the case, the pattern of Figure 11 likely also reflects a mechanical effect: there is simply less "space" to move around for the children with parents located in the tails. When moving to income below a more compelling argument of differing transmission across the distribution can be made.

use the Panel Study of Income Dynamics (PSID) to compute rank-rank estimates similar to those of Table 2. Whether they measure health by self-reported health or by self-reported number of diagnoses, the slopes are located approximately in the range 0.21-0.29, with the link between mothers and daughters similarly being the strongest. Abstracting from possible problems of the estimates being incommensurable due to differences in methods and samples this would suggest that intergenerational health mobility in Denmark is twice of that of the US. This stands in contrast to the conclusion drawn from using birth weight as a health measure. In this case the elasticity in birth weight is estimated to be almost similar[35] (0.18 in Denmark and 0.20 in US (compare Table 3 in Currie and Moretti (2007) with Table 2 in Kreiner and Sievertsen (2018)). In addition to comparing the mobility estimate to other health estimates from different countries, it can also be compared to transmission of other outcomes in a Danish setting. For years of education, an intergenerational correlation of 0.35 is found (Andrade and Thomsen (2018)); For wealth a rank-rank correlation of 0.25 Boserup et al. (2016); For income an intergenerational correlation of 0.20 and rank-rank slope of 0.27 (Landersø and Heckman (2017); for income, see also below). These studies are not identical to this one in terms of cohort selection, data restrictions and whether estimates are presented by gender. Nevertheless, a picture emerges of health being a dimension subject to less transmission of status than other outcomes, in particular for males.

5.1.2 A Brief History of Income (and Health)

Next, intergenerational associations for income are presented. Recall that the data, and specifically the estimation sample, is not set up for optimally measuring income mobility[36]. Nevertheless, it seems to add credibility to present results from the same sample as the health estimations are based on instead of exclusively relying on comparisons to other Danish studies (e.g., Landersø and Heckman (2017)) or comparable settings as Sweden (e.g., Björklund and Jäntti (2012)). Income serves as a natural benchmark partly because much of the previous literature has emphasized this measure. Table 3 shows parent-child associations in income, again rank-rank slopes in panel A and intergenerational correlations in panel B. Due to historical issues with respect to female labour force participation columns 1 and 3 with the father as the parent will receive the most attention. It is seen that the rank-rank slopes are 0.27 and 0.22 for sons and daughters, respectively. Thus, having a father at the bottom of the income distribution compared to at the top translates to a difference of 27 percentiles for the income distribution of sons. Again, it should be stressed that this is a purely descriptive estimate. Even if it was feasible to design an intervention that moved parents in this fashion (or, say, only 10 percentiles), it would not necessarily *cause* the child generation to have higher incomes themselves. Nothing in

---

[35]As the dispersion in birth weight probably does not vary a lot between generations the intergenerational elasticity should be close to the intergenerational correlation.

[36]The income of parents is measured at age 50-55, which is unlikely to be the optimal age at which to proxy lifetime income. Also, in terms of the birth cohorts included and trends in labour force participation, the income status of women born in the 1930s may be a suboptimal measure of social status: for example, very low income may signal having a very well-off spouse to a higher degree than it does today.

Table 2 or Table 3 rules out that the observed correlations are exclusively due to transmission of genes, intelligence, time preferences or a myriad of other factors acting in complex unison. Whatever the nature of the underlying forces may be there is a stronger association from father to son than from father to daughter which is in line with the literature. The rank-rank slope of 0.27 for the father-son combination also aligns perfectly with the estimate of 0.27 from other studies using Danish data (See first column of Table A17 in Landersø and Heckman (2017)). Turning to Panel B it is seen that there are some differences between rank-rank slopes and intergenerational correlations. For the father-son channel it is 0.21 and it is 0.13 between fathers and daughters which is relatively similar to statistics for Denmark (Summarized in Figure 1[37] in Corak (2013) and in Table 2, column 1 of Landersø and Heckman (2017)). This may indicate that transmission of economic status varies in strength throughout the distribution in a way that is not the case for health, or, it may be a more benign issue of noisy data due to somewhat suboptimal age of parent measurement. In either case it shows that the functional form of the relationship parent and offspring status may affect mobility statistics differently. Fundamentally, the intergenerational correlation measures the linear dependence between two normally distributed variables while the rank-rank slope effectively discards the cardinal information and only measures the degree of monotonicity between the variables. Thus, the two measures may vary for a variety of reasons, including skewness in the data and nonlinear transmission patterns. No single measure is "better", it makes more sense to view the two parameters as complementary as they entail differing assumptions and measurement objectives[38]. In this setting it has the consequence that conclusions for women are different, depending on the specification. According to the rank-rank slope, women are more tightly linked to their father in income status than they are in health status (0.217 versus 0.111) but the intergenerational correlations yield a comparison of 0.130 versus 0.115.

Figure 13 shows a heat map for income ranks between fathers and sons, similarly to how Figure 11 showed this for health ranks. As argued above, parent-child associations seem to be stronger for income than health in our sample, although for women this conclusion hinges on the preferred intergenerational statistic, intergenerational correlations or rank-rank slopes. However, one thing is to compare persistence coefficients, a supplementary analysis involves gauging relations throughout the distribution. On the point of simple strength of the association, Figure 13 expectedly shows higher diagonal densities than Figure 11 showed for health. This is not surprising given that the rank-rank slope for father-sons are more than twice as high for income as for health. However, it is striking how the figures differ in terms of strengths in the south west and north east corners of the map. For income, there is strong transmission in the top: only the north east corner exhibits these high densities[39]. If

---

[37]Figure 1 actually shows the intergenerational elasticites, not the intergenerational correlations. However, I find only moderate differences between elasticites and intergenerational correlations (I do not report any elasticities in the regression tables). A further caveat is that the elasticity is measured based on earnings, not income.

[38]In Corak et al. (2014) it is noted that if parent income is given as $x = 1, 2, 3, \ldots$ and child income as $y = \exp(x)$, then the rank-rank slope will be 1 but the correlation between x and y will be around 0.25. The authors compare mobility levels of Canada, Sweden and the US and for the latter two it is found that the rank-rank correlation exceeds the intergenerational correlation as is the case in Table 3.

[39]This is consistent with the first graph of Figure 12 which shows a steepening of the slope in the top part of the father

a son has a father in the top of the distribution his chance of ending up there himself is 2.5 percent (on a baseline of 1 percent), while the corresponding figure is 0.5 percent for a child with a father in the bottom of the income distribution. The color differences of the heat maps suggest that health and income are distinct domains in which to analyse mobility over and above the mean transmission across the distribution: for income, there is a specific reproduction going on in the elite. This may be transmission of work ethics, connections or certain job types[40]. For health, it seems to be more difficult to pass on superior health to your children. There is persistence at play between parents and children but it is much more uniform. In the sibling correlations below I ignore these non-linearity aspects of transmission, although it is certainly an interesting extension for future research[41].

5.2 Sibling and Twin Correlations

This subsection presents results from the second, and arguably more nuanced, way of evaluating mobility in health outcomes. Results consist of the sibling and twin correlations, accompanied by a discussion of how they relate to the intergenerational associations and other findings from the literature.

For expository purposes the results from the sibling estimations are displayed in figures rather than in tables. Figure 14 shows various sibling correlations, ranging from ordinary siblings on the left, to twins in the middle, and to identical twins on the right. Note that twins in the middle panel consist both of monozygotic and dizygotic twins. Recall that non-twin siblings and twins can be observed in the data, but that zygosity is unknown. Thus, the two left panels of the figure compute observed patterns in the data while the right panel relies on certain identifying assumptions explained in the methodology section. Also, recall that the key rationale for this whole exercise is to obtain plausible bounds on the total importance of family background.

First, attention should be given to the "Brothers" and "Sisters" category, i.e., the sibling correlations. These are the types of sibling correlations most studies evaluate as they do not require special twin databases or large sample sizes commonly found in register data. It can be seen that sibling correlations in health are given as 0.14 for brothers and 0.20 for sisters. In other words, 14 percent of the variance of male health outcomes can be explained by family background while this figure is 20 percent for women. Several things are worth noting here. First, it indicates that the intergenerational correlations (and the rank-rank slopes) capture persistence of health quite poorly. In the case of women, the highest intergenerational correlation coefficient was found to be 0.145 which implies an R squared of 0.021. The parent-child associations explain (0.021/0.20)*100= 11 percent of what

---

income distribution.

[40]I am not the first note the strong transmission at the top of the income or earnings distribution. Björklund et al. (2012) report very high transmission levels in the top of the Swedish income distribution. In terms of mechanisms such as parental networks Bingley et al. (2011) show that the fraction of children employed at a former employer of their father grows rapidly in the top of the child income distribution.

[41]Evaluating mobility throughout the distribution is very rare for the part of the literature employing sibling correlations.

tends to make sisters similar in terms of health status. For brothers (comparing with the highest intergenerational coefficient, that is, the mother-son association), making a similar calculation yields a figure of 11 percent. Recalling equation (3) this indicates that the second term, the "other factors", is quite important. It is in this sense that Björklund and Jäntti (2012) conclude that studying parent-child associations is like studying "the tip of the iceberg" if one seeks to understand why status is perpetuated in families. In the outcomes they study, they find the "tips of the icebergs" to be between 4 percent (non-cognitive skills) and 43 percent (height) of the sibling correlation. For the US Solon (1999) reports that forty percent of sibling correlations in earnings are explained by parental earnings. Hence, my conclusions for health indicate that, relatively speaking, a large fraction of the iceberg is beneath the surface. For example, it may be the case that sibling resemblance in health is the result of similar time preferences or educational levels rather than healthy parents somehow passing on health capital.

A second perspective on the estimated sibling correlations is how they compare with similar correlations in other outcomes. Bredtmann and Smith (2018) finds, in a Danish context, sibling correlations in years of education to be around 0.4 for both genders. This is certainly higher than my corresponding figures of 0.14 and 0.20 for health. The same holds true for a range of economic outcomes in the US, where Mazumder (2008) estimate sibling correlations to be around 0.5. Björklund and Jäntti (2012) report correlations of 0.3 for non-cognitive skills for males, 0.2 for brother earnings and 0.16 for sister earnings. So, only when comparing to this last study, health is not necessarily a domain of more mobility for females.

With respect to income in the current setting, Figure 15 shows the corresponding estimates for income. This figure will be briefly discussed before moving on to the second and third columns/blocks of figures 14-15 (i.e., twin correlations). Figure 15 shows sibling correlations in income to be around 0.27 for brothers and 0.18 for sisters. As with health, it confirms the general impression from the intergenerational associations that sibling correlations convey more information than parent-child associations. Moreover, it reinforces a story of gender differences: For men, family background seems to be substantially more important for income than for health (sibling correlations of 0.27 versus 0.14), while for women both correlations are around 0.19. Thus, from this perspective health is a domain with more equality of opportunity than found for income but this conclusion only holds for men; for women, the importance of family background is of similar magnitude across domains. Or, at least, the lower bound on the total importance of family background is similar.

Next, the remaining (twin) sibling correlations of Figure 14 and Figure 15 are considered. Unsurprisingly, Figure 14 shows that sibling correlations in health increase when their degree of similarity goes up (i.e., when moving left to right in the figure, conditional on gender). Their increasing similarity can be organized under headings as increased genetic overlap, increased interaction and increased shared exposure to various "shocks" to the family. The aim here is not so disentangle these forces empirically but to group them collectively as factors beyond the control of the individual and hence to take them into account when gauging equality of opportunity. Unfortunately, it is also seen that

the precision of the MZ-estimates (i.e., identical twins) is low. In the section on robustness checks I will show that employing a different, larger sample leads to more precisely estimated parameters for the health outcome, and the point estimates themselves not changing dramatically. Hence, these twin correlations can be taken at face value with at least some credence. They indicate that male identical twins have a health status correlation of 0.34 and female identical twins have a correlation of 0.38. An interesting first comparison to these figures is Herskind et al. (1996) who estimate identical twin correlations in longevity using the Danish Twin Register. They find correlations of 0.19-0.25 for males and 0.18-0.31 for females. Putting aside considerations of confidence intervals the estimates from Figure 14 are then relatively high[42]. However, comparing to (other) outcomes from other literature paint a different picture. For example, Björklund and Jäntti (2012) find identical twin correlations of around 0.75 for male educational attainment and male earnings for Sweden, and Nicoletti and Rabe (2013) report correlations of 0.9 in educational outcomes for both male and female identical twins in the UK. Revisiting Figure 15 it can be seen how twins correlate in income in the current sample. As for health in Figure 14, precision in Figure 15 is also less than ideal. However, it is somewhat reassuring that the point estimates are not too different from Björklund and Jäntti (2012), although it should be noted that income measures are not identical. Figure 15 indicates that identical male twins correlate around 0.62 in income and female identical twins around 0.42. This suggests that the upper bound on the importance of health and income is strikingly similar for women (0.38 versus 0.42), while markedly different for males (0.34 versus 0.62).

5.3   Conclusion on Results

Ultimately, the bounding exercise of Figure 14 of of siblings and identical twin correlations leads to the following tentative conclusion. For men, between 14 and 34 percent of health status is determined by family background in a broad sense. For women the corresponding bounds are 20 and 38 percent. In my summary of the literature on traditional parent-child associations, supplemented with estimates from the previous subsection, paints a picture of health, for males, overall being a domain of relatively high mobility[43]. At the very least, it seems warranted to conclude that persistence in outcomes between generations is higher in outcomes as education, earnings and income than for health[44]. In other words, offspring outcomes seem to be less determined by circumstances beyond their control (such as parent status) when health is the outcome compared to the case of more traditional labour market outcomes.

---

[42]This is consistent with the observation that inequality in health while alive is larger than inequality in longevity. In other words, inequality in morbidities is high but for a variety of reasons, individuals live with their ailments for relatively longer time before morbidity becomes fatal. Hence, health while alive, as is the outcome measure in the current setting, is arguably a more interesting avenue to gauge dependence on family background rather than simple longevity.

[43]This may run counter to a common conjecture that health "is largely genetically determined". While genetics undoubtedly plays a big role, the literature seems to suggest that genetic factors play an even bigger role for outcomes as education and income (although this conclusion may be too strong given the difficulty in differentiating genetic endowment from environmental factors facing the individual growing up).

[44]Halliday et al. (2018) conjecture that "[...] One explanation for this pattern is that policies and institutions in the U.S. may be much more effective at breaking intergenerational linkages in health than in labor market outcomes."

For women, mobility levels in health are lower than for educational outcomes but seemingly on the same footing as income and earnings. In addition, what is common for both outcomes studied here, is that sibling correlations provide a much fuller picture than intergenerational associations. In fact, only 11 percent of the sibling resemblance can be attributed to their parents having identical health.

6   Discussion and Conclusion

Studies of intergenerational mobility and equality of opportunity are subject to increasing interest as inequality seems to be on the rise in several domains and uniformly across most OECD countries. This paper is part of a literature that attempts to broaden the set of outcomes studied in contemporary society. Using Danish administrative data I present the first comprehensive estimates of general health mobility outside the US. On many accounts I replicate and confirm earlier findings in the literature. Like Halliday et al. (2018) I find that health mobility is higher than income mobility when relying on rank-rank slopes. Also, the pattern of transmission is quite different: for income there is a very concentrated transmission in the top of the distribution while it is much more uniformly distributed for health. For intergenerational correlations transmission of health is also weaker than for income, although the difference is very small for women.

In line with other literature I confirm that parent-child regressions can be seen as the tip of the iceberg when compared to sibling correlations, which is a broader measure of equality of opportunity (Solon (1999); Mazumder (2008); Björklund and Jäntti (2012); Bredtmann and Smith (2018)). As a methodological point in its own right this is yet another indication that traditional parent-child associations far from paint the full picture. For health I find that eleven percent of the sibling correlation can be explained by parental health. This may be a hint to policy makers that if they are worried about reproduction of health status in families, parental health is not necessarily the only avenue in which to consider interventions. Rather, it is the broad set of circumstances that are shared in families. Examples include educational attainment and economic status of the parents as well as values and preferences imparted on the offspring generation such as risk aversion and time preferences. It remains an important topic for future research to probe into how these variables affect health status. Besides the sheer amount of health information utilised another novel feature of this paper is the bounding analysis, where I assert that the total importance of family background is bounded between sibling correlations and identical twin correlations. Using this approach I find that 13-38 percent of the total variation in health can be explained by family background factors. While this is less than for other relevant outcomes such as income, earnings and educational attainment, it suggests that even in comparably egalitarian societies with universal health insurance, health status is still reproduced within families.

Accepting the premise that family background broadly understood is a factor beyond the control of the individual this suggests the absence of equality of opportunity in health for both genders, more so for females. Such figures should enter a debate on health policy prevalent in many arenas. It touches

on an even more general debate concerning to what extent individuals should be held accountable for their outcomes. For example, it has been debated whether bariatric surgery (weight loss surgery) should be paid for by the patients themselves, since it is primarily their behaviour, and hence their own responsibility, that there is need for an expensive operation (Lund et al. (2011)). Also, in a paper used as guidance by the World Health Organization, Whitehead (1991) lists various criteria for when health inequities can be considered fair, one of them being the degree of 'choice' involved in a certain health outcome. For example, lung cancer can more intuitively be attributed to a choice on the part of the individual to smoke than can, say, a birth defect. Without further assumptions and preferences, sibling correlations, and their magnitudes, in themselves do not give clear directions for policy, but they arguably inform the discussion. As Nielsen and Andersen (2014) argue it can still make sense to hold the obese (or generally, unhealthy) individuals *cost*-responsible for their health outcomes, although e.g., large sibling correlations may be an indication that their non-self-chosen family background may have played an instrumental role in bringing about their situation. In other words, a compromise between efficiency considerations and fairness considerations may have to be found. Thus, descriptive studies of social mobility by e.g., sibling correlation do not provide a clear-cut recipe for a further course of action. Instead, they avoid the near-impossible task of distinguishing certain diseases or behaviours as the result of either effort or circumstances, but pragmatically gauge the importance of the birth lottery by focusing on the one aspect that individuals never choose themselves: their parents, and the implied endowment of genes and (early) environment. These mobility estimates, along with fairness and efficiency considerations, and estimates of the responsiveness of individual behaviour then form a solid foundation on which to design polices and institutions.

The final policy conclusion to draw here is very tentative given the infancy of the literature with its associated difficulties in cross-country comparisons. Nevertheless, it is striking to note that most of the estimates from the literature on mobility in education and income are consistent with a rule of thumb saying that persistence in social status in Scandinavia is between one half and two thirds of what is found for the US[45]. It seems plausible that part of this difference can be attributed to varying levels of income inequality, economic mobility and availability of universal health insurance versus employer provided insurance[46]. This does not imply that other countries necessarily would attain similar mobility rates from one day to the other if they implemented policies from Scandinavia; however, on the other hand it may interpreted as an indicative pattern. Disentangling the many complex ways in which institutions, policies and societal norms play together with inequality and mobility remains an important, and formidable, task for social scientists in the future.

---

[45]When comparing with Halliday et al. (2018) my results for the rank-rank slopes in health fit neatly into this simple mathematical rule (0.21-0.29 versus 0.11-0.14). For rank-rank slopes in income I place them in the range 0.22-0.27 compared to at least 0.4 for the US (Mazumder (2015). For sibling correlations in income, I find brother correlations of 0.27 versus 0.5 for the US (Mazumder (2008)).

[46]Halliday et al. (2018) report lower rank-rank slopes (i.e., higher mobility) for those individuals with health insurance. Of course, having health insurance is likely not an exogenous variable but at least this finding is suggestive about the role of policies in fostering mobility.

7 Appendix: Robustness Checks

7.1 Life Cycle Bias and Attenuation Bias

This subsection evaluates measurement issues regarding the estimated intergenerational associations in health. Results are presented for rank-rank slopes only (identical results are obtained for intergenerational correlations). The literature on intergenerational mobility has been concerned with various biases in estimating the key parameters. *Attenuation bias* is an example of the well-known errors-in-variables downward bias in an explanatory variable. For instance, a single-year measure of earnings may be a noisy indicator of long-run parental earnings due to transitory shocks to income. Therefore, it has become commonplace to use multi-year averages of the outcome variable to increase the strength of the "signal" relative to the "noise" in the measured variable. For income, Mazumder (2005) recommends using at least 9 years of parent income to make the attenuation bias negligible. It may reasonably be expected that many of the considerations of earnings or income carry over to the case of health. An individual may be hit by a transitory shock to health as well as to labour market productivity. Figure 16 shows the estimated rank-rank slopes as a function of included years of parent health. As expected, the estimates initially increase when incrementing from a single year of health measurement. It seems to plateau at around six or seven years, although this is less clear for the mother-son combination. However, it seems overall the estimated rank-rank slopes pass the test of attenuation bias when including 10 years of parental health measurements. The literature performing this exercise for health mobility is sparse but it can be noted that Halliday et al. (2018) also finds that 10 years of health measurement is sufficient to stabilize the estimated parameters.

The other measurement concern of researchers in the field, *life-cycle bias*, typically concerns the age interval at which outcomes are measured over, conditional on the length of the window. The outcomes of the offspring generation are typically the principal concern although conceptually life cycle considerations should apply to both generations. However, for many datasets measurements at father prime earnings ages have been easier to obtain than, say, child's earnings at age 40. Solon (2002), among others, note that for a *fixed length* of earnings measurements, intergenerational elasticities/correlations seem to be higher when based on earnings from children well into their 30s rather than their late 20s. The reason is heterogeneous age-income profiles with individuals destined for higher incomes experiencing higher earnings growth which, expressed in the language of linear regression analysis, leads to mean-reverting measurement error in the dependent variable. There is wide agreement in the literature that income-related lifecycle bias is minimized when outcomes are measured around age 40 (Grawe (2006); Mazumder (2005)). Far less research has been done with respect to health, so we find ourselves in somewhat uncharted territory. Intuitively, one would conjecture that optimal health measurements should be done later in life for health than for income as the health status is more truly revealed later than age 40. That is, chronic diseases and severe health problems tend to arise later than age 40. Halliday et al. (2018) find some evidence that rank-rank slopes in health are higher when outcomes are measured later in life. However, standard errors are generally too big to draw firm

conclusions. In the current setting, I can observe children up till age 50. Thus, the window where I can perform life cycle bias checks is more limited which is simply a consequence of parent-child linkage in the Danish registers. However, it can be seen how intergenerational associations vary when health measurement is split in the current observation window. In the main specification health is measured over the age interval 36-50. Figure 17 shows how rank-rank slopes vary when health is measured over ages 36-43 vis-á-vis ages 43-50. For convenience, estimates from the main specification is placed as the left column in each "block". It would be clear signs of lifecycle biases if the right column of the four panels were markedly larger than the middle column, i.e., if health measurements from later in life gave rise to significantly higher persistence coefficients. Although three out of four specifications exhibit higher correlations for the later measurement (i.e., at ages 43-50) the differences seem negligible both from a statistical and a practical perspective.

7.2 Changing Sample Restrictions

The results obtained so far are all based on the main estimation sample. This sample, as defined in the data section, is subject to a number of restrictions. The most comprehensive of these is probably the restriction that all individuals from the offspring generation must have a father and a mother who are both alive at age 70. This restriction was necessary in part to be able to observe parent health for ten years (from age 60 to age 70). Also, it is convenient (and persuasive) to use the same basic sample for all the main estimations. However, the parent-child estimations performed are not contingent on having *both* parents available, and the sibling correlations obviously does not require parents to be alive till age 70, only for the parents to be alive until age 1986 such that the link can be observed. In the following I will loosen these restrictions. The consequences are twofold: on the one hand, it gives a larger sample size (increase from 71,460 113,654), but on the other hand it changes the sample from estimation to estimation. One concern one might have about the restrictions imposed on the main sample is that they lead to a positively selected sample. After all, it is certainly not random in which families both parents are alive at age 70. Furthermore, policy makers might be concerned about status perpetuation precisely in the *lower* part of the distribution. It is not immediately clear in which way leaving out, relatively speaking, disadvantaged individuals would affect the results. If transmission of health is stronger in the lower part of the distribution it should be expected that the importance of family background for health outcomes is actually stronger than what have been found in the analysis so far[47].

7.2.1 Parent-Child Associations

First, it will be tested how the intergenerational correlations react to loosening parent survival restrictions. Now, only the relevant pairwise restrictions are imposed. For example, to calculate the

---

[47]Even though the parent-child associations showed some "symmetry" concerning transmission across the distribution it does not necessarily carry over to sibling correlations.

father-son correlation it is necessary to impose father survival till age 70 while no restrictions are imposed on the mother. Figure 18 shows how the four different intergenerational correlations change[48]. Or, rather, it shows how they do *not* change, which is reassuring. In each of the four blocks, the estimate from the main estimation sample is placed to the left, while the estimate from the larger, unrestricted sample is placed to the right. All four estimates closely resemble the alternate specification. There are several explanations for this. One is that mobility is similar across the distribution (i.e., similar for relatively disadvantaged individuals compared to the ones with parents who are both alive at age 70). Another explanation is that the requirement that both parents are alive at age 70 is not that severe. It shows in the data that the main sample is slightly positively selected as the average health rank for the fully restricted sample is 51.81, and around 51.2 for both samples where only survival of one parent is conditioned on (all relative to being 50.5 in the unrestricted sample). In any case, estimated parent-child associations do not seem to fundamentally differ due to (very) selective survival of parents.

7.2.2   Sibling Correlations

This subsection repeats the above exercise for sibling correlations instead of for parent-child associations. As sibling correlations are interesting in themselves, their sensitivity is evaluated first before returning to the correlations between twins.

Figure 19 tests whether sibling correlations change for the above mentioned alternate sample restrictions. The left correlation of each block, "MAFA", is the estimate from the main estimation sample placed in the graph for comparison. Moving to the right in each block shows that happens when survival restrictions on parents are loosened. In terms of confidence intervals, the precision is increased when comparing the far left specification with the far right, with the two intermediates being in between. Concerning point estimates, the correlation to the far right, "NR", shows the estimates for the largest sample, where none of the parents need to be alive at age 70 for the child to be included in the sample. It is seen that estimates are very stable for women and quite stable for men. The increasing tendency from left to right in the left block of correlations is consistent with a story of more and more vulnerable individuals being included in the sample (and if those individuals are more dependent on family background). However, estimates are not statistically significant, and the differences in point estimates are also quite small.

7.2.3   Sibling and Twin Correlations

Next, we revisit the correlation between twins. Recall that the idea is to obtain an upper bound on the total impact of family background for various outcomes. Figure 14 showed that sibling correlations progressively increased when going from ordinary siblings to twins, ending with identical twins. The estimates for the identical twins were rather imprecise in this setting. The aim of this robustness

---

[48]Results are similar for rank-rank regressions.

check is twofold. First, to check whether point estimates are robust to an alternate sample (keeping the above considerations of a positively selected sample in mind). Second, if point estimates are not too different, it would be ideal to obtain more precise estimates, particularly for the identical twins. Figure 19 indicated that this was the case for siblings.

To maximise sample size, only the largest alternate sample is employed. Hence, no conditioning on parents is made (i.e., this is the "NR"-configuration from Figure 19). The results are displayed in Figure 20. Figure 20 is similar to Figure 14 except the estimations of the latter are carried out on a sample where parent survival is not conditioned on. Several findings emerge from the figure. First, point estimates are quite similar compared to Figure 14, although slightly higher for most specifications. Second, precision is improved although the confidence intervals of the identical twin correlations to the far right are still relatively wide. Björklund and Jäntti (2012), utilising a larger sample, also obtain somewhat wide confidence intervals in their similar estimations. Nevertheless, it does provide some confidence that different sample specifications yield nearly identical point estimates. The basic insight is the same: sister correlations are slightly higher than brother correlations and in a somewhat proportional fashion, they increase when the degree of similarity is increased, when moving to the right in the figure. Taking the point estimates as credible, the preliminary conclusion still remains that between 14 percent and 38 percent of health status is determined by family background. It seems that sister correlations are higher than brother correlations, although the difference is only statistically significant for the left block.

8 Figures and Tables

Figure 1: Overview of Cohort Selection and Data Sources

Parents
1930-1940

Children
1960-1963

1977-2014
Hospitalisations

1980-2014
Income

1986-2014
Mortality

1990-2014
GP visits
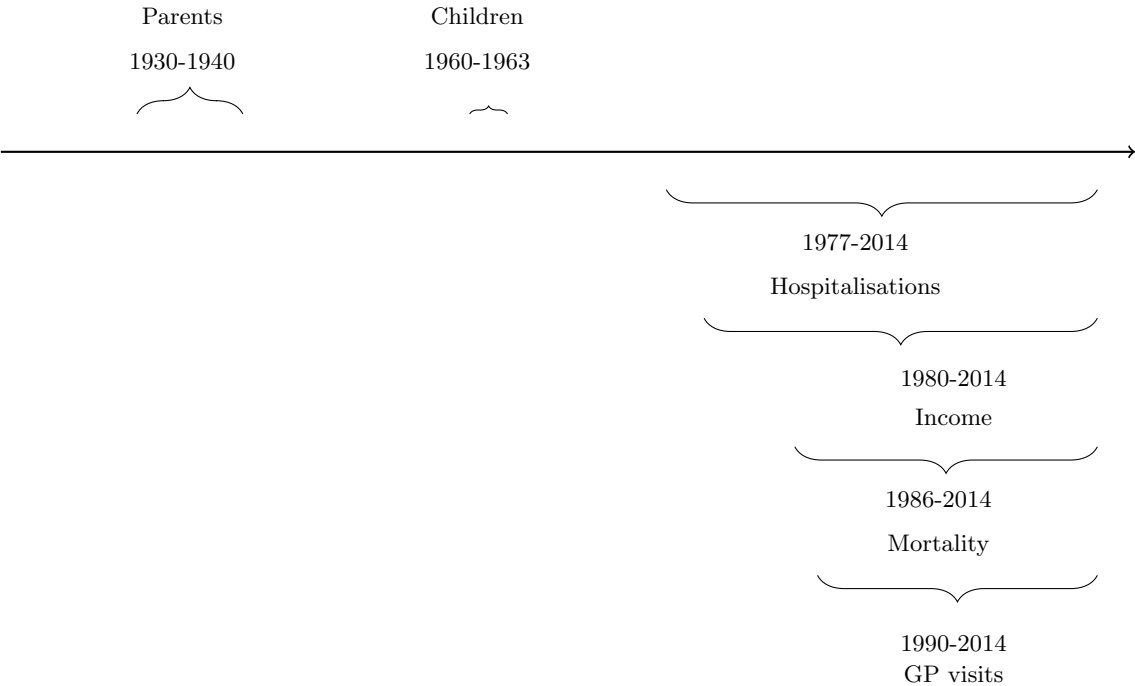
Table 1: Descriptive Statistics of Health Care Utilisation

	Sons	Daughters	Fathers	Mothers
Share with ailment				
Infectious Diseases	.03	.03	.03	.03
Lung Diseases	.09	.08	.15	.14
Nervous System	.1	.11	.14	.1
Cardiovascular Diseases	.09	.09	.29	.23
Arteries and Veins	.02	.03	.07	.05
Varicose Veins	.01	.04	.02	.04
Blood Disorders	.01	.04	.03	.03
Gastrointestinal Diseases	.22	.26	.29	.27
Rheumatic Diseases	.68	.61	.47	.55
Skin Diseases	.11	.13	.08	.08
Eye Diseases	.16	.1	.13	.15
Ear, Nose and Throat	.12	.12	.22	.15
Endocrine Diseases	.07	.12	.14	.16
Mental Illnesses	.05	.05	.03	.03
Mean GP Visits by Quintile				
GP Visits Q1	16	44	20	32
GP Visits Q2	43	95	53	78
GP Visits Q3	70	138	86	123
GP Visits Q4	108	197	132	181
GP Visits Q5	215	349	255	331
Observations	36,178	35,282	58,535	58,609

Notes: The upper part shows the share of individuals who have ever been in hospital care with the given ailment, over the years of observation (36-50 for the child generation and 60-70 for the parent generation). The lower part shows the mean general practitioner (GP) visits by quintile (i.e., doctor visits).

Figure 2: Scree plots
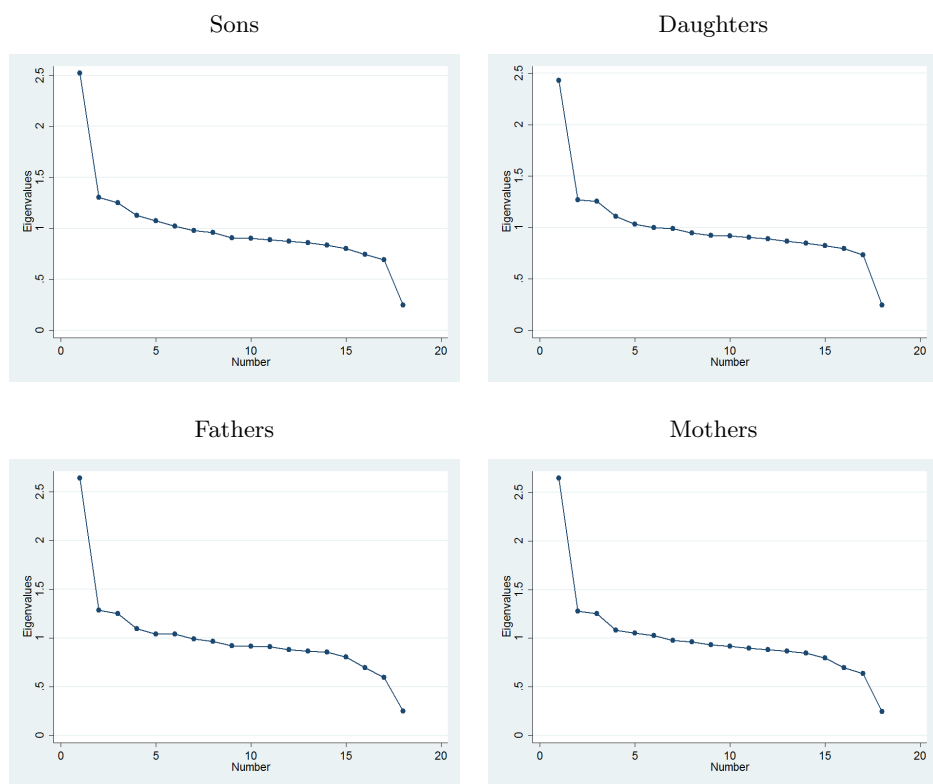
Notes: Scree plot for the principal components procedure for health by generation and sex.

Figure 3: Component Loadings
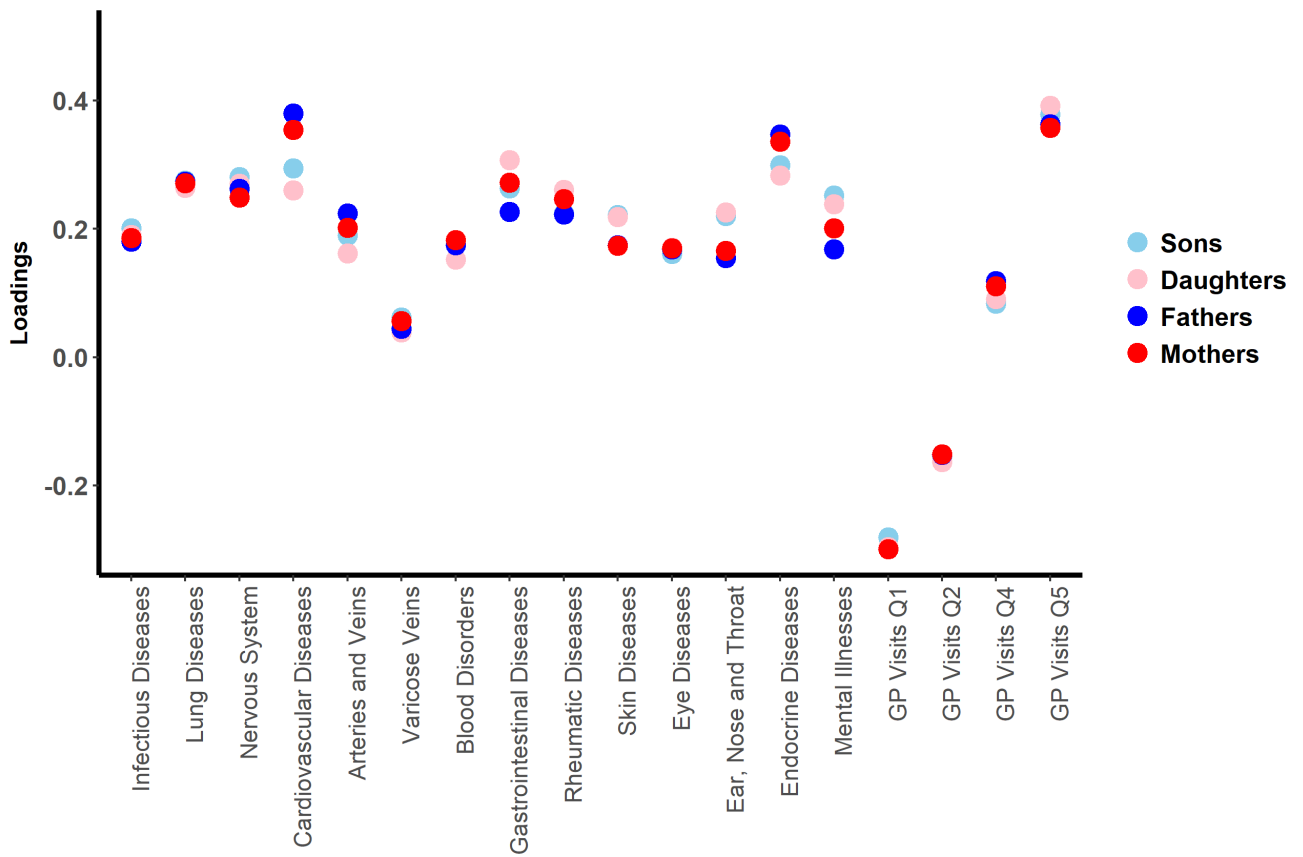
Notes: Component loadings from the principal components procedure by generation and sex.
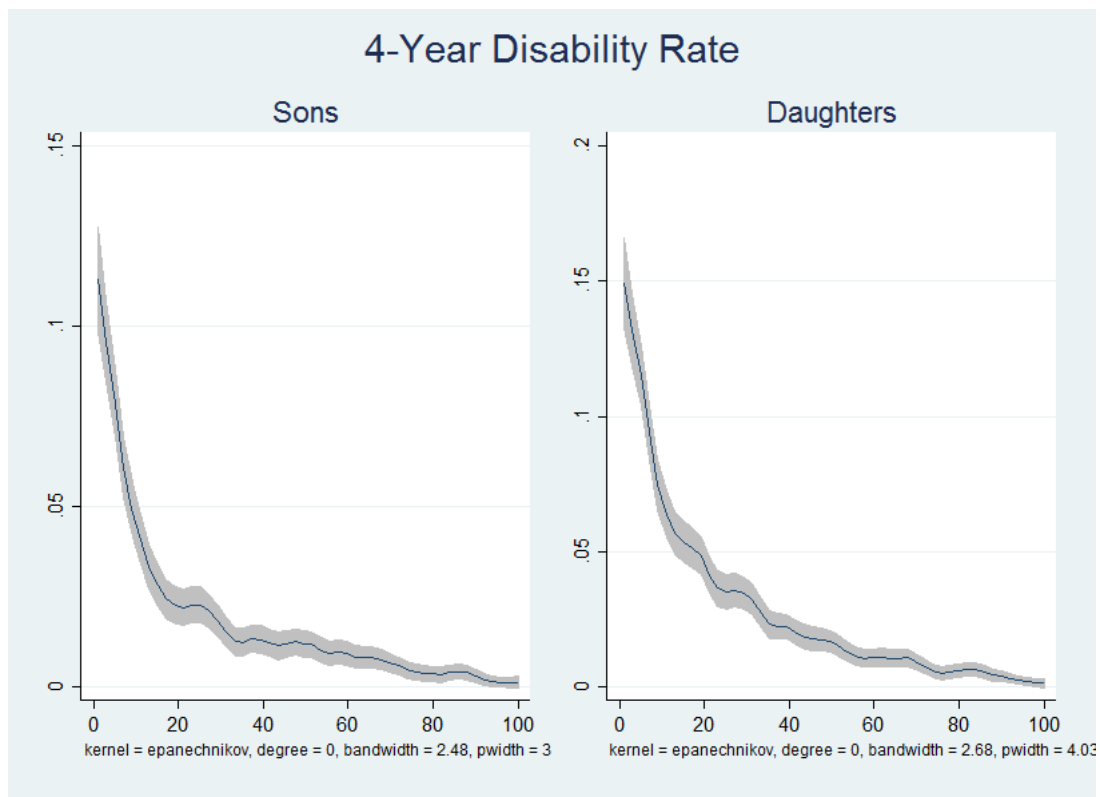
Figure 4: Parent Health and Mortality
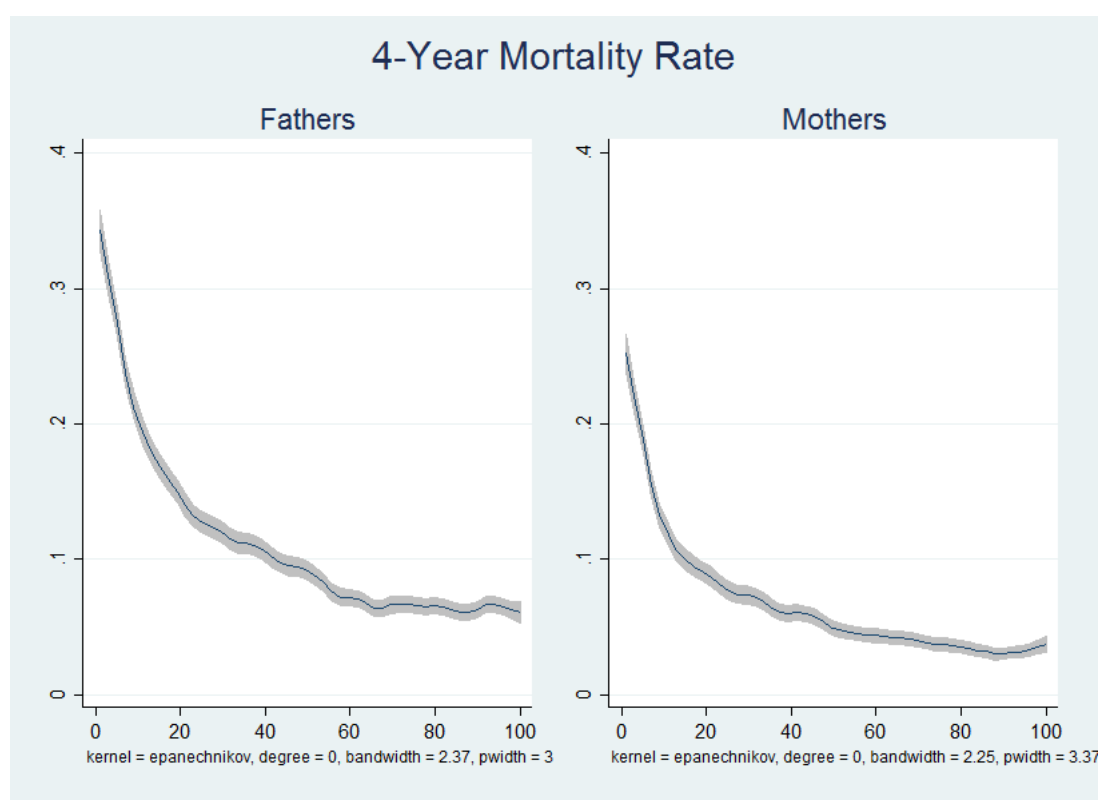
Notes: Nonparametric regression of parent mortality rate in age interval 70-74 on predicted health rank. Confidence bands are based on a 95 % Level.
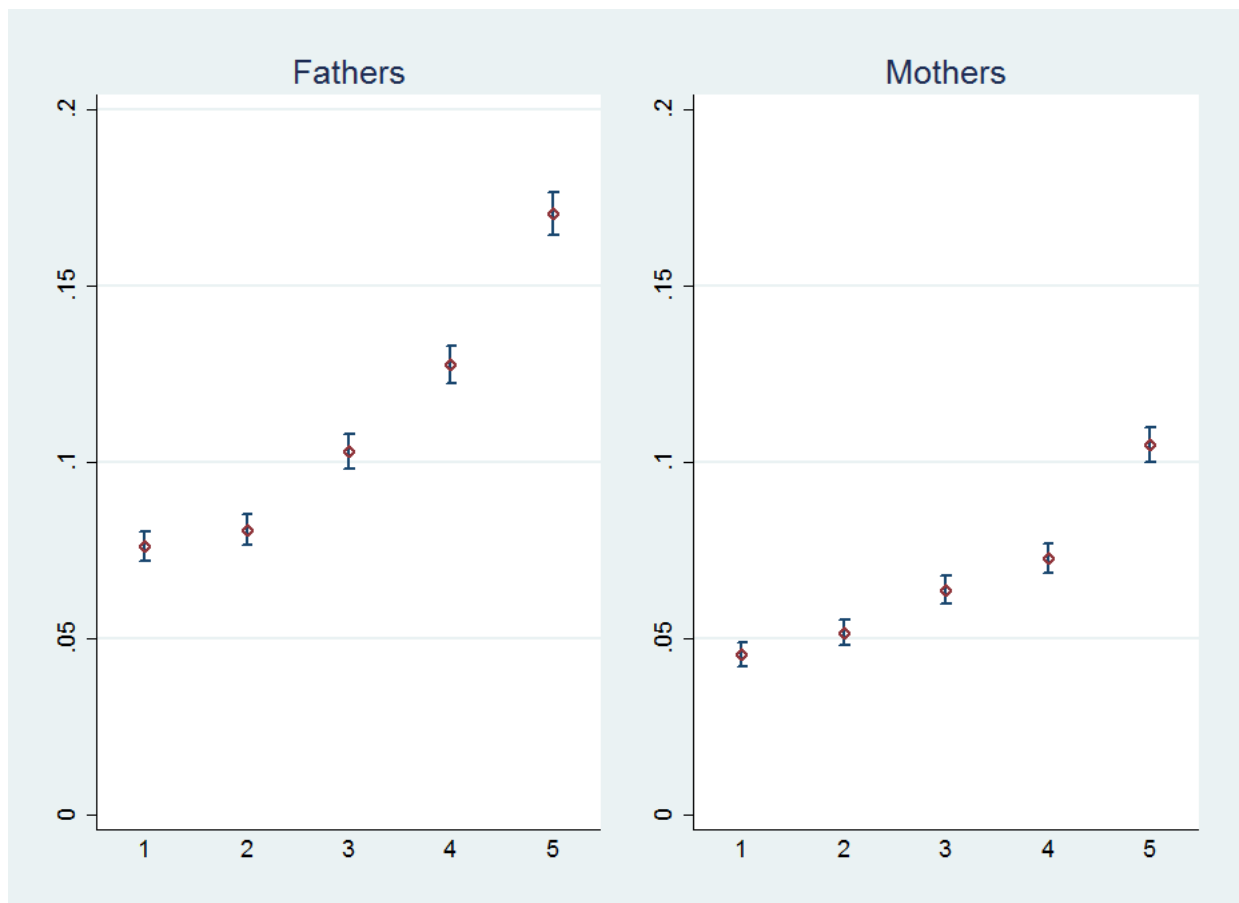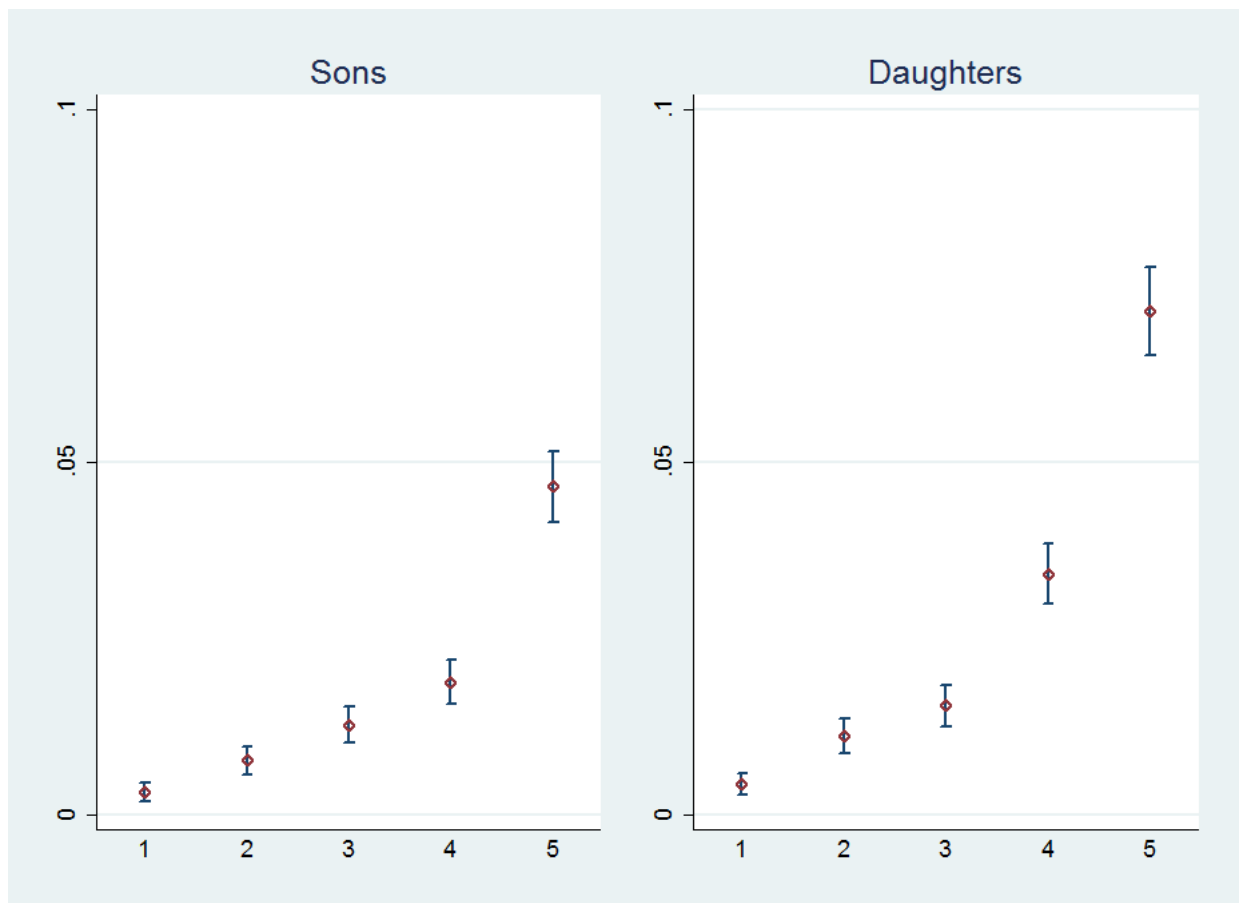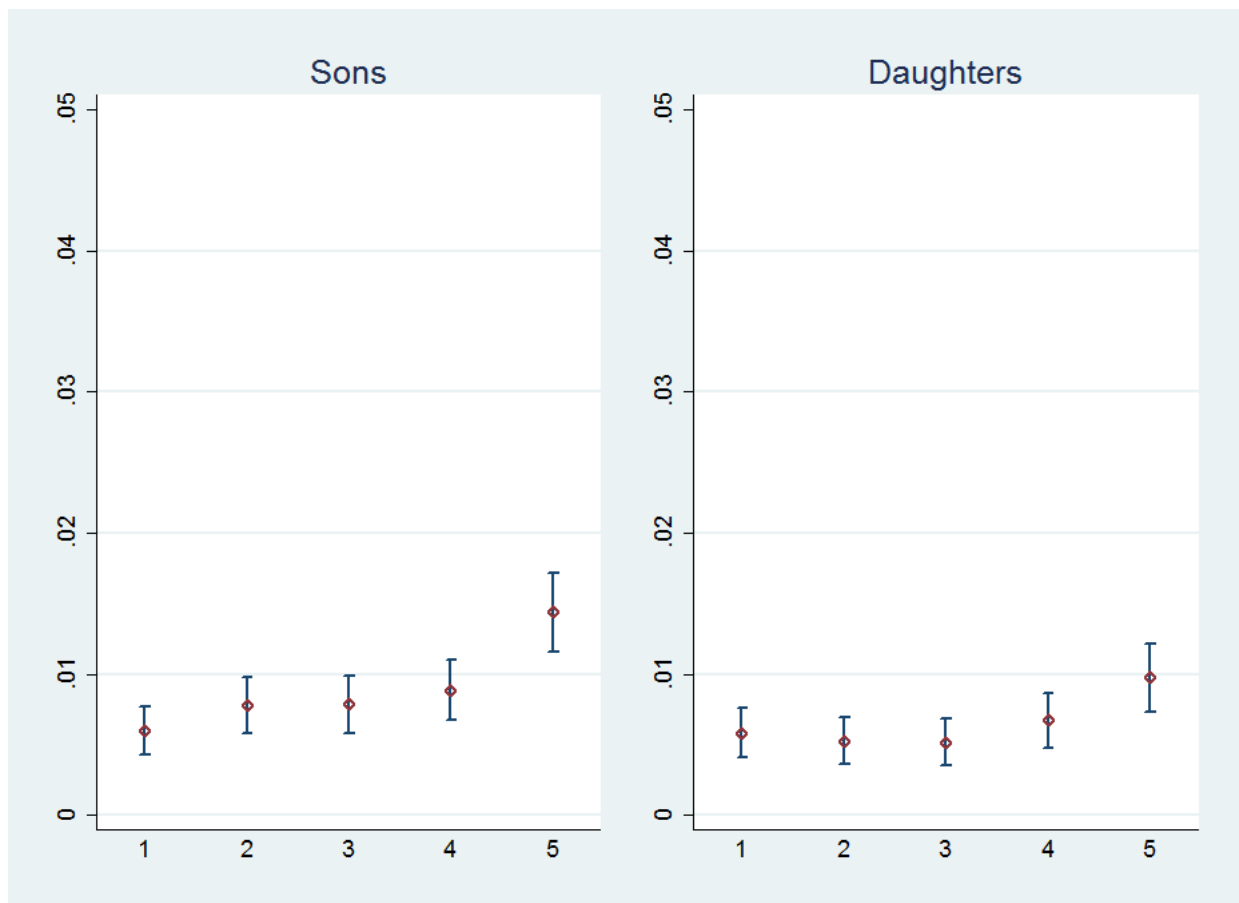
Figure 5: Child Health and Mortality
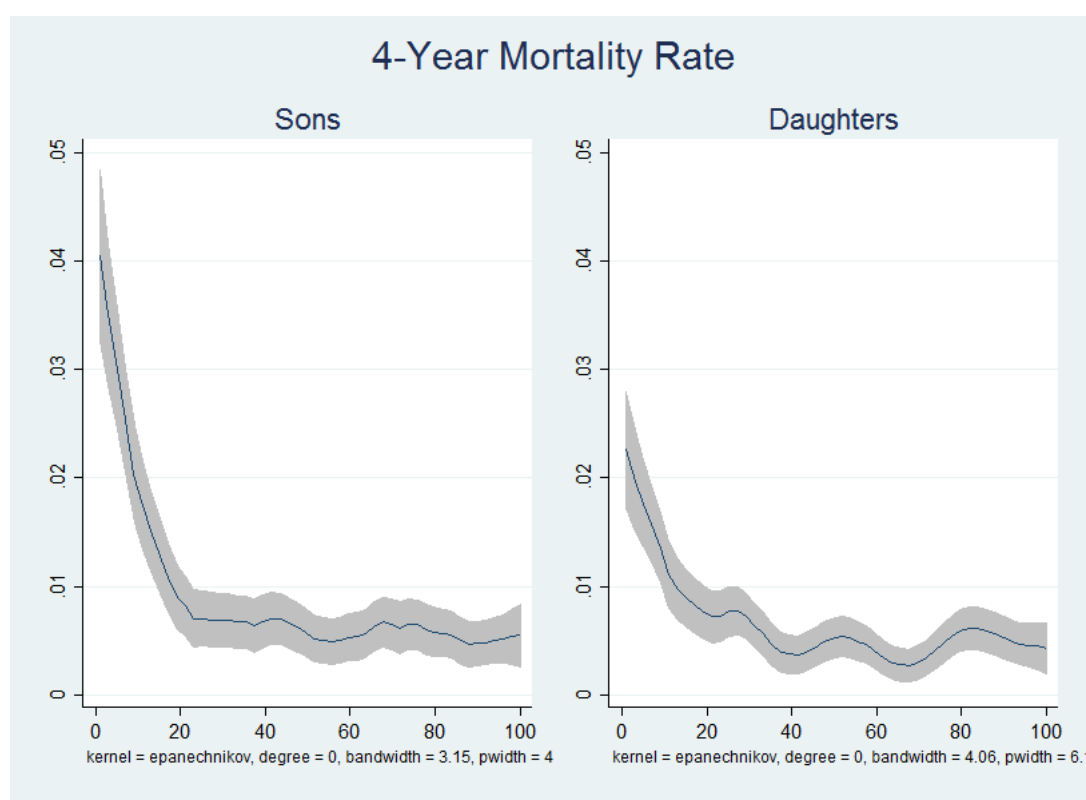
Notes: Nonparametric regression of child mortality rate in age interval 46-50 On predicted health rank. Confidence bands are based on a 95 % level.

Figure 6: Child Health and Disability Pension

Notes: Nonparametric regression of child disability pension rate in age interval 46-50 on predicted health rank. Confidence bands are based on a 95 % level.

Figure 7: Parent GP visits and Mortality

Notes: Nonparametric regression of parent mortality rate in age interval 70-74 on GP visits quintile. Confidence bands are based on a 95 % Level.

Figure 8: Child GP visits and Disability Pension

Notes: Nonparametric regression of child disability pension rate in age interval 46-50 on GP visits quintile. Confidence bands are based on a 95 % level.

Figure 9: Child GP visits and Mortality

Notes: Nonparametric regression of child mortality rate in age interval 46-50 on GP visits quintile. Confidence bands are based on a 95 % level.

Table 2: Intergenerational Associations In Health

	(i) Son-Father	(ii) Son-Mother	(iii) Daughter-Father	(iv) Daughter-Mother
Panel A: Rank-Rank Slopes (β_R)				
Health Rank Father	0.112 (0.005)		0.111 (0.005)	
Health Rank Mother		0.127 (0.005)		0.145 (0.005)
Panel B: Correlations (γ)				
Health Father	0.113 (0.005)		0.115 (0.005)	
Health Mother		0.125 (0.005)		0.142 (0.005)
Observations	36,178	36,178	35,282	35,282

Notes: All regressions are performed using Ordinary Least Squares. Robust standard errors clustered on family level are in parentheses.

43

Figure 10: Binned Scatter Plots for Health
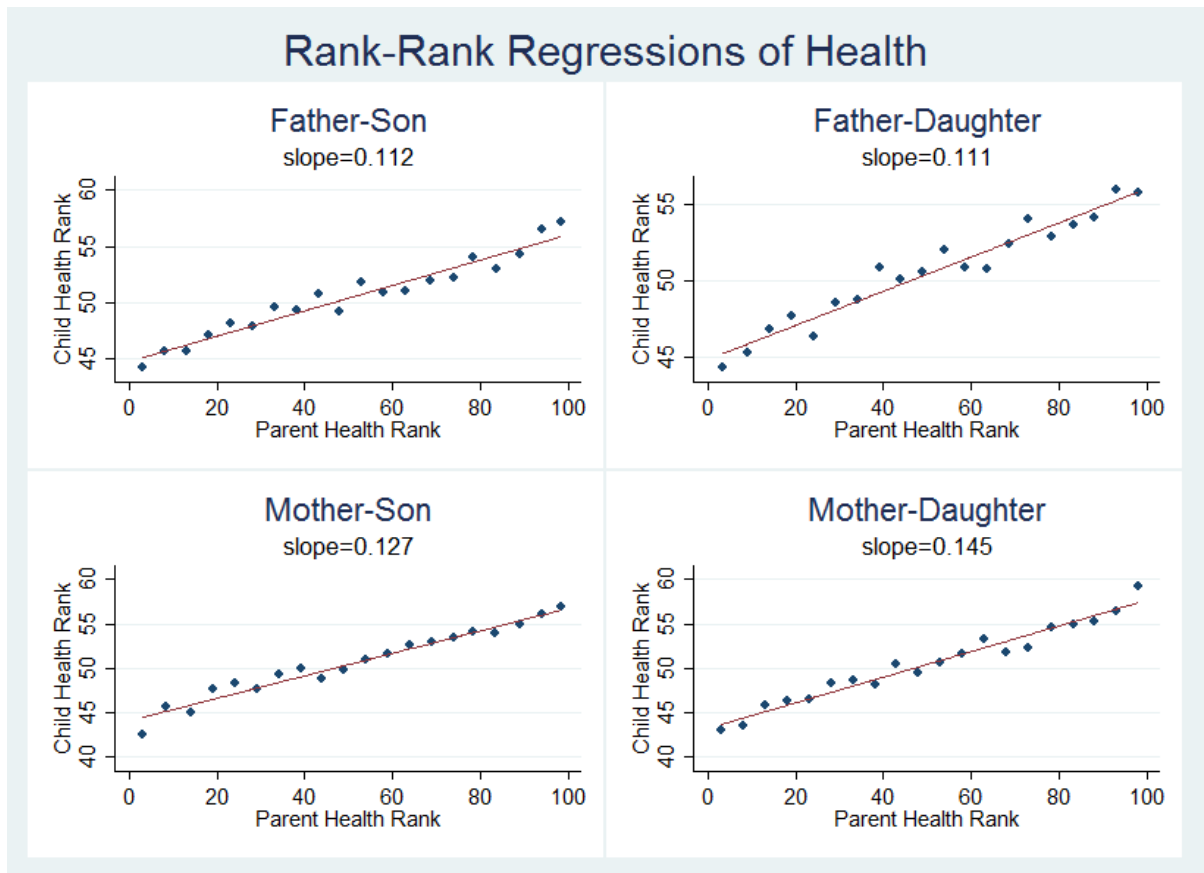
Notes: Estimations and samples are identical to Panel A of Table 2. Additionally, the graph shows a non-parametric plot of the conditional expectation of child health rank given parent health rank.
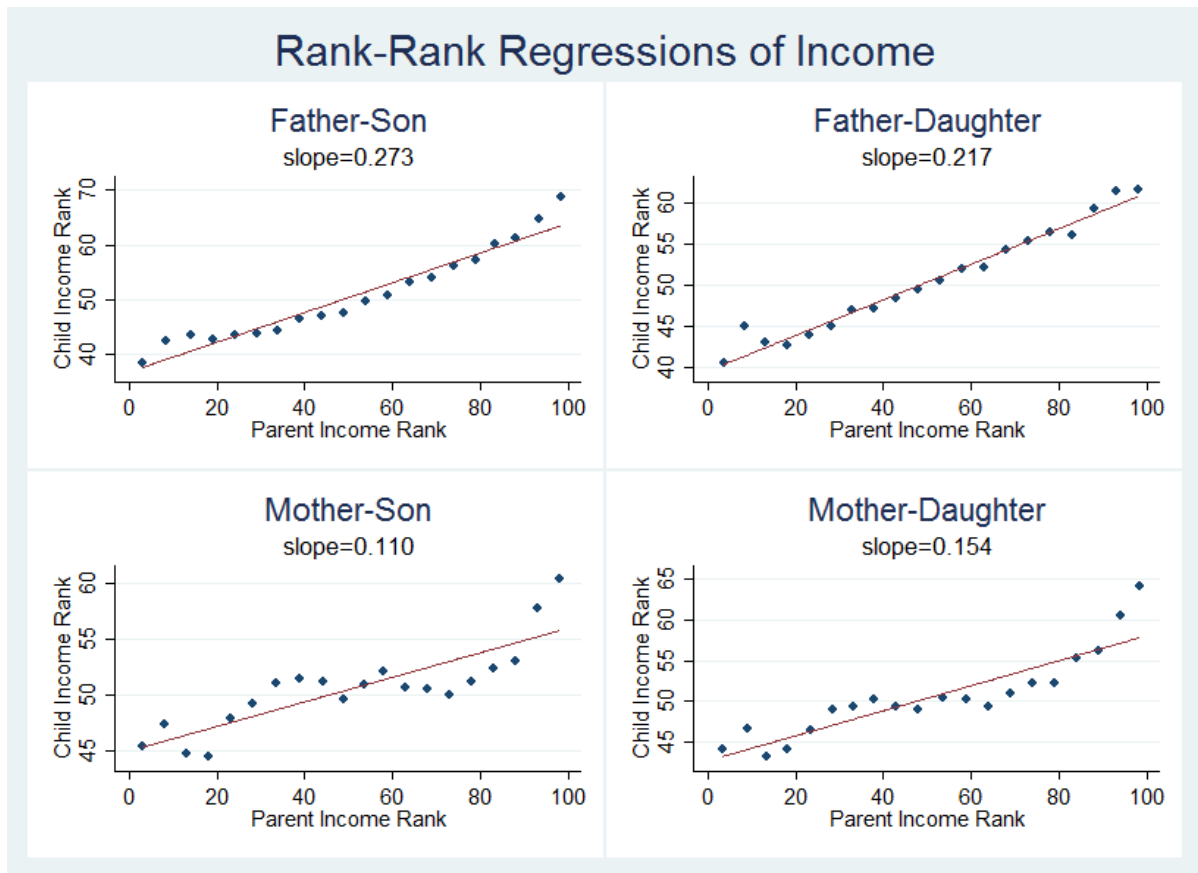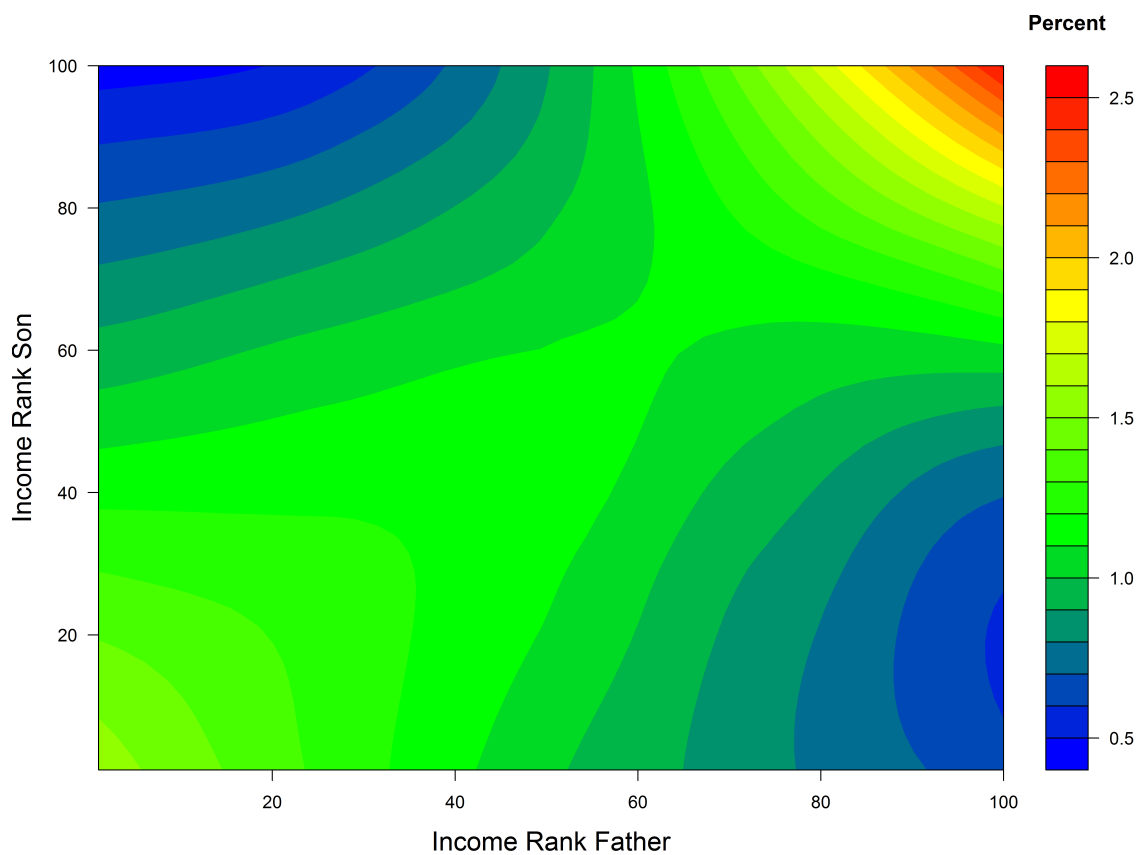
Figure 11: Contour Map for Health (Father-Son)
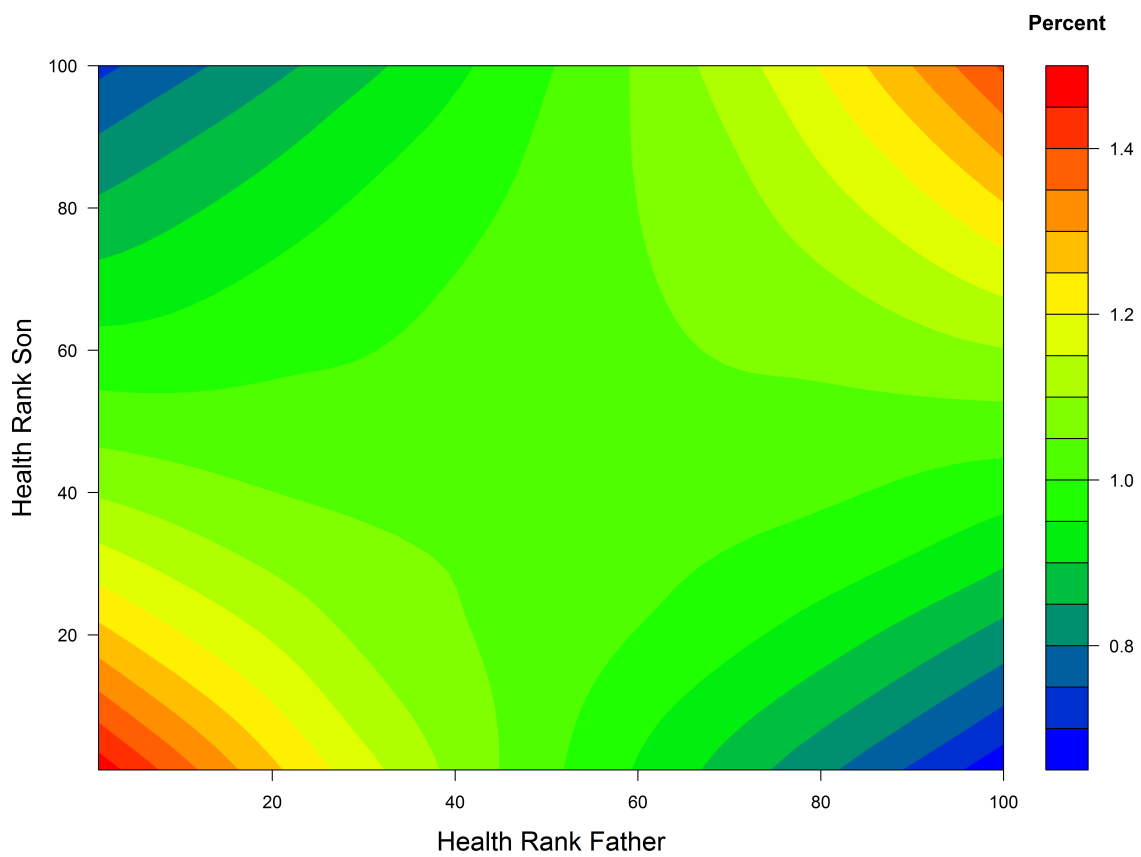
Notes: Estimations are based on a 100 x 100 grid of father and son health percentiles. Child densities are then calculated and interpolated to provide a continuous appearance.

Table 3: Intergenerational Associations In Log Of Income

	(i) Son-Father	(ii) Son-Mother	(iii) Daughter-Father	(iv) Daughter-Mother
Panel A: Rank-Rank Slopes (β_R)				
Income Rank Father	0.273 (0.005)		0.217 (0.005)	
Income Rank Mother		0.110 (0.005)		0.154 (0.005)
Panel B: Correlations (γ)				
Income Father	0.212 (0.005)		0.130 (0.005)	
Income Mother		0.048 (0.005)		0.061 (0.005)
Observations	36,178	36,178	35,282	35,282

Notes: All regressions are performed using Ordinary Least Squares. Robust standard errors clustered on family level are in parentheses.

Figure 12: Binned Scatter Plots for Income

Notes: Estimations and samples are identical to Panel A of Table 3. Additionally, the graph shows a non-parametric plot of the conditional expectation of child income rank given parent income rank.
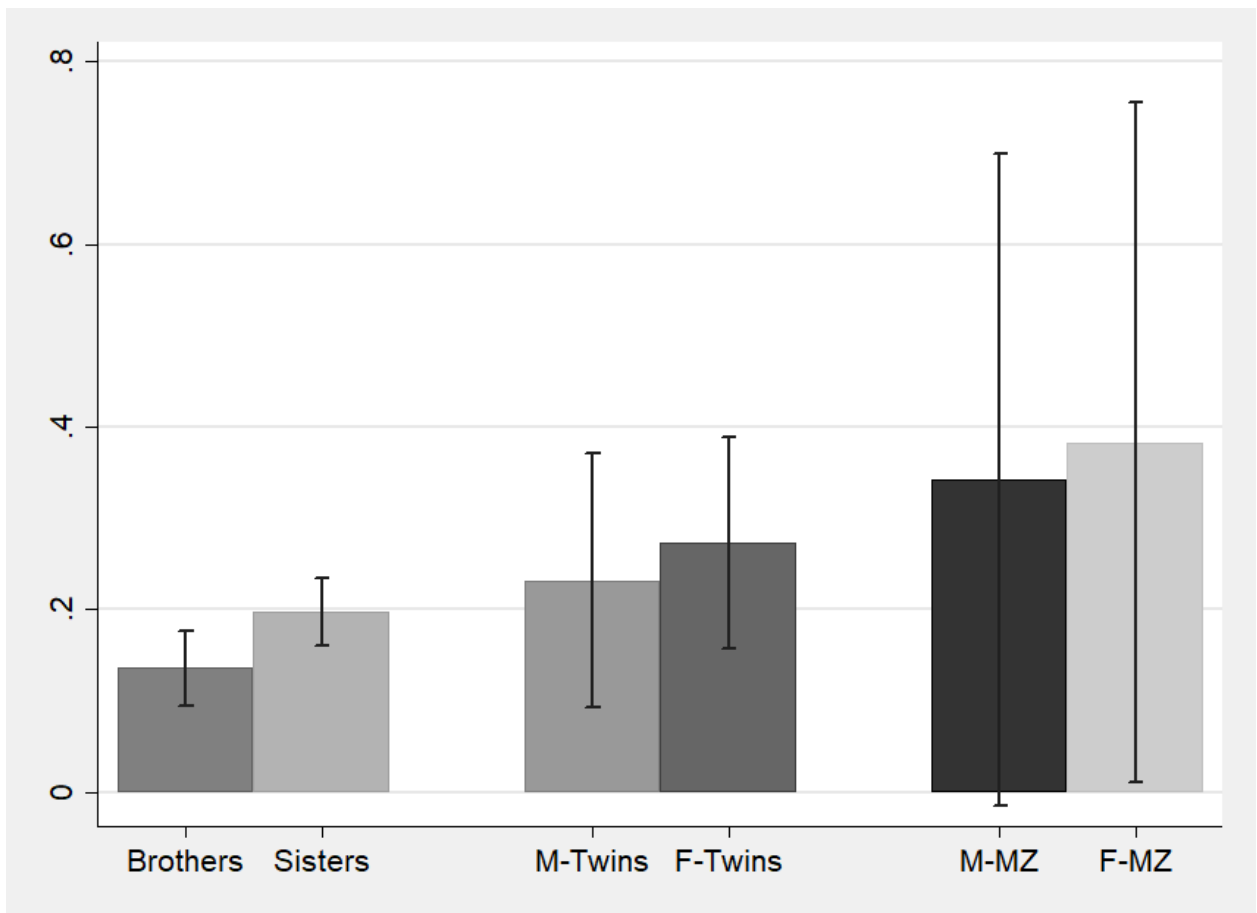
Figure 13: Contour Map for Income (Father-Son)

Notes: Estimations are based on a 100 x 100 grid of father and son income percentiles. Child densities are then calculated and interpolated to provide a continuous appearance.

Figure 14: Sibling Correlations in Health

Notes: Estimated sibling correlations for both genders. The first block is siblings, the second block is for twins, the third block is for identical twins. "M-" denotes males, and "F-" denotes females. Estimations are based on 36,178 sons from 32,674 families and 35,282 daughters from 31,983 families. The error bars show 95 % confidence intervals from bootstrapped standard errors.

Figure 15: Sibling Correlations in Income
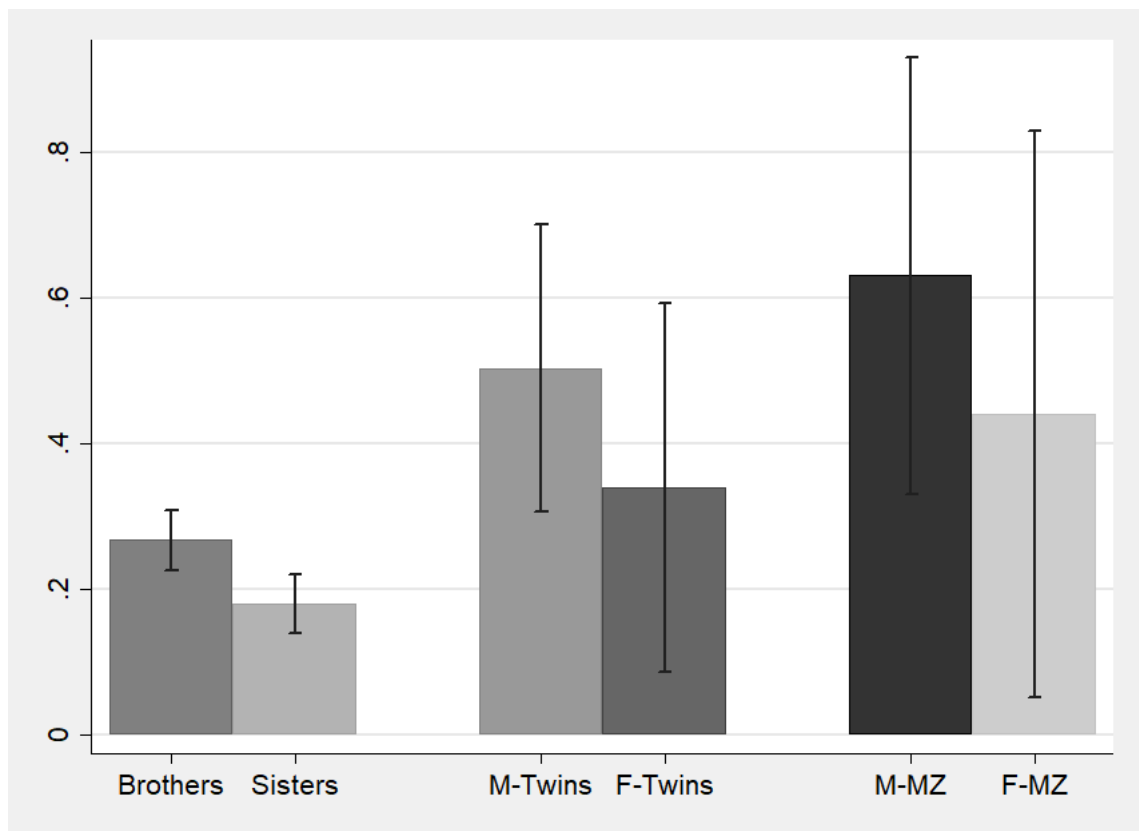
Notes: Estimated sibling correlations for both genders. The first block is siblings, the second block is for twins, the third block is for identical twins. "M-" denotes males, and "F-" denotes females. Estimations are based on 36,178 sons from 32,674 families and 35,282 daughters from 31,983 families. The error bars show 95 % confidence intervals from bootstrapped standard errors.

Figure 16: Attenuation Bias: Years of Parent Health Measurement
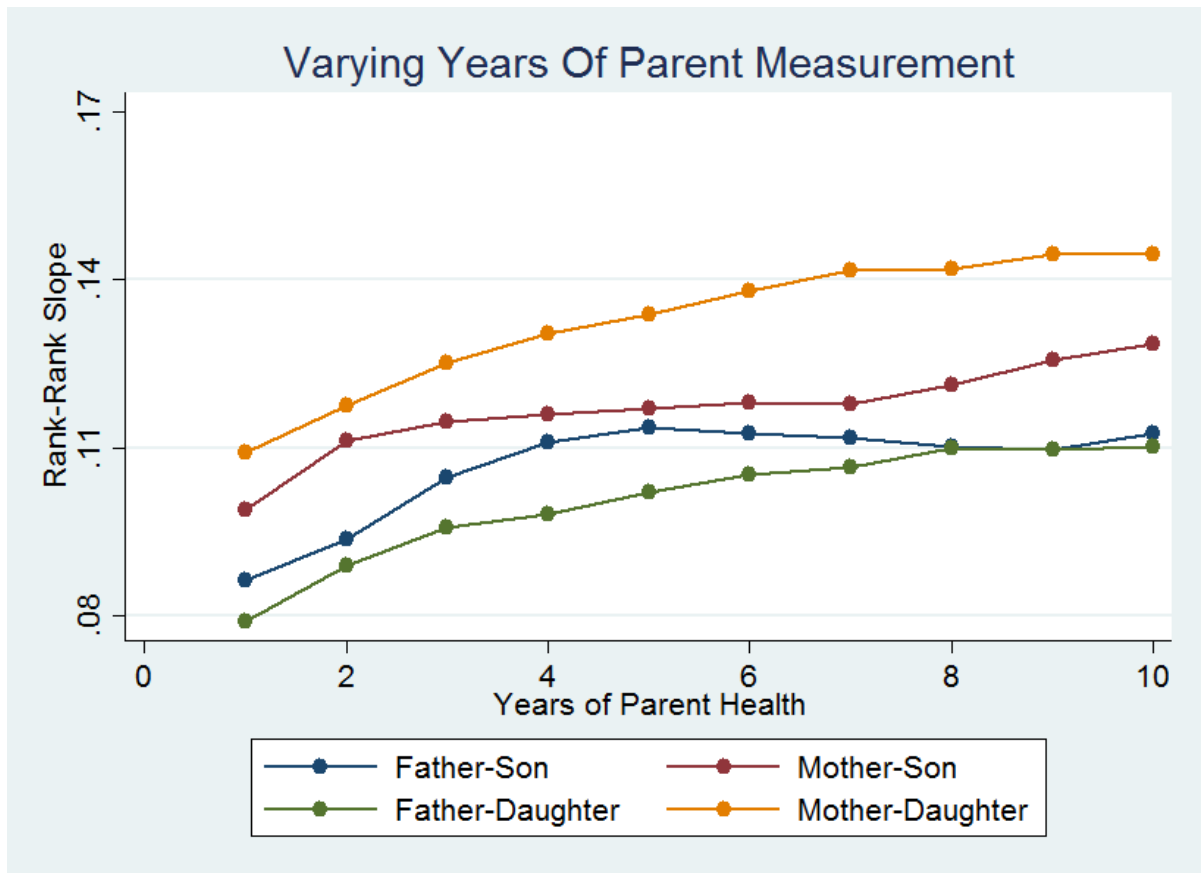
Notes: Estimations and samples are identical to Panel A of Table 2. Additionally, the graph shows a non-parametric plot of the conditional expectation of child health rank given parent health rank.

Figure 17: Testing Lifecycle Bias in Health
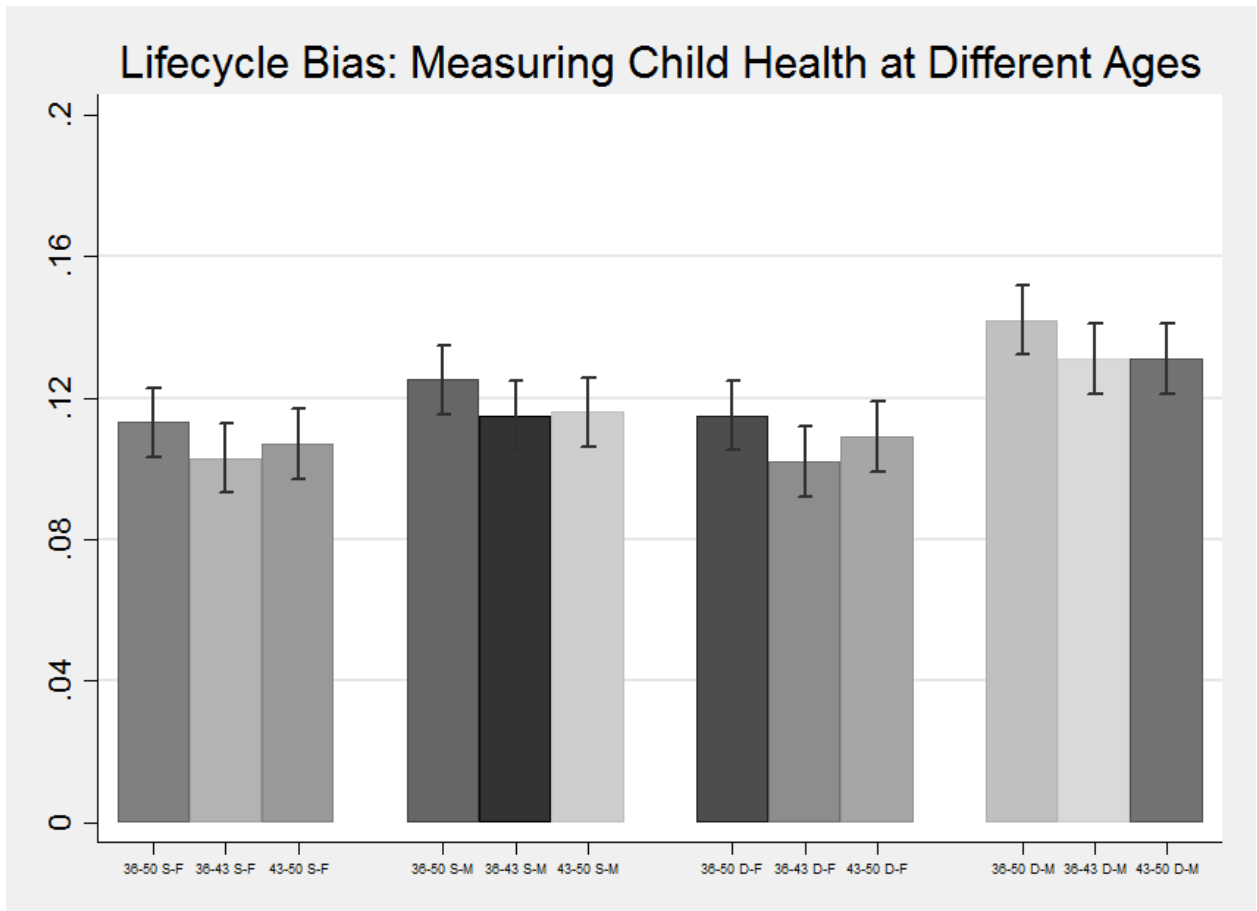
Notes: Estimations and samples are identical to Panel A of Table 2. The figure shows the intergenerational rank-rank slopes for the four different combinations of parent and child when the measurement window for health is split into two different segments of the life-cycle. "36-43 S-F" then means the estimated rank-rank slope between son and father when sons' health is averaged over age 36-43. "S-M" is son-mother, "D-F" is daughter-father, while "D-M" is daughter-mother.

Figure 18: Robustness Check: Intergenerational Correlations In Health
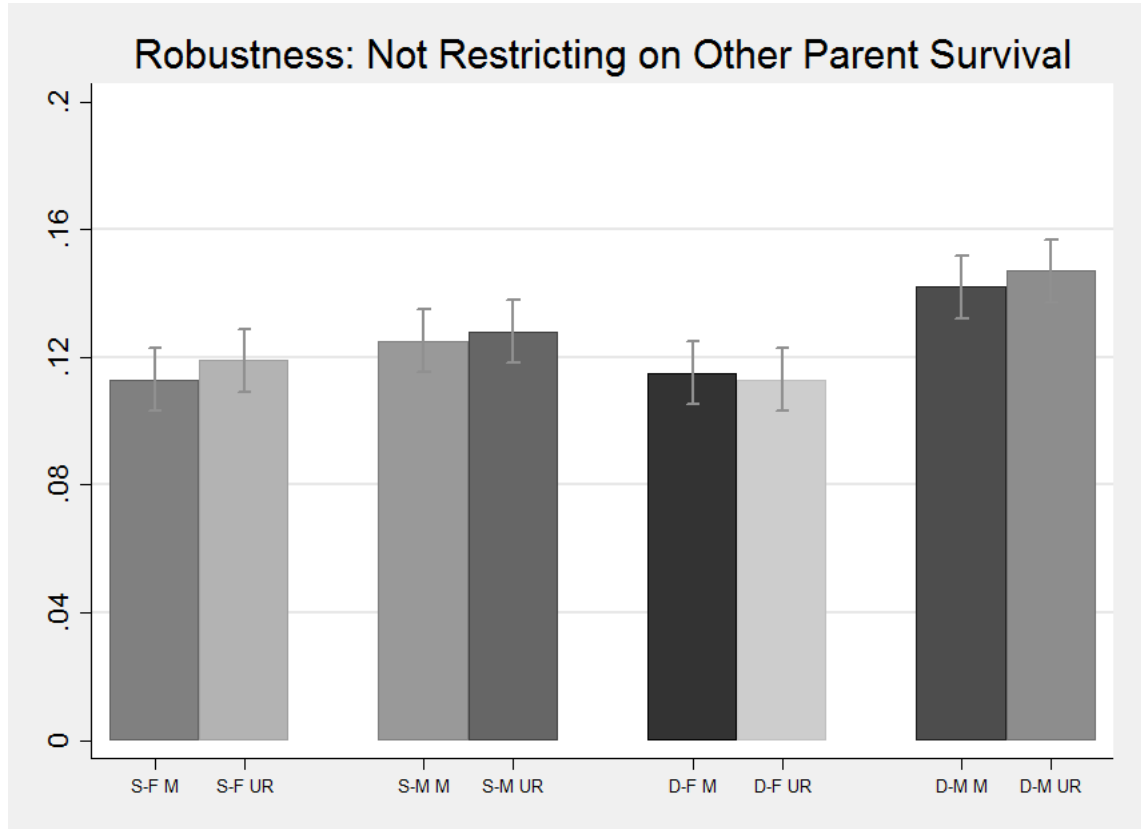
Notes: The intergenerational correlations for the four different combinations of parent and child when no survival restrictions are imposed on survival of the other parent. "S-F"=Son-Father, "S-M"=Son-Mother, "D-F"=Daughter-Father, "D-M"=Daughter-Mother. In each of the four blocks, the left panel is the main intergenerational correlation ("M") from Table 2, and the right is the unrestricted specification ("UR"). For example, "S-F UR" is the correlation between son and father where no restrictions are imposed on survival of the mother. "S-M" is son-mother, "D-F" is daughter-father, while "D-M" is daughter-mother. The sample size is 44,117 for "S-F", 48,928 for "S-M", 42,979 for "D-F" and 47,713 for "D-M".

Figure 19: Robustness Check: Sibling Correlations in Health
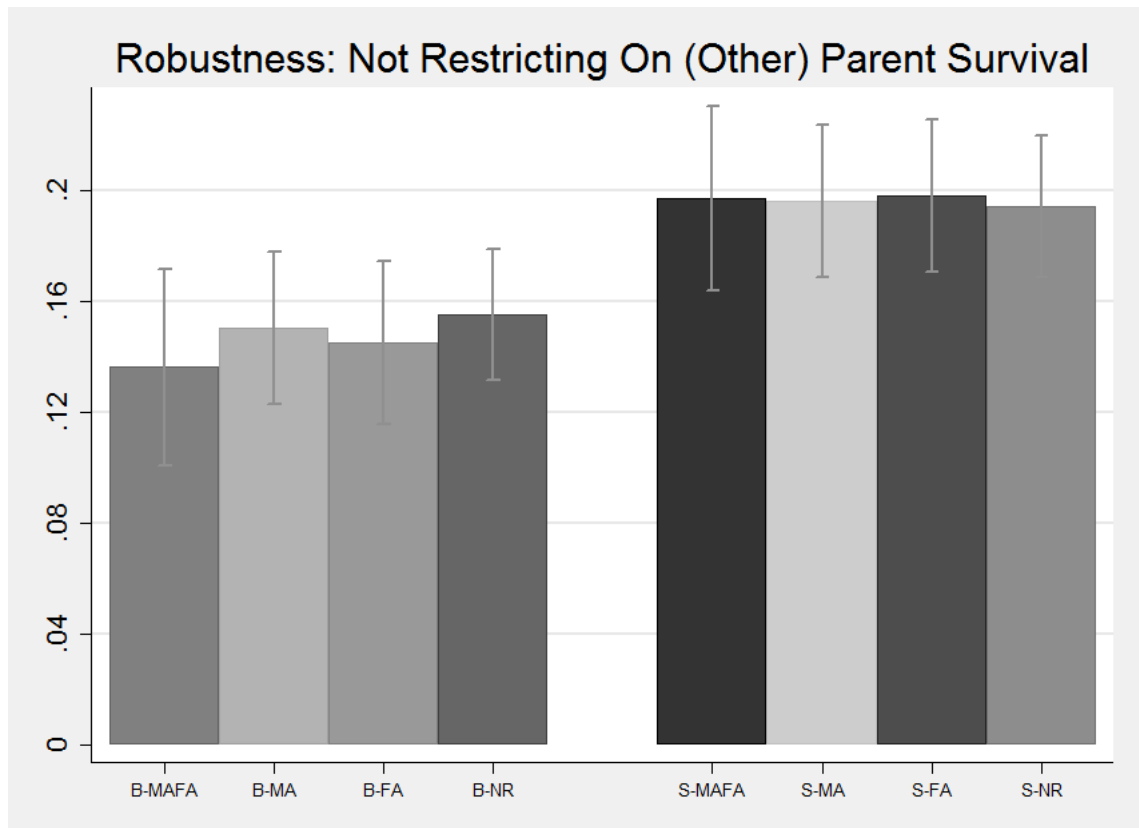
Notes: Estimated sibling correlations for both genders. The first block is Brothers ("B"), the second block is Sisters ("S"). MAFA="Mother And Father Alive", "MA"=Mother Alive, "FA"=Father Alive, "NR"=No Restrictions. The error bars show 95 % confidence intervals from bootstrapped standard errors. "MAFA"-estimations are identical to the main estimation sample. Hence, they are based on based on 36,178 sons from 32,674 families and 35,282 daughters from 31,983 families. "NR"-estimations are based on the largest sample available. Here, estimations are based on 57,095 sons from 52,501 from families and 56,559 daughters from 51,108 families.

Figure 20: Robustness Check: Sibling Correlations in Health
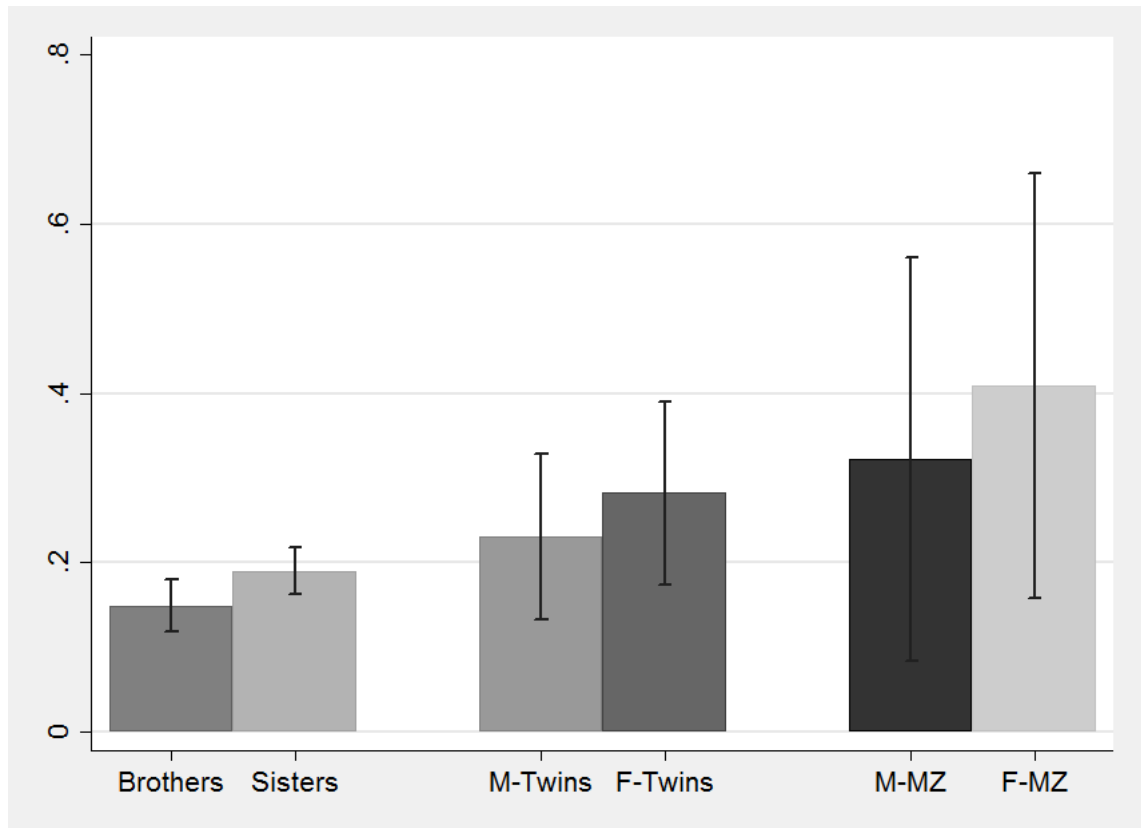
Notes: Estimated sibling correlations for both genders. The first block is siblings, the second block is for twins, the third block is for identical twins. "M-" denotes males, and "F-" denotes females. The error bars show 95 % confidence intervals from bootstrapped standard errors. Estimations are from from the sample, where no restrictions on parent survival are imposed. Hence, estimations are based on 57,095 sons from 52,501 from families and 56,559 daughters from 51,108 families.

References

Andrade, S. B. and Thomsen, J.-P. Intergenerational Educational Mobility in Denmark and the United States. *Sociological Science*, 5:93–113, 2018.

Avendano, M. and Galama, T. J. Inequalities in mortality in the US and Denmark: more alike than different. A commentary on Hoffmann. *Social Science & Medicine*, 73(11):1569–1572, 2011.

Baadsgaard, M. and Brønnum-Hansen, H. Stigende social ulighed i levetiden 2011. *Arbejder-bevægelsens Erhvervsråd og Statens Institut for Folkesundhed, København*, 2011.

Becker, G. S. and Tomes, N. Human capital and the rise and fall of families. *Journal of Labor Economics*, 4(3, Part 2):S1–S39, 1986.

Bertakis, K. D., Azari, R., Helms, L. J., Callahan, E. J., and Robbins, J. A. Gender differences in the utilization of health care services. *Journal of Family Practice*, 49(2):147–147, 2000.

Bingley, P., Corak, M., and Westergård-Nielsen, N. The intergenerational transmission of employers in Canada and Denmark. *IZA DP No. 5593*, 2011.

Björklund, A. and Jäntti, M. Intergenerational income mobility and the role of family background. *Oxford Handbook of Economic Inequality*, pages 491–521, 2009.

Björklund, A. and Jäntti, M. How important is family background for labor-economic outcomes? *Labour Economics*, 19(4):465–474, 2012.

Björklund, A. and Salvanes, K. G. Education and family background: Mechanisms and policies. In *Handbook of the Economics of Education*, volume 3, pages 201–247. Elsevier, 2011.

Björklund, A., Eriksson, T., Jäntti, M., Raaum, O., and Österbacka, E. Brother correlations in earnings in Denmark, Finland, Norway and Sweden compared to the United States. *Journal of Population Economics*, 15(4):757–772, 2002.

Björklund, A., Roine, J., and Waldenström, D. Intergenerational top income mobility in Sweden: Capitalist dynasties in the land of equal opportunity? *Journal of Public Economics*, 96(5-6):474–484, 2012.

Black, S. E. and Devereux, P. J. Recent developments in intergenerational mobility. Technical report, National Bureau of Economic Research, 2010.

Bond, S. and Sales, J. Household work in the UK: an analysis of the British Household Panel Survey 1994. *Work, Employment and Society*, 15(2):233–250, 2001.

Boserup, S. H., Kopczuk, W., and Kreiner, C. T. Born with a silver spoon? Danish evidence on wealth inequality in childhood. Technical report, National Bureau of Economic Research, 2016.

Bredtmann, J. and Smith, N. Inequalities in Educational Outcomes: How Important Is the Family? *Oxford Bulletin of Economics and Statistics*, 80(6):1117–1144, 2018.

Brown, T. A. *Confirmatory factor analysis for applied research*. Guilford Publications, 2014.

Bügelmayer, E. and Schnitzlein, D. D. Is it the family or the neighborhood? Evidence from sibling and neighbor correlations in youth education and health. *The Journal of Economic Inequality*, pages 1–20, 2018.

Chetty, R., Hendren, N., Kline, P., and Saez, E. Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics*, 129(4): 1553–1623, 2014.

Coneus, K. and Spiess, C. K. The intergenerational transmission of health in early childhood—Evidence from the German Socio-Economic Panel Study. *Economics & Human Biology*, 10(1):89–97, 2012.

Corak, M. *Generational income mobility in North America and Europe*. Cambridge University Press, 2004.

Corak, M. Inequality from generation to generation: The United States in comparison. *IZA Discussion Papers 9929*, 2012.

Corak, M. Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives*, 27(3):79–102, 2013.

Corak, M., Lindquist, M. J., and Mazumder, B. A comparison of upward and downward intergenerational mobility in Canada, Sweden and the United States. *Labour Economics*, 30:185–200, 2014.

Currie, J. and Madrian, B. C. Health, health insurance and the labor market. *Handbook of Labor Economics*, 3:3309–3416, 1999.

Currie, J. and Moretti, E. Biology as destiny? Short-and long-run determinants of intergenerational transmission of birth weight. *Journal of Labor economics*, 25(2):231–264, 2007.

Datta Gupta, N., Kleinjans, K., and Larsen, M. The effect of a severe health shock on work behavior: Evidence from different health care regimes. *Social Science & Medicine*, 136-137:44 – 51, 2015.

Devaux, M. and De Looper, M. Income-related inequalities in health service utilisation in 19 OECD countries, 2008-2009. *OECD Health Working Papers, No. 58*, 2012.

Dolton, P. and Xiao, M. The intergenerational transmission of body mass index across countries. *Economics & Human Biology*, 24:140–152, 2017.

Dwyer, D. S. and Mitchell, O. S. Health problems as determinants of retirement: Are self-rated measures endogenous? *Journal of Health Economics*, 18(2):173–193, 1999.

Erikson, R. and Goldthorpe, J. H. *The constant flux: A study of class mobility in industrial societies.* Oxford University Press, USA, 1992.

Galton, F. Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 15:246–263, 1886.

García-Gómez, P., Van Kippersluis, H., O'Donnell, O., and Van Doorslaer, E. Long-term and spillover effects of health shocks on employment and income. *Journal of Human Resources*, 48(4):873–909, 2013.

Goode, A., Mavromaras, K. G., and Smith, M. D. Intergenerational transmission of healthy eating behaviour and the role of household income. *Available at SSRN 1145914*, 2008.

Grawe, N. D. Lifecycle bias in estimates of intergenerational earnings persistence. *Labour Economics*, 13(5):551–570, 2006.

Grossman, M. The demand for health: a theoretical and empirical investigation. *NBER Books*, 1972.

Halliday, T. and Mazumder, B. An Analysis of Sibling Correlations in Health using Latent Variable Models. *Health Economics*, 26(12):108–125, 2017.

Halliday, T., Mazumder, B., and Wong, A. Intergenerational Health Mobility in the US. *IZA DP No. 11304*, 2018.

Helliwell, J., Layard, R., and Sachs, J. World Happiness Report 2016 Volume I. Technical report, The United Nations, 01 2016.

Herskind, A. M., McGue, M., Holm, N. V., Sörensen, T. I., Harvald, B., and Vaupel, J. W. The heritability of human longevity: a population-based study of 2872 Danish twin pairs born 1870–1900. *Human Genetics*, 97(3):319–323, 1996.

Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N., and Verashchagina, A. The inheritance of educational inequality: International comparisons and fifty-year trends. *The BE Journal of Economic Analysis & Policy*, 7(2), 2007.

Jantti, M., Bratsberg, B., Roed, K., Raaum, O., Naylor, R., Osterbacka, E., Bjorklund, A., and Eriksson, T. American exceptionalism in a new light: a comparison of intergenerational earnings mobility in the Nordic countries, the United Kingdom and the United States. *IZA DP No. 1938*, 2006.

Jencks, C. and Tach, L. Would Equal Opportunity Mean More Mobility? *Mobility and Inequality: Frontiers of Research in Sociology and Economics*, pages 23–58, 2006.

Jones, C. I. and Klenow, P. J. Beyond GDP? Welfare across countries and time. *American Economic Review*, 106(9):2426–57, 2016.

Kessler, R. C., Gilman, S. E., Thornton, L. M., and Kendler, K. S. Health, well-being, and social responsibility in the MIDUS twin and sibling subsamples. *The John D. and Catherine T. MacArthur foundation series on mental health and development. Studies on successful midlife development. How healthy are we?: A national study of well-being at midlife*, pages 124–152, 2004.

Kreiner, C. T. and Sievertsen, H. H. Neonatal Health of Parents and Cognitive Development of Children. *CEBI Working Paper 02/18*, 2018.

Landersø, R. and Heckman, J. J. The Scandinavian Fantasy: The Sources of Intergenerational Mobility in Denmark and the US. *The Scandinavian Journal of Economics*, 119(1):178–230, 2017.

Loureiro, M. L., Sanz-de Galdeano, A., and Vuri, D. Smoking habits: like father, like son, like mother, like daughter? *Oxford Bulletin of Economics and Statistics*, 72(6):717–743, 2010.

Lund, T. B., Sandøe, P., and Lassen, J. Attitudes to publicly funded obesity treatment and prevention. *Obesity*, 19(8):1580–1585, 2011.

Marmot, M. G., Stansfeld, S., Patel, C., North, F., Head, J., White, I., Brunner, E., Feeney, A., and Smith, G. D. Health inequalities among British civil servants: the Whitehall II study. *The Lancet*, 337(8754):1387–1393, 1991.

Mazumder, B. Fortunate sons: New estimates of intergenerational mobility in the United States using social security earnings data. *Review of Economics and Statistics*, 87(2):235–255, 2005.

Mazumder, B. Sibling similarities and economic inequality in the US. *Journal of Population Economics*, 21:685–701, 02 2008.

Mazumder, B. Estimating the intergenerational elasticity and rank association in the US: Overcoming the current limitations of tax data. *Federal Reserve Bank of Chicago, Working Paper No. 2015-04*, 2015.

McClellan, M. B. Health events, health insurance, and labor supply: Evidence from the health and retirement survey. In *Frontiers in the Economics of Aging*, pages 301–350. University of Chicago Press, 1998.

Mill, J., Dempster, E., Caspi, A., Williams, B., Moffitt, T., and Craig, I. Evidence for monozygotic twin (MZ) discordance in methylation level at two CpG sites in the promoter region of the catechol-O-methyltransferase (COMT) gene. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 141(4):421–425, 2006.

Nicoletti, C. and Rabe, B. Inequality in pupils' test scores: How much do family, sibling type and neighbourhood matter? *Economica*, 80(318):197–218, 2013.

Nielsen, M. E. J. and Andersen, M. M. Should We Hold the Obese Responsible?: Some Key Issues. *Cambridge Quarterly of Healthcare Ethics*, 23(4):443–451, 2014.

OECD. An overview of growing income inequalities in OECD countries: main findings. *Divided We Stand: Why Inequality Keeps Rising*, 2011.

Parman, J. et al. Gender and intergenerational mobility: Using health outcomes to compare intergenerational mobility across gender and over time. Technical report, 2010.

Piketty, T. Theories of persistent inequality and intergenerational mobility. *Handbook of Income Distribution*, 1:429–476, 2000.

Piraino, P., Muller, S., Cilliers, J., and Fourie, J. The transmission of longevity across generations: The case of the settler Cape Colony. *Research in Social Stratification and Mobility*, 35:105–119, 2014.

Redondo-Sendino, Á., Guallar-Castillón, P., Banegas, J. R., and Rodríguez-Artalejo, F. Gender differences in the utilization of health-care services among the older adult population of Spain. *BMC Public Health*, 6(1):155, 2006.

Reeves, R. V. *Dream hoarders: How the American upper middle class is leaving everyone else in the dust, why that is a problem, and what to do about it.* Brookings Institution Press, 2018.

Roemer, J. E. *Equality of opportunity.* Harvard University Press, 2009.

Roos, L. L., Walld, R., and Witt, J. Adolescent outcomes and opportunities in a Canadian province: looking at siblings and neighbors. *BMC Public Health*, 14(1):506, 2014.

Sahl Andersen, J., De Fine Olivarius, N., and Krasnik, A. The Danish national health service register. *Scandinavian Journal of Public Health*, 39(7):34–37, 2011.

Schanzenbach, D. W., Nunn, R., and Bauer, L. The changing landscape of American life expectancy. *Washington, DC: The Hamilton Project, June*, 2016.

Schmidt, M., Schmidt, S. A. J., Sandegaard, J. L., Ehrenstein, V., Pedersen, L., and Sørensen, H. T. The Danish National Patient Registry: a review of content, data quality, and research potential. *Clinical Epidemiology*, 7:449, 2015.

Shah, A. R. and Gardner, T. W. Diabetic retinopathy: research to clinical practice. *Clinical Diabetes and Endocrinology*, 3(1):9, 2017.

Smith, G. D. Intergenerational influences on health: how far back do we have to go? *International Journal of Epidemiology*, (38):617–618, 2009.

Solon, G. Intergenerational income mobility in the United States. *The American Economic Review*, pages 393–408, 1992.

Solon, G. Intergenerational mobility in the labor market. In *Handbook of Labor Economics*, volume 3, pages 1761–1800. Elsevier, 1999.

Solon, G. Cross-country differences in intergenerational earnings mobility. *Journal of Economic Perspectives*, 16(3):59–66, 2002.

Statistics Denmark. DIAG23, 2019. URL `https://www.dst.dk/da/Statistik/dokumentation/Times/sygehusbenyttelse/diag23`.

Torche, F. Analyses of intergenerational mobility: An interdisciplinary review. *The ANNALS of the American Academy of Political and Social Science*, 657(1):37–62, 2015.

Wall Street Journal. The Summer's Most Unread Book Is..., 2014. URL `http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm`.

Whitehead, M. The concepts and principles of equity and health. *Health Promotion International*, 6 (3):217–228, 1991.

Zimmerman, D. J. Regression toward mediocrity in economic stature. *The American Economic Review*, pages 409–429, 1992.

Economics Working Papers

2018-04:   Martin Paldam: The strategies of economic research - An empirical study

2018-05:   Ingo Geishecker, Philipp J.H. Schröder, and Allan Sørensen: One-off Export Events

2018-06:   Jesper Bagger, Mads Hejlesen, Kazuhiko Sumiya and Rune Vejlin: Income Taxation and the Equilibrium Allocation of Labor

2018-07:   Tom Engsted: Frekvensbaserede versus bayesianske metoder i empirisk økonomi

2018-08:   John Kennes, Daniel le Maire and Sebastian Roelsgaard: Equivalence of Canonical Matching Models

2018-09:   Rune V. Lesner, Anna Piil Damm, Preben Bertelsen and Mads Uffe Pedersen: Life Skills Development of Teenagers through Spare-Time Jobs

2018-10:   Alex Xi He, John Kennes and Daniel le Maire: Complementarity and Advantage in the Competing Auctions of Skills

2018-11:   Ninja Ritter Klejnstrup, Anna Folke Larsen, Helene Bie Lilleør and Marianne Simonsen: Social emotional learning in the classroom: Study protocol for a randomized controlled trial of PERSPEKT 2.0

2018-12:   Benjamin U. Friedrich and Michał Zator: Adaptation to Shocks and The Role of Capital Structure: Danish Exporters During the Cartoon Crisis

2019-01   Ran Sun Lyng and Jie Zhou: Household Portfolio Choice Before and After a House Purchase

2019-02:   Mongoljin Batsaikhan, Mette Gørtz, John Kennes, Ran Sun Lyng, Daniel Monte and Norovsambuu Tumennasan: Daycare Choice and Ethnic Diversity: Evidence from a Randomized Survey

2019-03:   Dan Anderberg, Jesper Bagger, V. Bhaskar and Tanya Wilson: Marriage Market Equilibrium, Qualifications, and Ability

2019-04:   Carsten Andersen: Intergenerational Health Mobility: Evidence from Danish Registers