# Intergenerational Wealth Mobility: Evidence from Danish Wealth Records of Three Generations*

Simon Boserup
University of Copenhagen

Wojciech Kopczuk
Columbia University, CEPR and NBER

Claus Thustrup Kreiner
University of Copenhagen and CEPR

March 1, 2013

Abstract

We provide empirical evidence on the intergenerational mobility of wealth using administrative wealth records for three generations of Danes. Our preferred estimate for the intergenerational wealth elasticity (IWE) is 0.19. The elasticity estimate obtained from the selected sample with positive wealth (reported elsewhere in the literature) is 0.27. We construct a theoretical framework that allows for understanding the variability of the IWE across time, samples and countries. Our framework highlights that the IWE can be interpreted as the weighted average of elasticities corresponding to different sources of intergenerational correlation that may in principle vary in importance across different contexts. However, we find that the IWE estimate are more or less the same for different age groups, when using parents-grandparents instead of child-parents, when eliminating bequests, and when explicitly shutting down many of the potential channels behind intergenerational wealth mobility, including income and education. This suggests that parental wealth is a sufficient statistic for the channels that we control for and those that vary across different samples: that is, the effect of these parental characteristics on wealth of children can be summarized by their effect on wealth of parents. As Charles and Hurst (2003), we find that financial characteristics affect IWE so that their effect goes beyond the impact on parental wealth, suggesting that correlation in risk attitudes and investment behavior is an important factor (although the magnitude of this effect is smaller than in PSID). Finally, we find robust evidence that grandparental characteristics have an effect beyond parental ones. This finding indicates that focusing on a two-generation IWE leads to severe underestimation of the long-term persistence in the formation of wealth across generations.

**Preliminary and incomplete draft**

*We thank Søren Leth-Petersen for helpful discussion. Benjamin Ly Serena provided excellent research assistance. Financial support from WEST is gratefully acknowledged. Contact information: simon.h.boserup@econ.ku.dk, wojciech.kopczuk@columbia.edu, ctk@econ.ku.dk.

# 1 Introduction

The main objective of the literature studying intergenerational mobility is to analyze the extent to which economic well-being is related across generations, and to try to understand the underlying mechanisms (Piketty, 2000). In other words, of interest is both the strength of relationship between lifetime economic resources/consumption possibilities of children and parents, and whether the statistical association is governed by mechanical transmission of abilities and traits from parents to children or by active decision-making of parents such as investment in human capital of their children. Intergenerational correlations in earnings and income have been extensively studied (see e.g. the surveys of Solon (1999) and Black and Devereux (2011)), but much less evidence exists on intergenerational wealth mobility, even though wealth measured at some point in life may be as good as income (or even better) to proxy for lifetime economic resources. An important exception is Charles and Hurst (2003), hereafter C&H, who use wealth data from the Panel Study of Income Dynamics (PSID) to estimate the elasticity of child wealth with respect to parental wealth for the United States. C&H obtain an age-adjusted intergenerational wealth elasticity (IWE) of 0.37 before transfer of bequest (i.e., both parents are alive), which is at the low end of the range of estimates of the intergenerational income elasticity for the US, and find that around 2/3 of the elasticity is accounted for by income and asset ownership.[1]

In this paper, we provide new empirical evidence on the intergenerational mobility of wealth using Danish administrative wealth records of three generations observed for the entire Danish population. Our data has several advantages. First, in our basic sample we have more than 1 million child-parents pairs compared to around 1400 in C&H. This allows us to include a rich number of controls and to split the sample into age-cohorts while still obtaining very precise estimates. Second, the use of administrative data removes problems of attrition and measurement errors that often plaque survey studies.[2] Third, we are able to link comprehensive data for three generations (child-parents-grandparents). The information about an additional generation allows us to test whether child-parent relationships are stable and may also be used to address whether the simple IWE estimate provides an accurate measure of the degree of persistence in the wealth process across generations.

We contribute to the literature in a number of ways. We start by providing a theoretical framework capturing the relevant mechanisms that may create correlation of wealth across generations and that lays the foundation for the empirical analysis. The standard Becker and Tomes (1979, 1986) framework focuses on links in economic outcomes across generations due to correlation of abilities across generations and due to parental investment in human capital of children. These channels determine the association of

[1] Charles and Hurst (2003) review a few older studies looking at the intergenerational correlation of wealth. These studied have looked at small non-representative samples with few observations and poor data quality.

[2] Recent research has documented large survey measurement errors in key economic variables such as income and that the errors are correlated with conventional covariates implying that errors are non-classical (Kreiner et al., 2012).

income across generations and therefore are also likely to influence the relationship between wealth across generations. In addition to these mechanisms, however, wealth may be correlated across generations because of direct transfer of wealth from previous generations (inter vivo or through inheritance), because of correlation in patience and risk preferences creating differences in saving propensities (Stiglitz, 1969) or because of correlation in investment ability and the corresponding return (for example, due to stock market participation, entrepreneurship or attitudes toward borrowing). Individual wealth observed at a given point in time reflects therefore in a complicated way economic resources, abilities and traits inherited from the previous generation, and we show theoretically how the IWE coefficient is related to these underlying mechanisms.

The first part of our empirical analysis follows C&H. We find an (unconditional) child-parent age-adjusted wealth elasticity equal to 0.27 when focusing on indviduals with positive wealth as they do. This estimate is considerably lower than the IWE estimate of C&H. When we address the sample selection bias by using use the full population, i.e. do not remove child-parents pair with negative wealth of either child or parents (this is our baseline specification), the estimate is even lower at 0.19.[3] Recent studies for France (Arrondel, 2009) and Denmark (Kolodziejczyk, 2011) also find an IWE estimate of 0.2. The lower IWE estimate for Denmark is compatible with cross-country studies of intergenerational earnings elasticities finding low elasticities in the Nordic countries and the lowest elasticity for Denmark equal to 0.12 (Björklund and Jännti, 2009).

Life-cycle variation has proven to be an important consideration in the measurement of intergenerational mobility in income that appears to be age-dependent (Haider and Solon, 2006). Similarly, there may be substantial variation in the IWE depending on the age of the child and the parents. To address this issue we run separate regressions for each age-cohort of the children and for each age-cohort of the parents measured at the year when the child was born, respectively. The IWE coefficients are very precisely estimated and lie all within the range 0.16–0.22, revealing a surprisingly stable relationship between child wealth and parental wealth. In our baseline estimates, therefore, we can abstract from age-dependent differences in wealth correlation and proceed by pooling all cohorts while flexibly control for age of parents and children to adjust for the life-cycle patterns.

We obtain a similar conclusion when introducing additional covariates and when splitting the sample. When we add covariates that themselves have a strong explanatory power such as number of siblings dummies, income level, education length and portfolio composition dummies then the child-parent wealth elasticity falls but not a lot. For example, income and education of children and parents can only explain

[3] C&H use a standard log transformation of the data and therefore remove non-positive values. However, negative wealth may be optimal from an economic theory point of view in certain circumstances, in particular for young individuals, and many individuals in our data have negative wealth. To allow for negative wealth, we apply the inverse hyperbolic sine transformation (IHS) of the data, which yields identical results as the log transformation for both positive wealth and the absolute value of negative wealth taken in isolation, but in addition allows for combining the two.

8 percent of the original IWE estimate. The strongest effect comes from financial composition dummies that may explain 25 percent of the IWE, indicating that correlation in household finance behavior across generations may be important, but all explanatory variables taken together can only explain up to 30 percent of the IWE. We also split the sample according to parental age. This has quite large impact on the estimates of the effect of child and parental income on wealth but only small effect on the IWE estimate.

Our theoretical framework suggests that robustness of our estimates to inclusion of income and education (and to a lesser extent financial composition dummies) is consistent with parental wealth being a sufficient statistic for the mechanisms behind correlation of wealth that these variables proxy for.

In the second part of our empirical analysis, we exploit availability of information about grandparents. We start by addressing whether the relationship between child wealth and parental wealth is stable over time/generations. Recent research has documented substantial changes over the long run in top income shares, in the relative importance of capital and labor income at the top of the income distribution, and in the evolution of inheritance (Atkinson, Piketty and Saez, 2011; Piketty, 2011) and, similarly, there may be long run forces that reduce or increase the strength of this relationship. In light of our framework, this would be so when parental wealth is not a sufficient statistic for parental influence so that changing importance of different mechanisms behind intergenerational correlation of wealth translates into changing IWE. Consistenly with our original approach that amounted to controlling for different mechanisms, when estimating the intergenerational elasticity of parental wealth with respect to grandparental wealth we obtain an estimate of the IWE that is very similar to the estimate obtained from children-parents pairs. This is even more striking when taking into consideration that the generations differ in many other respects than just age, and indicates that the cumulative importance of the underlying mechanisms governing the intergenerational relationship in wealth is quite stable.

Next, we look at the correlation between children and grandparents in isolation. If parental wealth is a sufficient statistic for all previous generations then the coefficient on grandparental wealth should be the child-parent IWE raised to the power of two. We obtain a coefficient that is more than three times as high and very precisely estimated. This indicates that the degree of persistence across generations is higher than what is reflected in the IWE estimate. A reason may be that the underlying processes relating wealth across generations have more memory than just one generation, for example because grandparents have a direct impact on their grandchildren (Solon, 2012). When including both parental and grandparental wealth in the regression, we obtain a significant and sizable coefficient on grandparental wealth, while the child-parents coefficient only falls a little. This is consistent with the hypothesis of more persistence and underlying processes relating wealth across generations having more memory than just one generation.

Another likely reason for the significant coefficient on grandparents is measurement errors in wealth

or, in the same vein, that wealth is a noisy signal of abilities and traits transmitted across generations and of the size of money transfers given from generation to generation. Attenuation bias created by transitory components in the economic outcomes has been a major concern in the intergenerational income mobility literature since the influential contribution of Solon (1992), and it is common to take averages over some years as we have done to reduce the importance of transitory components. However, this method may only remove a small part of the attenuation bias if the transitory component has some persistence (Mazumder, 2005) and would not work at all to remove a bias from random "fixed effects", e.g. you are born lucky in terms of ability given your family background without transmitting this to your children. In the theory section, we demonstrate how idiosyncratic variation in wealth may bias the IWE estimates downward but also provide conditions under which it is appropriate to use wealth of grandparents an instrument for parental wealth in order to obtain consistent estimates. When we redo the first part of the empirical analysis, we consistently obtain estimates of the IWE in the range 0.6-0.7, which is more than three times as high as the original ordinary least square estimates. This points again to much higher degree of persistence in the wealth formation across generations.

The exclusion restriction that allows for using grandparental wealth as an instrument for parental one is strong. Hence, the results indicate one of the two possibilities: either the extent of correlation between generations is significantly underestimated when using the conventional approach or the effects extend for more than just a single generation. In either case, we interpret it as indicating that the extent of intergenerational mobility is substantially lower than a naïve estimate would indicate.

As an alternative approach, we further inquire into the degree of persistence by allowing for dynasty fixed effects. The standard Becker and Tomes framework, underlying intergenerational empirical analyses, assumes that economic outcomes of all generations converge towards the same constant. But as demonstrated by Stiglitz (1969) and in our theory section, people may belong to different classes/dynasties who differ with respect to productivity or savings behavior, implying that wealth, without any new shocks, converges towards a steady state distribution rather than a constant. Our empirical analysis allowing for fixed effects indicates that the model with a common constant is misspecified and provides therefore further evidence to the conclusion that the conventional IWE estimates underestimate the degree of persistence.

The remaining part of the paper is organized as follows. Section 2 provides a theoretical framework to understand the relevant mechanisms underlying the correlation of wealth across generations and the challenges in estimating the degree of persistence across generations. Section 3 describes the construction of the data sets and provides summary statistics of key variables. Section 4 describes the results of the empirical analysis. Finally, Section 5 offers concluding remarks and an appendix provides additional details concerning the data.

## 2 A theory of the correlation of wealth across generations

We start by providing a simple conceptual framework for understanding the relationship between wealth of different generations. Let's denote $w_g$ to be wealth of generation $g$. A member of family $i$ that belongs to generation $g$ maximizes lifetime utility $u(\{C_a^g, q_a^g\}_{a=o}^{a=A}; \varepsilon_{ig})$ where $C_a^g$ is consumption at age $a$, where $0 \leq a \leq A$, $q_a$ are transfers to the subsequent generation made at age $a$ (inter vivos when $a < A$ and bequests for $a = A$) and $\varepsilon_{ig}$ is the set of taste parameters characterizing preferences of this individual. Optimization is subject to the initial wealth of $W_0^g = 0$ and the set of budget constraints for $a > 0$

$$W_a^g = W_{a-1}^g \cdot (1 + r_{ig} + \gamma_a) + y_a^g - C_a^g + B_a^g - q_a^g,$$

where $W_a$ is wealth at age of $a$, $y_a^g$ is income at the age of $a$, $B_a^g$ is the bequest received from the previous generation at age of $a$ (with one member per generation, $B_a^g = q_{\tilde{a}}^{g-1}$ where $\tilde{a} - a$ is the age difference between the parent and the child), $r_{ig}$ is the mean rate of return specific to that particular individual (to allow for varying investment strategies) and $\gamma_a$ is the mean-zero deviation from the normal rate of return. Consequently, the level of wealth can be expressed in terms of the exogenous parameters of the problem as

$$W_a^g \left( \{q_b^{g-1}\}_{b=0}^A, \{y_b^g\}_{b=0}^A, r_{ig}, \varepsilon_{ig}, \{\gamma_b\}_{b=0}^a \right) . \tag{1}$$

In other words, wealth at any given age depends on the history and future (expected) transfers from parents, the history and future own income, the rate of return, preference parameters and stochastic shocks.

We are interested in understanding the relationship between wealth of different generations. Most simply, one can observe the statistical association that may be, for example, described using the correlation coefficient $\mathrm{corr}(W_a^g, W_{\tilde{a}}^{g-1})$, where $a$ and $\tilde{a}$ are ages at which we observe wealth of children and parents, respectively, or the empirical association between some transformation of observed wealth levels $\frac{dw_a^g}{dw_{\tilde{a}}^{g-1}}$, where (if we focus on positive wealth) we could use for example $w = \ln(W)$ or (as we will do in what follows) we can use inverse hyperbolic sine transformation $w = \log(W + \sqrt{W^2+1})$ that behaves as $\pm \log(|W|)$ everywhere with the exception of in the neighborhood of zero.

Let's focus attention on

$$\frac{dw_a^g}{dw_{\tilde{a}}^{g-1}} .$$

It should be clear that this is not a well-specified concept unless we take a stand on why $w_{\tilde{a}}^{g-1}$ varies. We can estimate a statistical relationship between wealth of different generations that combines the effect of various reasons why wealth levels could co-vary, but in general this relationship may be specific to the sample choice reflecting varying importance of different determinants of wealth.

From the formula (1), we can identify a number of reasons why wealth would co-vary:

- transfers from parents $q_b^{g-1}$ are a function of their own characteristics and hence correlated with parental wealth

- incomes of parents and children may be correlated for many reasons extensively analyzed in the literature

- rates of returns and preferences may be correlated across generations

Hence, depending on the relative importance of these factors, we may expect to see different relationship at different times and places and different sample choices unless they happen to correspond to precisely the same strength of correlation in wealth (the notion that we will expand on in what follows).

In order to make progress and introduce our empirical specification, we simplify our framework to resemble what has been sometimes dubbed the "mechanical" approach in the literature (Goldberger, 1989). It amounts to positing a statistical relationship between variables of different generations without explicitly specifying the nature of the individual optimization problem. We adopt that terminology but note that the framework we described is an optimization framework, except that we have refrained so far from imposing additional assumptions so that the framework serves only to identify the relevant variables. Our main simplification in what follows is approximation of the (potentially nonlinear) relationship rather than assuming away optimization.

In order to operationalize our empirical approach, we linearize our original wealth formula as

$$w_a^g \approx \alpha + \beta_q q_a^{g-1} + \beta_y y_a^g + \beta_r(1 + r_{ig}) + \beta_\varepsilon \varepsilon_{ig} + \eta_a^g, \tag{2}$$

where $\eta$ is an error term incorporating rate of return shocks and approximation errors that is assumed (very strongly) orthogonal to the other variables and we posit that the current values of right hand side variables are sufficient to summarize the relationship with wealth. While we write all terms in the linear fashion to simplify notation, they can correspond to transformations (e.g., log or IHS) of the original variables.

When we estimate the statistical relationship between $w_a^g$ and $w_a^{g-1}$ as

$$w_a^g = \beta_w w_a^{g-1} + \omega,$$

the coefficient $\beta_w$ is going to reflect the average impact of $w_a$ running through all four possible channels: bequests, income, preferences and rate of return.

In what follows, we abstract from the age aspect of the inter-generational relationship (we will return to considering it in the empirical work). It will prove useful to write the determinants of wealth as a vector $x_g = (q^{g-1}, y^g, 1 + r_{ig}, \varepsilon_{ig}, \eta_g)$, with wealth expressed as $w_g = x_g\beta + \zeta_g$, where $\zeta_g$ is white noise measurement error.

Finally, we specify the law of motion for determinants of wealth as $x_g = x_{g-1}\Xi + \nu_g$. We assume for now that it is autoregressive with the order one, but deviations from this assumption will be important to consider in what follows.

Note that we have introduced many potentially unobservable sources of variation in wealth. First, taste parameters $\varepsilon$ are likely unobservable directly (though perhaps they may be proxied for). It is natural to consider and test whether correlation in preferences is a source of intergenerational persistence. Second, we split the rate of return into two separate components. One is the normal rate of return — this is akin to a preference parameter in the sense of reflecting traits of individuals that lead them to select investments with varying outcome, although it is conceptually distinct since it represents manifestation of preferences through choices rather than preferences themselves. The other one is $\eta$ that corresponds to deviations from the normal rate of returns (and approximation errors that we abstract from). We assume that it is an idiosyncratic components so that $\eta$ is assumed not to be correlated across generations ($\Xi_{\eta\eta} = 0$). We include it though in $x_g$ because it is certainly possible that such random shocks to wealth do in fact have an impact on the subsequent generations by influencing other variables. The final source of randomness is $\zeta$ — this is assumed to be measurement error or other sources of variation in wealth that have no consequence for the subsequent generations.

Using this notation (and expressing all variables in terms of their deviation from the mean in order to eliminate the constant term), we can re-write wealth as follows:

$$w_g = x_g\beta + \zeta_g = x_{g-1}\Xi\beta + \nu_g\beta + \zeta_g = \beta_w x_{g-1}\beta + \xi_g = \beta_w w_{g-1} - \beta_w \zeta_{g-1} + \xi_g \tag{3}$$

where $\beta_w$ is a linear projection (linear regression coefficient) of $x_{g-1}\Xi\beta + \nu_g\beta + \zeta_g$ on $x_{g-1}\beta$. Using the standard OLS formula $(X'X)^{-1}X'Y$ with $X = x_{g-1}\beta$ and $Y = x_{g-1}\Xi\beta + \nu_g\beta + \zeta_g$ and noting that $\nu_g\beta + \zeta$ is orthogonal to $x_{g-1}\beta$, we can write the expected value of $\beta_w$ as

$$\mathrm{E}[\beta_w] = ((x_{g-1}\beta)'x_{g-1}\beta))^{-1} \cdot (x_{g-1}\beta)' \cdot (x_{g-1}\Xi\beta) = (\beta' \cdot (x'_{g-1}x_{g-1}) \cdot \beta)^{-1} \cdot \beta' \cdot (x'_{g-1}x_{g-1}) \cdot \Xi\beta, \tag{4}$$

showing how $\beta_w$ summarizes the complicated web of influences that links wealth across generations. It reflects the causal relationship between relevant arguments of formula (2) ($\Xi$) that can be mapped to one-dimensional wealth within each generation via $\beta$. In general, the mapping of wealth across generations depends on the distribution of characteristics in the population $x$. This is because there are many channels through which characteristics of subsequent generations relate and the population relationship between wealth of different generations reflects different components of $\Xi$ depending on variation of $x_{g-1}$ in the population. This is reflected by the presence of $x'_{g-1}x_{g-1}$ in the formula.

The coefficient $\beta_w$ is as close as one can get to a "structural" intergenerational correlation of wealth if one insists on summarizing it by a single parameter. In the special case where $x_{g-1}$ is one dimensional,

so that $\Xi$ is a scalar, it is easy to show that $\beta_w = \Xi$, implying that the $\beta_w-$coefficient reveals the (single) structural parameter.

## 2.1 Interpretation of $\beta_w$

More generally, it is a relationship that is specific to a given population and it does not have a direct causal interpretation. This is because the source of variation in wealth ($w_{g-1}$) matters: depending on which component of $x_{g-1}$ is responsible, the effect will vary accordingly. $\beta_w$ does though reflect the effect of variation in wealth in the population that combines influences from variety of sources.

To illustrate the logic of this formula, consider the following example

**Example 1** Suppose that elements of $x$ are uncorrelated with each other ($x_g' x_g$ is diagonal), $\Xi$ is diagonal and only two elements of $\beta$, indexed by $i$ and $j$ are non-zero. Then straightforward manipulation yields

$$\beta_w = \frac{\beta_i \sigma_{ii}^2 \Xi_{ii} + \beta_j \sigma_{jj}^2 \Xi_{jj}}{\beta_i \sigma_{ii}^2 + \beta_j \sigma_{jj}^2}$$

where $\sigma_{ii}^2$ is the $(i,i)$ element of $x_g' x_g$ (variance of $i$th element of $x_g$, denoted by $x_{gi}$) and $\Xi_{ii}$ is the $(i,i)$ element of $\Xi$

In this particular example, the correlation of wealth has its source in two different channels, $i$ and $j$. A marginal increase in $x_{g-1,i}$, translates into $\beta_i$ increase in wealth of generation $g-1$ ($w_{g-1}$) and into $\beta_i \Xi_{ii}$ increase in wealth of generation $w_g$. Hence, the intergenerational elasticity driven by this source of variation is $\Xi_{ii}$. Analogously, the intergenerational elasticity driven by $j$th element is $\Xi_{jj}$. When both of the sources are present at the same time, $\beta_w$ has to reflect both of these sources and their contribution depend on the relative magnitude of $\beta_k \sigma_{kk}^2$ $(k = i, j)$. This is intuitive: what matters is the extent to which a given channel is responsible for variation in wealth — this is affected by the extent to which the given channel influences wealth ($\beta_k$) and the extent of variation in the underlying characteristics, $\sigma_{kk}$.

Note that even when one is willing to assume that $\beta$ and $\Xi$ are structural parameters, there is no single structural intergenerational elasticity here. Different societies at different points in time may differ with respect to the extent of variation in determinants of wealth — variation in bequests, tastes, education, rate of returns may all vary over time with institutions, policies, culture etc. Each of such situation will correspond to a different weighted average of $\Xi_{ii}$ and $\Xi_{jj}$ as the measured IWE.

Another way of phrasing it is that the intergenerational wealth elasticity is the average treatment effect corresponding to changes in wealth of the prior generation. Since prior wealth may be varying for a multitude of reasons (corresponding to different impact of the treatment — change in prior wealth), the corresponding estimate will reflect a weighed average impact with the weights being sample-dependent.

The example is of course restrictive, but illustrative of the logic of formula (4) that characterizes the general case.

## 2.2 Parental wealth as a sufficient statistic

As discussed above, in general there is no single intergenerational elasticity of wealth because different sources of variation in parental characteristics may in general translate into different strength of intergenerational association. However, it is possible to identify a special case when it is not so. Imagine that parental wealth is a sufficient statistic for the effect of children. That is, retaining the linear structure, suppose that

$$x_g = (w_{g-1} - \zeta_{g-1}) \cdot \Gamma = w_{g-1}\Gamma - \zeta_{g-1}\Gamma$$

for some vector $\Gamma$ (and note that we are using wealth net of $\zeta$ — the measurement error term). Then, $x_g = x_{g-1}\beta\Gamma - \zeta\Gamma$, so that — using our general notation – $\Xi = \beta\Gamma$ and $\nu = \zeta\Gamma$. Substituting into eq. (4) yields

$$\mathrm{E}[\beta_w] = (\beta' \cdot (x'_{g-1}x_{g-1}) \cdot \beta)^{-1} \cdot \beta' \cdot (x'_{g-1}x_{g-1}) \cdot \beta \cdot \Gamma \cdot \beta = \Gamma\beta$$

which can also be seen directly from $w_g = x_g\beta + \zeta_g = w_{g-1}\Gamma\beta + \zeta_g - \zeta_{g-1}\Gamma$.

When wealth is a sufficient statistic for characteristics of grandparents, the IWE is equal to $\Gamma\beta$ regardless of the source of variation in wealth — the formula for $\mathrm{E}[\beta_w]$ is independent of $x_{g-1}$.

## 2.3 Testing importance of different channels

So far, we have clarified the theoretical relationship between wealth of different generations while allowing for many different channels of interactions. Consider now the possibility of observing some of them. Specifically, suppose that we partition $x = (x^1, x^2)$ with $x^1$ unobservable and $x^2$ observed. Similarly, let's denote by $\Xi_{11}$ and $\Xi_{12}$ the partitions of matrix $\Xi$ that determine $x^1$: $x^1_g = x^1_{g-1}\Xi_{11} + x^2_{g-1}\Xi_{12} + \nu^1_g$. Then,

$$\begin{aligned} w_g &= x_g\beta + \zeta_g = x^1_g\beta + x^2_g\beta + \zeta_g = x^1_{g-1}\Xi_{11}\beta + x^2_{g-1}\Xi_{12}\beta + x^2_g\beta + \nu^1_g\beta + \zeta_g \\ &= \beta^1_w x_{g-1}\beta + x^2_{g-1}\beta^2_{w,g-1} + x^2_g\beta^2_{w,g} + \xi^1 \\ &= \beta^1_w w_{g-1} + x^2_{g-1}\beta^2_{w,g-1} + x^2_g\beta^2_{w,g} - \beta^1_w\zeta_{g-1} + \xi^1 \end{aligned} \quad (5)$$

where $\beta^1_w$ is a linear projection of $x^1_{g-1}\Xi_{11}\beta$ on $x_{g-1}\beta$ while partialling out the effect of $x^2_{g-1}$ and $x^2_g$ (with $\beta^2_{w,g-1}$ and $\beta^2_{w,g}$ being the resulting coefficients) . In other words, by controlling for some subset of characteristics of both parents and children (note that controlling for both at the same time is important), we can zoom in on the effect that goes through the remaining characteristics.

## 2.4 Measurement error

Simple inspection of relationships (3) or (5) reveals that OLS of $w_g$ on $w_{g-1}$ will not in general estimate $\beta_w$ (or $\beta^1_w$. This is because the error term $\zeta_{g-1}$ will be correlated with $w_{g-1}$ — recall that $w_{g-1} = \beta \cdot x_{g-1} + \zeta_{g-1}$ or, more explicitly, the relationship is described by formula (2). The attenuation bias due

to the presence of $\zeta$ will not be a problem only when $\text{var}(\zeta) = 0$. One may of course make assumptions to that effect but absent these assumptions, our estimate of $\beta_w$ will be biased.

Assuming that the measurement error term $\zeta_{g-1}$ is not serially correlated, characteristics of preceding generation $x_{g-2}$ can serve as an instrument for $w_{g-1}$. In particular, a natural instrument to consider is $w_{g-2}$. Under the simple autoregressive structure of order one that we assumed, the exclusion restriction is satisfied and this is a valid instrument.

Note though that different instruments will identify different average treatment effects, i.e. in our case they will identify the weighed average of effects going through different channels.

### 2.5 Multiple generations

We also have assumed so far that the characteristics follow the simple autoregressive structure of order one. Two natural generalizations are to consider a higher order autoregressive structure and to consider fixed (or very persistent) family component.

Suppose first that $x_g$ follows second-order autoregressive process, $x_g = x_{g-1}\Xi^1 + x_{g-2}\Xi^2 + \nu_g$, while we continue to have $w_g = x_g\beta + \zeta_g$. We can define $y_g = (x_g, x_{g-1})$, $y_{g-1} = (x_{g-1}, x_{g-2}$, $\Xi^y = [\Xi^1, \Xi^2]$ so that $y_g = y_{g-1}\Xi^y + \nu_g$ and $w_g = y_g\beta^y + \zeta_g$ where $\beta^y = (\beta, 0, \cdots, 0)$. Substituting $y_g$, $\Xi^y$, $\beta^y$ throughout in place of $x_g$, $\Xi$ and $\beta$ allows for repeating the whole argument but now with the set of characteristics expanded to include grandparental ones. On the conceptual level, grandparental characteristics become yet another determinant of IWE.

In the presence of a family fixed effect, we can analogously expand the definition of $x_g$ to include the constant family term.

While both of these extensions are straightforward for the purpose of understanding the IWE $\beta_w$, they raise additional estimation issues. First, they invalidate the possibility of using grandparental characteristics as an instrument. Second, eliminating the fixed family component in order to understand the importance of this channel is complicated due to the presence of the lagged dependent variable structure.

### 2.6 Empirical strategy

We will proceed as follows. First, we will estimate population $\beta_w$. We will follow it by considering various subsamples and different generations in order to test the robustness of this estimate. Sensitivity of the results to different sample choices would indicate that the relative importance of different channels varies depending on the subsample. Lack of sensitivity would indicate either that wealth is a sufficient statistic or that, in fact, different samples considered happen to correspond to a similar mix of channels behind intergenerational correlation of wealth.

Next, we will consider controlling for different channels by including the corresponding characteristics of parents and children, and investigating sensitivity of IWE to these choices. As we discussed, this

approach allows for shedding light on whether the effect of the particular channel can be summarized through its impact on wealth.

Finally, we will consider grandparental characteristics to investigate the relevance of interactions that run across multiple generations.

# 3 Data

# 4 Data

Our empirical analysis is based on data from several public administrative registers gathered by Statistics Denmark and linked together using personal identification numbers. Every citizen in Denmark is assigned a unique personal identification number at birth and the identification numbers of the mother and the father are registered for all Danes born in 1960 and onwards.[4] This enables us to combine different data sources at the individual level and to link data across generations.

The data on individual wealth and income is based on administrative tax return records, rather than survey questions as in Charles and Hurst (2003) and most other studies on intergenerational mobility. The Danish Tax Agency (SKAT) collects, in addition to information of various income sources, information about the values of asset holdings and liabilities measured at the last day of the year for all Danes, and the bulk of the wealth components are third-party reported.[5] The available pieces of information at Statistics Denmark are the aggregate value of assets and liabilities, respectively, covering the period 1980 to 2011, and from 1998 and onwards it is also possible to obtain portfolio information, e.g. value of bonds, stocks, cash in banks, mortgage loans and the sum of other loans. An attractive feature of the wealth data is that the information is not top coded.

The information about the value of financial assets and liabilities at the end of the year is reported to the tax authorities by banks, other financial institutions, and some government institutions, while the cash value of property is assessed by the tax authorities, based on detailed information of the property, and used for taxation of the imputed rent on the property. The third-party reported value of assets includes all deposits, stocks, bonds, value of property, and deposited mortgages. Pension funds are not included. The third-party reported value of liabilities includes debt in financial institutions, mortgage credit debt, credit and debit card debt, deposited mortgage debt, student debt and debt in The Mortgage Bank (a public institution), debt to financial corporations, debt to the Danish municipalities and other liabilities

[4] Registrations of parents exist before 1960 but are incomplete.

[5] The tax authorities use the income information to generate pre-populated tax returns and the wealth data to cross check if the reported income level is consistent with the change in net-wealth during the year under the assumption of a given estimated consumption level. A recent study by Kleven et al. (2011) reveals, using a large scale randomized tax auditing experiment constructed in collaboration with the Danish tax authorities, only small differences between the third-party reported income items and the corresponding items on the final tax return. This indicates that the third party reported information of the Danish Tax Agency is of a very high quality.

such as unpaid taxes and mortgage debt, which are not deposited.

From 1980 to 1996, Denmark had a wealth tax, and tax payers had to self-report car value, boat value, caravan value, title deed of cooperative dwelling, premium bonds, cash deposits, stocks (both listed and non-listed thereby including privately held companies), and private debt. These components are not included in the computations after 1996. Until 1996 the value of stocks was self-reported, while afterwards it became third-party reported by banks and financial institutions (excluding non-listed stocks). The registration of the company value of self-employed has been changed several times but has stayed unchanged since 1997, where assets and liabilities of the firm were registered separately and included, respectively, in the assets and liabilities of the owner. Another definitional change occurs in 1983. Before 1983 all family wealth of a married couple was assigned to the husband, while the wealth of husbands and wives have been registered separately afterwards.

Ideally we would like to observe wealth, income etc. of the individual in the middle age and observe the different generations at the same age because of the life-cycle variation in economic outcomes (Haider and Solon, 2006). If, for example, parents are around 25 years old when their children are born and we observe the wealth of the child generation in 2011 then we should observe the wealth of the parents and grandparents in 1986 and 1961, respectively. This goal has to be balanced against data availability (grandparents) and data quality (parents). Our main empirical analyses are based on parental wealth observed in 1997-1999, where the definition of the wealth measure is the same as that used for children in 2009-2011, and grandparental wealth measured in 1983-1985 where wealth of biological grandfathers and grandmothers are more accurately measured than the years before. We take three year averages of wealth of each individual to reduce the importance of transitory components, as often done in the literature following Solon (1992). The effects on our estimates are very small.

The largest change in the definition of wealth occurs around 1997 where the wealth tax was abolished. However, for 1995 and 1996 Statistics Denmark computed assets and liabilities of each individual using both the new definition of wealth (used for children and parents) and the old definition (used for grandparents). In Appendix A, we exploit this overlap to show that the new wealth measure is well approximated by the old way of measuring wealth, and we provide more details on the wealth data. Additional information on the Danish wealth and income-tax data may be found in Leth-Petersen (2010) and Chetty et al. (2011).

In the empirical analysis, we consider two types of samples: a child-parents sample (CP) and a child-parents-grandparents (CPG) sample. The CP sample focuses only on child-parents relationships without exploiting information of grandparents. In this sample, we consider all children of age 19-51 in 2011, where both parents are alive in 2011, and where both parents are between 19 and 66 years old in 1998. This is very similar to Charles and Hurst (2003) with the exception that they limit the sample to children that are

older than 25 years and require that both children and parents have positive wealth. The importance of these two differences in sample selection will be discussed. The child-parents-grandparents (CPG) sample is based on the CP sample but with the additional requirement that at least one grandparent is alive in 1983. To avoid selection problems, we further require that parents are born in 1960 or later, corresponding to a maximum age of 38 in 1998, implying that the personal identifiers of all the grandparents are known.[6]

Table 1 provides summary statistics of the two samples. In the CP sample, we have 1.16 million child-parent pairs and the CPG sample consists of 97 thousand observations. In both samples, parents are significantly older than their children at the time where wealth levels are observed in the data, and grandparents are older than parents in the CPG sample. Since households normally accumulate wealth over the life cycle up to retirement, we should expect to observe the highest wealth for grandparents and higher wealth for parents than for children, which is also the pattern we see in Table 1. The large sample size allows us to account for the life-cycle effects by age-adjusting the wealth levels using age dummies and to estimate separate effects for different child cohort–parent age constellations in the empirical analyses. Moreover, following the existing literature measuring intergenerational effects, we will relate relative differences within a generation to the relative difference of another generation, which is not sensitive to scale effects.

Notice that wealth is negative for many individuals. This is in particular the case among the child generation where around half of the individuals have negative wealth compared to eight percent in the sample of C&H. One reason is that we do not restrict the sample to children above 25 as C&H, which significantly increases the wealth of children but also of parents, who will on average be older. Another reason is that Danish households have very high debt-to-income ratios (the liability-gross income ratio is around 200 percent for children and 150 percent for parents in the CP sample) compared to other countries, which has received international attention recently (IMF, 2012; European Commission, 2012). The difference to the US and other countries probably reflects that Denmark has a reasonably high universal public pension benefit level, substantial labor market pension savings by international standards, and an extensive social safety net that reduces the need for precautionary savings.

Labor earnings and gross income are on average higher for parents than for children, while earnings for grandparents are lower than for parents reflecting that some of the individuals have retired. The table also reports the years of education counting completed education. Parents are clearly more educated than grandparents but also somewhat more educated than their children in the CPG sample. However, this reflects that many of the children have not yet completed their education.

[6] This reduces the sample considerably. Without the restriction, we obtain the same regression coefficients and a higher statistical precision but we prefer the restricted sample to avoid sample selection bias.

# 5 Empirical analysis

## 5.1 Elasticity of child wealth with respect to parental wealth

We first use the child-parents (CP) sample described above to estimate the (unconditional) elasticity of child wealth with respect to the average wealth of the (biological) parents.[7] This corresponds to estimating an ordinary least squares regression after using the natural log transformation on the wealth measures of children and parents. Column 1 of Table 2 reports the estimated elasticity without age adjustment, while column 2 reports the age-adjusted elasticity obtained by including age dummies of both children and parents in the regression. The finding of a child-parents age-adjusted elasticity equal to 0.268 implies that children born of parents with a wealth level that is 10 percent above the mean of the parent generation can expect to obtain a wealth level that is 2.7 percent above the mean of the child generation. C&H obtained an intergenerational wealth elasticity (IWE) of 0.365 for the United States using the PSID survey data. The lower estimate for Denmark is not surprising. Denmark has a very homogenous population and a high degree of redistribution, and comparative studies have found that Denmark has the lowest intergenerational elasticity of earnings/income.[8]

When applying the log transformation, we are throwing away all child-parents pairs where either the child or the parents have zero or negative net wealth. Most of the empirical literature analyzing intergenerational relationships have looked at economic outcomes that do not attain negative values by definition, for example earnings. In this case, it is natural to apply the log transformation, which has appealing properties. This is, however, not the case when analyzing net wealth which may well be negative and where standard life cycle theory predicts negative values for young persons who have increasing earnings profiles. Another reason for observing negative wealth of households in our case, and also in C&H, is that we are unable to include pension wealth. In order to avoid the potential selection problem of using the log transformation, we will for the remaining part of the paper be using the inverse hyperbolic sine transformation (IHS), $w = \log(W + \sqrt{W^2 + 1})$, that behaves as $\pm \log(|W|)$ everywhere with the exception of in the neighborhood of zero. Column 3 shows the IWE estimate after using the IHS transformation on the sample where wealth is restricted to be positive. The estimate is completely identical to the result based on the log transformation in column 2. Next, we consider the full sample with 1.16 million child-parents observations that include negative wealth of children and/or parents. This gives an IWE of 0.19, which is considerably lower than the estimate obtained from the restricted sample, showing that it may create severe selection bias to remove observations with negative wealth. Note finally from Table 2 that all regression coefficients are very precisely estimated because of the large sample size as illustrated by

[7] Results are unchanged if we use the aggregate wealth of parents instead of the average wealth of parents, which is due to the fact that the log transformation and the IHS transformation of the data remove any scale effects.

[8] An overview of estimates of intergenerational earnings elasticities for different countries may be found in Björklund and Jännti (2009). They report an elasticity for Denmark of 0.12, which makes Denmark the country with the lowest correlation across generations.

the tiny standard deviations of the estimates.[9]

Life-cycle variation may be important when measuring intergenerational mobility in economic outcomes (Haider and Solon; 2006). Thus, there may be substantial variation in the IWE, depending on the age of the child and the parents, that are not removed by including age dummies. The large sample size allows to address this issue by running separate regressions for each age-cohort of the children and for each age-cohort of the parents measured at the year where the child was born, respectively. Figure 1a shows the IWE estimate as a function of the age of the child in 2011 starting from the early twenties and going up to an age of fifty years. The diagram is constructed by running separate regressions for each cohort of the children, including in each regression age dummies of the parents, and then plotting the estimates and the 95 percent confidence interval for each cohort. The confidence interval shows that the IWE is very precisely estimated for each age group and the graph displays a remarkably stable IWE. For all thirty age groups, all estimates lie within a narrow interval from 0.16 to 0.22 without any systematic trend. Figure 1b is constructed in the same way but this time subsamples are created for each age level of the parents at the time when the child was born and child age dummies are included in each regression. The graph shows a weakly increasing correlation of wealth between children and parents when the age-difference between parents and children is larger but the main conclusion is again that the IWE is remarkably stable and within the same interval as in Figure 1a.[10]

## 5.2 Decomposition of the intergenerational wealth elasticity

In Table 3, we have added income controls. Column 1 is the baseline without income controls and is identical to column 4 in Table 2. We first include the income of the child. It is natural to expect that a high income level of the child is associated with a high level of wealth. This is also what we see from column 2 showing that the elasticity of child wealth with respect to child income equals 1.4, implying that a ten percent higher income level is associated with 14 percent higher wealth. Our main interest is how it influences the size of the IWE. If parents with high wealth invest more in the human capital formation of their children, as in the theory of Becker and Tomes (1979) then this would raise the income and wealth of their children. This would generate a positive correlation between child wealth and parental wealth working through child income, and if so the IWE estimate should fall when we introduce child income as a regressor. Table 2 shows that the IWE is unchanged after introducing child income, and does therefore not lend support to a mechanism working through child income. Next, we also introduce parental income in the OLS regression. Column 2 shows that the elasticity of child wealth with respect to

[9] Another concern may be that outliers or observations with zero or close to zero wealth may be very important for the estimates. We have run sensitivity analyses, which revealed no effects on the estimates of removing observations in the tale of the distribution and around zero wealth.

[10] We have also estimated the IWE for each child-parent age constellations. All estimates are significant with the exception of constellations where the age-difference between the child and the parents is very small or very large. The conclusion is again that estimates are roughly the same for all combinations.

parental income is 2.2, which is therefore more important than the child's own income. The reason for an unconditional correlation between wealth of children and parents could arise if the true relationship was from parental income to child wealth, where parental wealth works as a proxy for parental income. If so, the IWE would fall when controlling for parental income, but we observe nearly the same estimate as in column 2. Note also that the unconditional elasticity in our baseline scenario could arise because income of children and parents are related without there being any correlation in wealth, if wealth is working as a proxy for income. If so, the IWE should fall when going from column 1 to column 3, but we observe only a reasonably small decrease. All in all, income explains less than 8 percent of the IWE, and this conclusion is completely unchanged if we use a more flexible specification with, for example, fourth degree polynomials in child income and parental income, respectively.

The last two columns, report the result of splitting the sample depending on parental age. This has a sizable impact on the coefficients on income but only a small impact on the IWE. Note that the coefficient on the child's own income becomes negative. This may reflect that the children are young in this group and that students, who currently have a low income (student benefits), come from a wealthy background and are wealthy themselves compared to young persons earning income in the labor market.

In Table 4, we include number of siblings dummies, years of schooling dummies and financial composition dummies. Column 1 is the baseline with only age dummies included. In column 2, we include number of siblings dummies, so that we only exploit the variation within families of a given size to estimate the IWE. This has nearly no impact on the estimate. Dummies for education length of both parents and children reduce the IWE a little, and the effect is of the same magnitude as when introducing income controls (see column 3 of Table 3). The largest reduction in the IWE arises when we introduce financial composition dummies for children and parents, i.e., dummies for homeownership, stock ownership, bonds ownership, and for being self-employed. This reduces the IWE by 25 percent, which is much more than what is explained by the income controls alone. Finally, when including all control variables the IWE becomes 0.134, which is a reduction of the IWE by less than 30 percent when starting from the baseline with only age controls. Thus, the explanatory variables explain only up to 30 percent of the IWE, which is considerably less than in C&H where income alone explains more than 50 percent of the IWE, and all variables together explain nearly 2/3 of the IWE.

## 5.3 Wealth across three generations

In this subsection, we use the child-parents-grandparents (CPG) sample, which enables us to analyse wealth across three generations. The children and parents are younger in the CPG subsample than in the full CP sample. For comparison, we therefore start by running a simple regression of child wealth with respect to parental wealth for this subsample. The result is reported in column 1 of Table 5 and gives an age-adjusted elasticity of 0.177, which is only a little smaller that the elasticity obtained when estimating

the relationship on the full sample (column 4 of Table 2).

It is important to know whether the relationship between child wealth and parental wealth is stable over time/generations. There may be long run forces that reduce or increase the strength of the relationship, and recent research has documented substantial changes over the long run in top income shares, in the relative importance of capital and labor income at the top of the income distribution, and in the evolution of inheritance (Atkinson, Piketty and Saez, 2011; Piketty, 2011). In column 2, we report the relationship between parental wealth and grandparental wealth in the data set. The age-adjusted child-parents wealth elasticity from this estimation is 0.160, which is only slightly lower than the IWE when running the regression on the child-parents pairs in the sample, and the standard deviations on the estimates from these two regressions are small and completely identical. This indicates that the underlying intergenerational relationship is quite stable.

Next, we look separately at the correlation between children and grandparents. From a structural relationship only relating generation $g$ to generation $g-1$ such as $w_g = \beta_0 + \beta_1 w_{g-1} + \varepsilon$, we would expect that a regression of $w_g$ on $w_{g-2}$ would give a coefficient equal to $(\beta_1)^2$. From the IWE estimates in columns 1 and 2, we should therefore expect a coefficient around 0.03 (i.e., $0.177^2 = 0.031$ and $0.160^2 = 0.026$). The estimated coefficient we obtain from regressing child wealth on grandparental wealth is 0.094 (column 3 of Table 5) and is therefore more than three times as large. One reason for this difference could be that the underlying stochastic processes relating wealth across generations have more memory than just one generation, for example because grandparents have a direct impact on their grandchildren (Solon, 2012). Another possible reason, which we pursue in the next subsection, is measurement error/omitted variable that creates a downward bias in the estimated coefficient.[11]

Column 4 of Table 5 reports the results from including both parental wealth and grandparental wealth in the estimation. When compared to the univariate relationships (columns 1 and 2), we see that the coefficient on parents fall a little, while the coefficient on grandparents falls by 1/3. If we consider a ten percent increase in the wealth of grandparents then this regression predicts nearly one percent higher wealth of the grandchildren, which is again three times as high as the prediction we obtain from a standard estimation of IWE exploiting data from only two generations. In Table 6, we add the same controls as we did in Table 4 plus dummy variables for number of cousins and for number of living grandparents in 2011.[12] Introducing the demography variables in isolation reduces the coefficients on parents and grandparents a little (going from column 1 to column 2). The effects on the coefficients are larger when

[11] If, for example, only 50 percent of the measured variation in wealth is governed by the true variation while the rest is noise then we would estimate a $\beta_1-$ coefficient around 0.2, when the structural coefficient is 0.4. The true coefficient on grandparents would be $0.4\hat{}2 = 0,16$, but we would estimate $0.16 * 50\% = 0.08$, which is higher than $0.2\hat{}2 = 0.04$.

[12] If grandparents transfer money to grandchildren then the effect of many cousins would be less money received on average per grandchild. By including dummy variables for number of cousins, we are only exploiting the variation within families of same size (measured by the number of cousins). We include dummy variables for number of living grandparents in 2011 because grandparents may die after we have observed their wealth in 1983-85, implying that parents and children may inherit their wealth.

including education and financial composition dummies (column 3), as was also the case when we looked at the relationship across only two generations in Table 4. The effect of grandparental wealth on child wealth is again three times as big as the predicted effect from a standard two-generation estimation of the IWE.

## 5.4 Using grandparental wealth as instrument

Attenuation bias caused my measurement problems has been a main issue in the intergenerational income mobility literature since the influential contribution by Solon (1992). Measurement problems may for example be caused by response errors in surveys and by transitory components in income implying that current income is a pure measure of the permanent income. The standard method to reduce the influence of the transitory component is to compute three or five year averages, as we have also done in our measurement of wealth. However, this method may only remove a small part of the attenuation bias if the transitory component has some persistence. For example, Mazumder (2005) provides simulations suggesting that it may require an average over 20 to 30 years in order to bring the attenuation factor down to 90 percent.

The theory in Section 2 illustrates cases where child wealth only depends on parental wealth and where ordinary least squares estimation of the relationship provides a downward biased estimate of the IWE, but where two-stage least squares estimation using wealth of grandparents as instrument for wealth of parents provides a consistent estimator. Table 7 reports results from such 2SLS estimations. The two first columns report 1st stage and 2nd stage results with only age dummies while the following columns report results when other variables are added as regressors. The strong correlation of parental wealth with grandparental wealth and F-test values well above 10 in all cases indicate that we do not have a 'weak instrument' problem in the 1st stage regressions.[13] The 2nd stage results give estimates of the IWE in the range 0.6-0.7, which is more than three times as high as the ordinary least squares estimates described in Subsection 4.1. For robustness, we have redone the age-dependency graphs in Figure 1a and 1b, but this time displaying the 2SLS estimates. This is done in Figure 2a and 2b, which for comparison also include the OLS estimates of Figure 1a and 1b. The graphs based on the 2SLS estimates have the same shapes as the graphs based on the OLS estimates, but of course with a more wide confidence interval. More importantly, the IWE estimates are consistently around 3 times as high at each age level.

[13] As a further test of the strength of the first stage, we have constructed a worst case scenario where grandparental wealth has no predictive power for parental wealth following the approach of Bound, Jaeger and Baker (1995). We achieve this by randomly assigning the observations of grandparental wealth to the observations in data. The resulting pseudo instrument is by construction uncorrelated with the endogenous regressor, yet, it retains the marginal distribution of the actual instrumental variable. Using the pseudo-instrument gives an IV estimate close to the OLS estimate.

## 5.5 Dynasty fixed effects

In Section 2, we also considered the possibility of a dynasty/social class fixed effect leading to the relationship

$$w_g^d = b_0^d + b_1 w_{g-1}^d + \varepsilon_g, \quad g = 1, 2, \tag{6}$$

where $w_g^d$ is the wealth level of generation $g$ in dynasty $d$, $b_0^d$ is the fixed effect of the dynasty and $\varepsilon_g$ is an error term. A direct estimation of equation (6) will provide an inconsistent estimate of $b_1$ because the assumption of strict exogeneity of the regressors is violated by construction when using the lagged dependent variable as regressor. The Within estimator is also biased and so is an estimation of the first differences:

$$\Delta w_g^d = b_1 \Delta w_{g-1}^d + \Delta\varepsilon_g, \tag{7}$$

where $\Delta w_g^d \equiv w_g^d - w_{g-1}^d$. However, our assumption that $\varepsilon$ is serially uncorrelated allows us to follow the empirical approach first suggested by Anderson and Hsiao (1982) and use grandparental wealth $w_{g-2}^d$ as an instrument for $\Delta w_{g-1}^d$. This avoids the bias of the lagged dependent variable and provides a consistent estimate of $b_1$. Table 8 shows the results from this exercise where we have allowed each child-parents-grandparents to have a unique constant $b_0^d$.[14] The first column reports the OLS estimate of $b_1$ under the assumption of a common constant $b_0^d = b_0$ for $d$, while the second column reports the first-differences 2SLS estimate of $b_1$. The table shows that the 2SLS estimate of the coefficient is only around 1/3 of the OLS estimate. Thus, allowing for fixed effects instead of a common constant reduces the estimate substantially, strongly suggesting more persistence in wealth formation across generations and that the standard measurement of the IWE is underestimating the level of intergenerational fluidity.

## 5.6 Robustness

In the empirical analysis, we have imposed the condition that both parents of a child should be alive in 2011 and that some grandparents should be alive in 1985. We have therefore allowed for the possibility that grandparents die between 1985 and 2011 and leave wealth to their heirs, and we have allowed for variation in the number of grandparents alive in 1985. In order to study the importance of these sample selection criteria, we have redone the baseline regressions for the subsample where all four grandparents are alive in 2011. The results are shown in Table 9. The basic child-parents wealth elasticity equals 0.17 (0.19 before), the basic child-grandparents ela

# 6 Concluding remarks

.... to be added

# A Additional description of the wealth data

...... Registration of company values of self-employed: In the period 1981 to 1985, firm assets such as buildings and operating fixture, equipment, machines, and cars are included. In 1981, buildings and operating fixture, equipment, machines, cars etc. are registered at 80 pct. of the cash value or the balance sheet book value. It is calculated as 75 pct. of the cash value in 1982, 70 pct. of the cash value in 1983-1988, and 60 pct. of the cash value in 1988-1996. In the period 1986 to 1996, the equity of the firm is computed separately and included in the assets of a self-employed. In addition to the above assets, the computation of firm equity also includes financial assets of the firm, inventory etc. and company debt is subtracted. From 1997 only firm assets are included in the assets of the owner while firm liabilities are included in the liabilities of the owner.

# References

[1] Anderson, T.W. and C. Hsiao (1982). "Formulation and Estimation of Dynamic Models Using Panel Data." Journal of Econometrics 18, 47–82.

[2] Arrondel, Luc (2009). "My Father Was Right: The Transmission of Values between Generations." PSE Working Paper 2009-12.

[3] Atkinson, Anthony B., Thomas Piketty and Emmanuel Saez (2011). "Top Incomes in the Long Run of History." Journal of Economic Literature, 49, 3–71.

[4] Becker, Gary and Nigel Tomes (1979). "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." Journal of Political Economy 87, 1153 - 89.

[5] Becker, Gary and Nigel Tomes (1986). "Human Capital and the Rise and Fall of Families." Journal of Labor Economics 4, 1-39.

[6] Björklund, Anders and Jännti, Markus (2009). "Intergenerational Income Mobility and the Role of Family Background." in W. Salverda, B. Nolan and T. Smeeding, eds., Oxford Handbook of Economic Inequality, Oxford: Oxford University Press, 491-521.

[7] Black, Sandra and Paul J. Devereux (2011). "Recent Developments in Intergenerational Mobility." In: Ashenfelter, Orley, Card, David (Eds.), Handbook of Labor Economics. Vol. IVb. North-Holland, Amsterdam, 1487–1541.

[8] Bound, John, David A. Jaeger and Regina M. Baker (1995). "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogeneous Explanatory Variable is Weak." Journal of the American Statistical Association 90, 443-450.

[9] Charles, Kerwin K., Hurst, Erik (2003). "The Correlation of Wealth Across Generations." The Journal of Political Economy 111, 1155–1182.

[10] Chetty, Raj, John N. Friedman and Tore Olsen and Luigi Pistaferri (2011). "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records." Quarterly Journal of Economics, 126(2), 749-804.

[11] European Commission (2012). "Alert mechanism report." http://ec.europa.eu/economy_finance/economic_governance/documents/alert_mechanism_report_2012_en.pdf.

[12] Goldberger, Arthur S. (1989). "Economic and Mechanical Models of Intergenerational Transmission." American Economic Review 79, 504-513.

[13] Haider, Steven and Gary Solon (2006). "Life-Cycle Variation in the Association between Current and Lifetime Earnings." American Economic Review 96, 1308–1320.

[14] International Monetary Fund (2012). "World Economic Outlook (WEO): Growth Resuming, Dangers Remain." April 2012 (Washington).

[15] Kleven, Henrik Jacobsen, Martin Knudsen, Claus Thustrup Kreiner, Søren Pedersen and Emmanuel Saez (2011). "Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark." Econometrica 79, 651–692.

[16] Kreiner, Claus Thustrup, David Dreyer Lassen and Søren Leth-Petersen (2012). "Measuring the Accuracy of Survey Responses using Administrative Register Data: Evidence from Denmark." Manuscript prepared for "Improving the Measurement of Household Consumption Expenditures," eds. Christopher Carroll, Thomas Crossley and John Sabelhaus. Forthcoming NBER Book Series Studies in Income and Wealth.

[17] Kolodziejczyk, Christophe (2011). "The Intergenerational Correlation of Wealth and Portfolio Allocation." Unpublished working paper.

[18] Leth-Petersen, Søren (2010). "Intertemporal Consumption and Credit Constraints: Does Consumption Respond to An Exogenous Shock to Credit?" American Economic Review, 100(3), pp. 1080-1103.

[19] Mazumder, Bhashkar (2005). "Fortunate sons: new estimates of intergenerational mobility in the US using social security earnings data." Review of Economics and Statistics 87, 235-255.

[20] Piketty, Thomas (2000). "Theories of Persistent Inequality and Intergenerational Mobility." In: Atkinson, A.B. and F. Bourguignon (Eds.), *Handbook of Income Distribution*. Vol I. North-Holland, Amsterdam, 430–476.

[21] Piketty, Thomas (2011). "On the Long-Run Evolution of Inheritance: France 1820–2050." *Quarterly Journal of Economics* 126, 1071-1131.

[22] Solon, Gary (1992). "Intergenerational income mobility in the United States." *American Economic Review* 82, 393–408.

[23] Solon, Gary (1999). "Intergenerational mobility in the labor market." In: Ashenfelter, Orley, Card, David (Eds.), *Handbook of Labor Economics*. Vol III. North-Holland, Amsterdam, 1761–1800.

[24] Solon, Gary (2012). "Theoretical Models of Inequality Transmission across Multiple Generations." Working paper. Michigan State University.

[25] Stiglitz, Joseph E. (1969). "Distribution of Income and Wealth Among Individuals." *Econometrica* 37, 382-397.

Table 1: Summary Statistics

| | Child-Parent (CP) Sample | | Child-Parent-Grandparent (CPG) Sample | | |
|---|---|---|---|---|---|
| | Children 2009-2011 | Parents 1997-1999 | Children 2009-2011 | Parents 1997-1999 | Grandparents 1983-1985 |
| Age | 33.9 | 48.5 | 23.4 | 35.0 | 47.1 |
| | (8.2) | (7.6) | (2.3) | (2.2) | (5.1) |
| Years of education[a] | 12.9 | 12.3 | 11.0 | 12.6 | 8.9 |
| | (2.4) | (2.6) | (1.7) | (1.9) | (2.3) |
| Labor income | 256,025 | 288,873 | 129,398 | 282,080 | 247,334 |
| | (207,597) | (174,194) | (103,486) | (135,885) | (133,734) |
| Gross income | 326,867 | 396,203 | 172,004 | 358,322 | 371,134 |
| | (233,078) | (345,531) | (92,620) | (127,748) | (163,437) |
| Share owning stocks | 0.22 | 0.26 | 0.15 | 0.14 | 0.07 |
| | (0.41) | (0.36) | (0.36) | (0.28) | (0.14) |
| Share owning property[b] | 0.47 | 0.59 | 0.11 | 0.59 | 0.41 |
| | (0.50) | (0.34) | (0.32) | (0.41) | (0.23) |
| Share owning bonds[c] | 0.06 | 0.06 | 0.03 | 0.01 | 0.11 |
| | (0.23) | (0.19) | (0.16) | (0.08) | (0.20) |
| Share self-employed[d] | 0.04 | 0.08 | 0.01 | 0.06 | 0.11 |
| | (0.19) | (0.20) | (0.10) | (0.17) | (0.16) |
| Value of assets | 655,952 | 923,242 | 95,113 | 584,975 | 828,626 |
| | (1,988,246) | (2,413,591) | (367,708) | (883,319) | (1,094,917) |
| Value of liabilities | 636,575 | 596,021 | 129,221 | 605,768 | 550,996 |
| | (1,702,573) | (1,207,167) | (343,321) | (714,934) | (822,594) |
| Net wealth | 19,377 | 327,220 | -34,108 | -20,793 | 277,894 |
| | (1,117,250) | (1,928,688) | (215,500) | (536,283) | (701,004) |
| Percentiles of wealth | | | | | |
| 20th | -232,694 | -89,456 | -90,626 | -185,697 | -10,472 |
| 40th | -68,913 | 47,364 | -24,166 | -87,330 | 111,541 |
| 60th | 8,240 | 255,708 | 3,969 | -14,398 | 273,039 |
| 80th | 147,696 | 585,444 | 28,575 | 123,348 | 499,783 |
| Observations | 1,155,564 | 1,155,564 | 97,438 | 97,438 | 97,438 |

Notes: The table reports mean values and standard deviations (in parentheses) of the variables. Age, education and ownership variables are as of 2011 for children, 1998 for parents and 1983 for grandparents.

Child-Parent (CP) sample: Children are aged 21-51 in 2011, both parents are alive in 2011 and aged 21-66 in 1998, and children are neither immigrants nor descendants of immigrants.

Child-Parents-Grandparents (CPG) sample: Children are aged 19-51 in 2011, both parents are alive in 2011 and aged 19-38 in 1998, children are neither immigrants nor descendants of immigrants, and have at least one grandparent alive in 1983. Parent variables are averages of biological parents. Grandparent variables are averages of biological grandparents. All monetary variables are measured in DKK and deflated with 2010 prices.

a) Measures years of completed education. The variable is based on 2010 data for children.

b) Property ownership dummy for grandparents is based on 1987 data.

c) Bond ownership dummy for grandparents is based on 1995 data.

d) Self-employed dummy for children is based on 2010 data.

Table 2: Child-Parent Wealth Elasticity

| | (1)<br>Child wealth<br>2009-2011<br>log | (2)<br>Child wealth<br>2009-2011<br>log | (3)<br>Child wealth<br>2009-2011<br>IHS | (4)<br>Child wealth<br>2009-2011<br>IHS |
|---|---|---|---|---|
| Parental per cap. wealth (1997-1999, log) | 0.379 | 0.268 | | |
| | (0.002) | (0.002) | | |
| Parental per cap. wealth (1997-1999, IHS) | | | 0.268 | 0.190 |
| | | | (0.002) | (0.001) |
| Child age dummies | | X | X | X |
| Parental age (avg., rounded) dummies | | X | X | X |
| Observations | 385,338 | 385,338 | 385,338 | 1,155,564 |
| R-squared | 0.086 | 0.267 | 0.268 | 0.102 |
| Adj. R-squared | 0.086 | 0.267 | 0.267 | 0.102 |

Notes: All regressions are on the CP sample described in table 1. Difference in number of observations from (1)-(2) to (4) is that the log transform in (1)-(2) excludes observations of zero or negative wealth. The regression in (3) is run on the same sample as in (1)-(2). Robust standard errors are reported in the parentheses.

Table 3: Child-Parent Wealth Elasticity and the Importance of Own and Parental Income.

| | (1) Baseline | (2) | (3) | (4) Par. age <= 48 | (5) Par. age > 48 |
|---|---|---|---|---|---|
| Parental per cap. wealth 1997-1999 | 0.190 | 0.186 | 0.175 | 0.165 | 0.185 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) |
| Child income 2009-2011 | | 1.354 | 1.134 | -0.299 | 2.772 |
| | | (0.022) | (0.022) | (0.027) | (0.035) |
| Parental per cap. income 1997-1999 | | | 2.215 | 3.057 | 1.276 |
| | | | (0.029) | (0.041) | (0.042) |
| Child age dummies | X | X | X | X | X |
| Parental age (avg., rounded) dummies | X | X | X | X | X |
| Observations | 1,155,564 | 1,155,564 | 1,155,564 | 611,918 | 543,646 |
| R-squared | 0.102 | 0.105 | 0.110 | 0.147 | 0.067 |
| Adj. R-squared | 0.102 | 0.105 | 0.110 | 0.147 | 0.067 |

Notes: Dependent variable is the average 2009-2011 wealth of children. The IHS transformation is used on wealth and income variables. All regressions are on the CP sample described in table 1. Robust standard errors are reported in the parentheses.
Columns (4) and (5) split the sample by median parental age (50 in 1999).

Table 4: Child-Parent Wealth Elasticity and the Importance of Controls.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Parental per cap. wealth 1997-1999 | 0.190 | 0.183 | 0.173 | 0.141 | 0.133 | 0.134 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Child income 2009-2011 | | | | | | -0.407 |
| | | | | | | (0.023) |
| Parental per cap. income 1997-1999 | | | | | | 0.671 |
| | | | | | | (0.033) |
| No. of siblings dummies[a] | | X | | | X | X |
| Years of schooling dummies[b] | | | X | | X | X |
| Financial composition dummies[c] | | | | X | X | X |
| Child age and parental age dummies | X | X | X | X | X | X |
| Observations | 1,155,564 | 1,155,564 | 1,155,564 | 1,155,564 | 1,155,564 | 1,155,564 |
| R-squared | 0.102 | 0.105 | 0.111 | 0.155 | 0.161 | 0.162 |
| Adj. R-squared | 0.102 | 0.105 | 0.111 | 0.155 | 0.161 | 0.162 |

Notes: Dependent variable is the average 2009-2011 wealth of children. The IHS transformation is used on wealth and income variables. All regressions are on the CP sample described in table 1. Robust standard errors are reported in the parentheses.
a) No. of siblings is calculated as the average of mother's no. of children and father's no. of children (rounded).
b) Years of schooling dummies for both child and parent are included. Years of schooling is recorded as of 2010.
c) Financial composition consists of dummies for both child and parents for homeownership, stock ownership, bonds ownership, and for being self-employed. All dummies are included for both child and parents. Parents are noted as owning stocks, say, if at least one parent owns stocks. Self-employment status for children is taken in 2010.

Table 5: Child-Parent and Child-Grandparent Wealth Elasticities

| | (1)<br>Child wealth<br>2009-2011 | (2)<br>Parental per cap.<br>wealth<br>1997-1999 | (3)<br>Child wealth<br>2009-2011 | (4)<br>Child wealth<br>2009-2011 |
|---|---|---|---|---|
| Parental per cap. wealth 1997-1999 | 0.177 | | | 0.168 |
| | (0.003) | | | (0.003) |
| Grandparental per cap. wealth 1983-1985 | | 0.160 | 0.094 | 0.062 |
| | | (0.003) | (0.003) | (0.003) |
| Child age dummies | X | | X | X |
| Parental age (avg., rounded) dummies | X | X | | X |
| Grandparent age (avg., rounded) dummies | | X | X | X |
| Observations | 97,438 | 97,438 | 97,438 | 97,438 |
| R-squared | 0.194 | 0.055 | 0.161 | 0.200 |
| Adj. R-squared | 0.194 | 0.055 | 0.160 | 0.199 |

Notes: The IHS transformation is used on wealth variables. All regressions are on the CPG sample described in table 1. Robust standard errors are reported in the parentheses.

Table 6: Child-Parent and Child-Grandparent Wealth Elasticities, Including Various Control Variables.

| | (1) Child wealth 2009-2011 | (2) Child wealth 2009-2011 | (3) Child wealth 2009-2011 |
|---|---|---|---|
| Parental per cap. wealth 1997-1999 | 0.168 | 0.161 | 0.126 |
| | (0.003) | (0.003) | (0.003) |
| Grandparental per cap. wealth 1983-1985 | 0.062 | 0.054 | 0.038 |
| | (0.003) | (0.003) | (0.003) |
| No. of siblings dummies[a] | | X | X |
| Number of cousins dummies[b] | | X | X |
| Number of living grandparents in 2011 dummies[c] | | X | X |
| Years of schooling dummies[d] | | | X |
| Financial composition dummies[e] | | | X |
| Full set of age dummies | X | X | X |
| Observations | 97,438 | 97,438 | 97,438 |
| R-squared | 0.200 | 0.206 | 0.256 |
| Adj. R-squared | 0.199 | 0.205 | 0.255 |

Notes: The IHS transformation is used on wealth and income variables. All regressions are on the CPG sample described in table 1. Robust standard errors are reported in the parentheses.
a) No. of siblings is calculated as the average of mother's no. of children and father's no. of children (rounded).
b) No. of cousins is calculated as the average of the (at most) four grandparents' no. of grandchildren (rounded).
c) Al combinations of four dummies (one for each grandparent) denoting whether a grandparent is alive in 2011.
d) Years of schooling dummies for both child and parent are included. Years of schooling is recorded as of 2010.
e) Financial composition consists of dummies for both child and parents for homeownership, stock ownership, bonds ownership, and for being self-employed. All dummies are included for both child and parents. Parents are noted as owning stocks, say, if at least one parent owns stocks. Self-employment status for children is taken in 2010.

Table 7: Child-Parent Wealth Elasticities, 2SLS Estimation Using Grandparental Wealth As Instrument.

| | (1)<br>1st stage<br>IV | (2)<br>2nd stage<br>IV | (3)<br>1st stage<br>IV | (4)<br>2nd stage<br>IV | (5)<br>1st stage<br>IV | (6)<br>2nd stage<br>IV |
|---|---|---|---|---|---|---|
| Grandparental per cap. wealth 1983-1985 | 0.155 | | 0.151 | | 0.087 | |
| | (0.003) | | (0.003) | | (0.003) | |
| Parental per cap. wealth 1997-1999 | | 0.608 | | 0.603 | | 0.631 |
| | | (0.022) | | (0.023) | | (0.040) |
| Number of siblings dummies | | | X | X | X | X |
| Number of cousins dummies | | | X | X | X | X |
| Extra control variables | | | | | X | X |
| Observations | 97,438 | 97,438 | 97,438 | 97,438 | 97,438 | 97,438 |
| R-squared | 0.067 | . | 0.074 | . | 0.182 | 0.016 |
| Adj. R-squared | 0.067 | . | 0.073 | . | 0.181 | 0.015 |
| F-test (1st stage) | 2,097.88 | | 1,976.82 | | 700.42 | |
| F-test p.-val. (1st stage) | 0.0000 | | 0.0000 | | 0.0000 | |
| Partial R-sq (1st stage) | 0.0186 | | 0.0176 | | 0.0064 | |
| R-sq (1st stage) | 0.0670 | | 0.0739 | | 0.1819 | |
| Adj. R-sq (1st stage) | 0.0667 | | 0.0732 | | 0.1809 | |

Notes: The IHS transformation is used on wealth and income variables. All regressions are on the CPG sample described in table 1. Robust standard errors are reported in the parentheses. All regressions include both child and parental age dummies.
The extra control variables used are identical to the ones in regression (6) of table 4.

Table 8: Wealth Elasticities With Child Fixed Effects.

| | (1) OLS | (2) FD 2SLS |
|---|---|---|
| Parental per cap. Wealth 1997-1999 | 0.220 | 0.080 |
| | (0.003) | (0.005) |
| Observations | 97,438 | 97,438 |
| R-squared | 0.055 | . |

Notes: Dependent variable is the average 2009-2011 wealth of children. The IHS transformation is used on wealth variables. All regressions are on the CPG sample described in table 1. Robust standard errors are reported in the parentheses. Estimation details: OLS estimation of child wealth regressed on parental wealth and a constant. FD 2SLS is the difference between child and parents regressed on the difference between parents and grandparents, where the difference between parents and grandparents is instrumented using grandparental (i.e., double lagged) per capita wealth.

Table 9: Child-Parent Wealth Elasticities, Robustness Check Using Families With 4 Living Grandparents.

| | (1) OLS | (2) OLS | (3) 1st stage IV | (4) 2nd stage IV |
|---|---|---|---|---|
| Parental per cap. wealth 1997-1999 | 0.165 | | | 0.521 |
| | (0.006) | | | (0.062) |
| Grandparental per cap. wealth 1983-1985 | | 0.059 | 0.137 | |
| | | (0.008) | (0.009) | |
| Child age dummies | X | X | X | X |
| Parental age (avg., rounded) dummies | X | | X | X |
| Grandparent age (avg., rounded) dummies | | X | | |
| Observations | 24,383 | 15,091 | 15,091 | 15,091 |
| R-squared | 0.193 | 0.166 | 0.069 | 0.042 |
| Adj. R-squared | 0.192 | 0.163 | 0.068 | 0.040 |
| F-test (1st stage) | | | 246.61 | |
| F-test p.-val. (1st stage) | | | 0.0000 | |
| Partial R-sq (1st stage) | | | 0.0147 | |
| R-sq (1st stage) | | | 0.0692 | |
| Adj. R-sq (1st stage) | | | 0.0675 | |

Notes: The IHS transformation is used on wealth variables. The results of the regression in column 1 correspond to column 4 of table 2 but now using only families with 4 living grandparents in 2011. Likewise the results in column 2 correspond to column 3 of table 5 and results in column 3 and 4 correspond to column 1 and 2 of table 7. Robust standard errors are reported in parentheses.

Figure 1a. OLS estimates of the IWE as a function of child age.

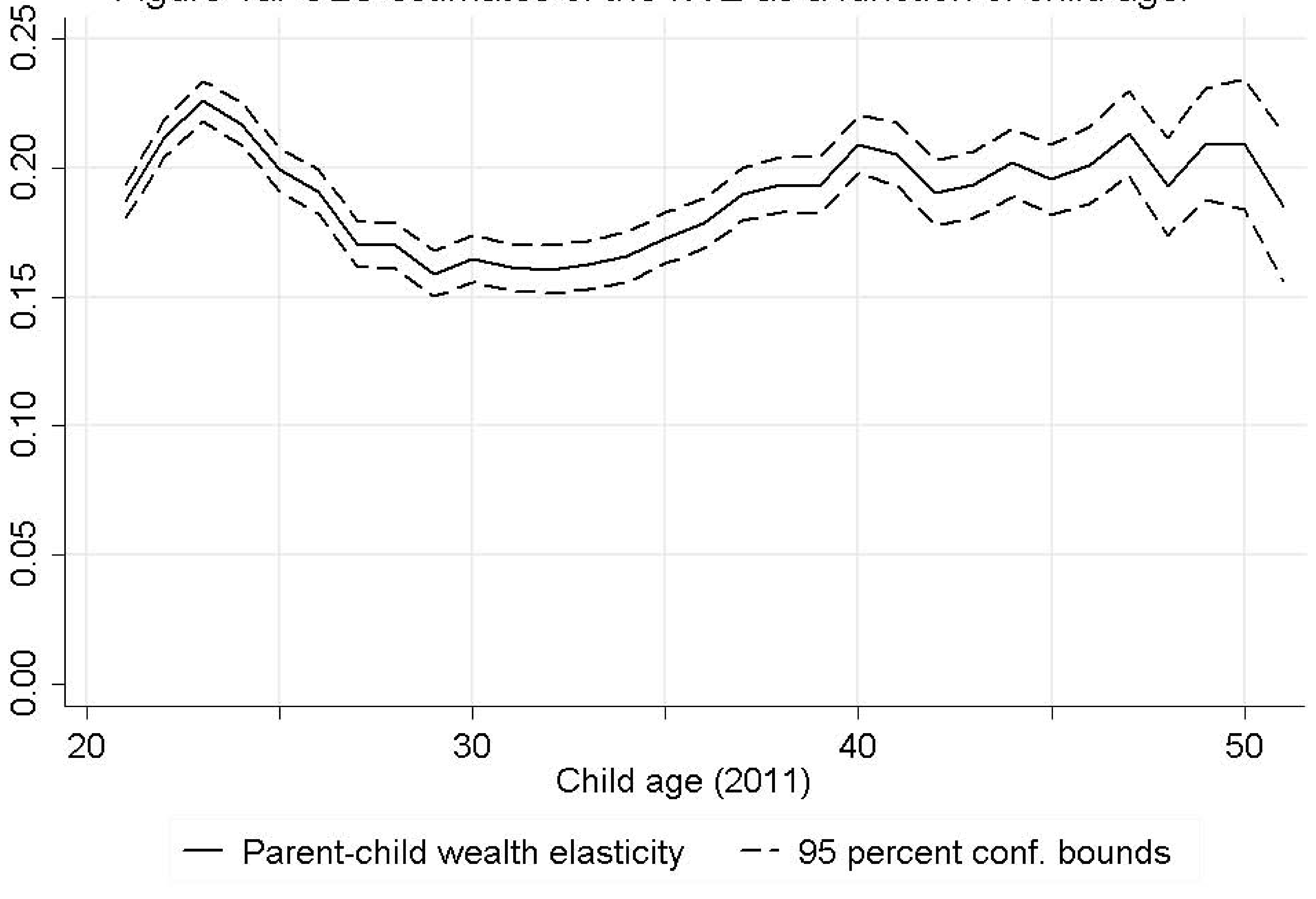

Figure 1b. OLS estimates of the IWE as a function of parental age.

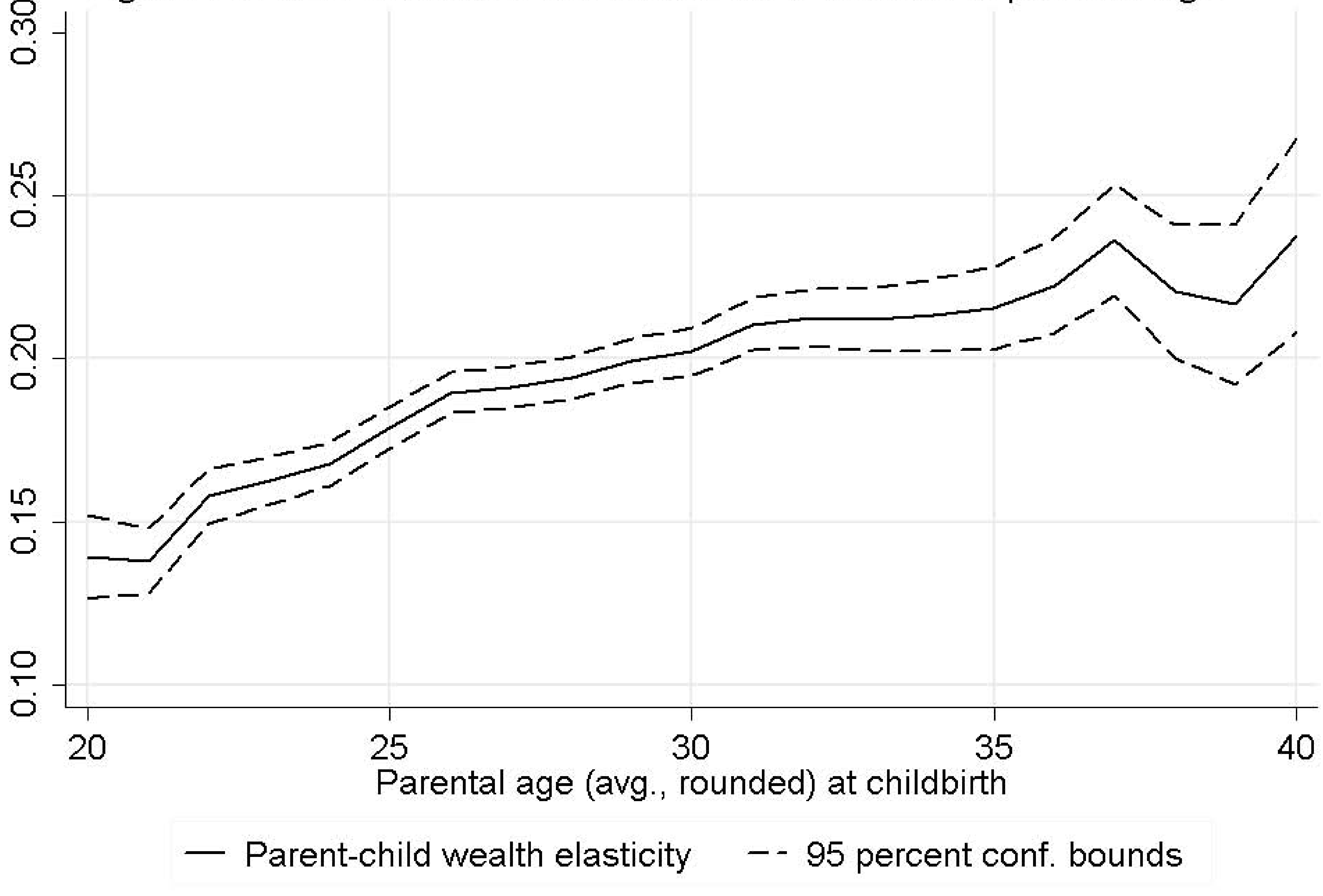

Figure 2a. 2SLS and OLS estimates of the IWE as a function of child age.

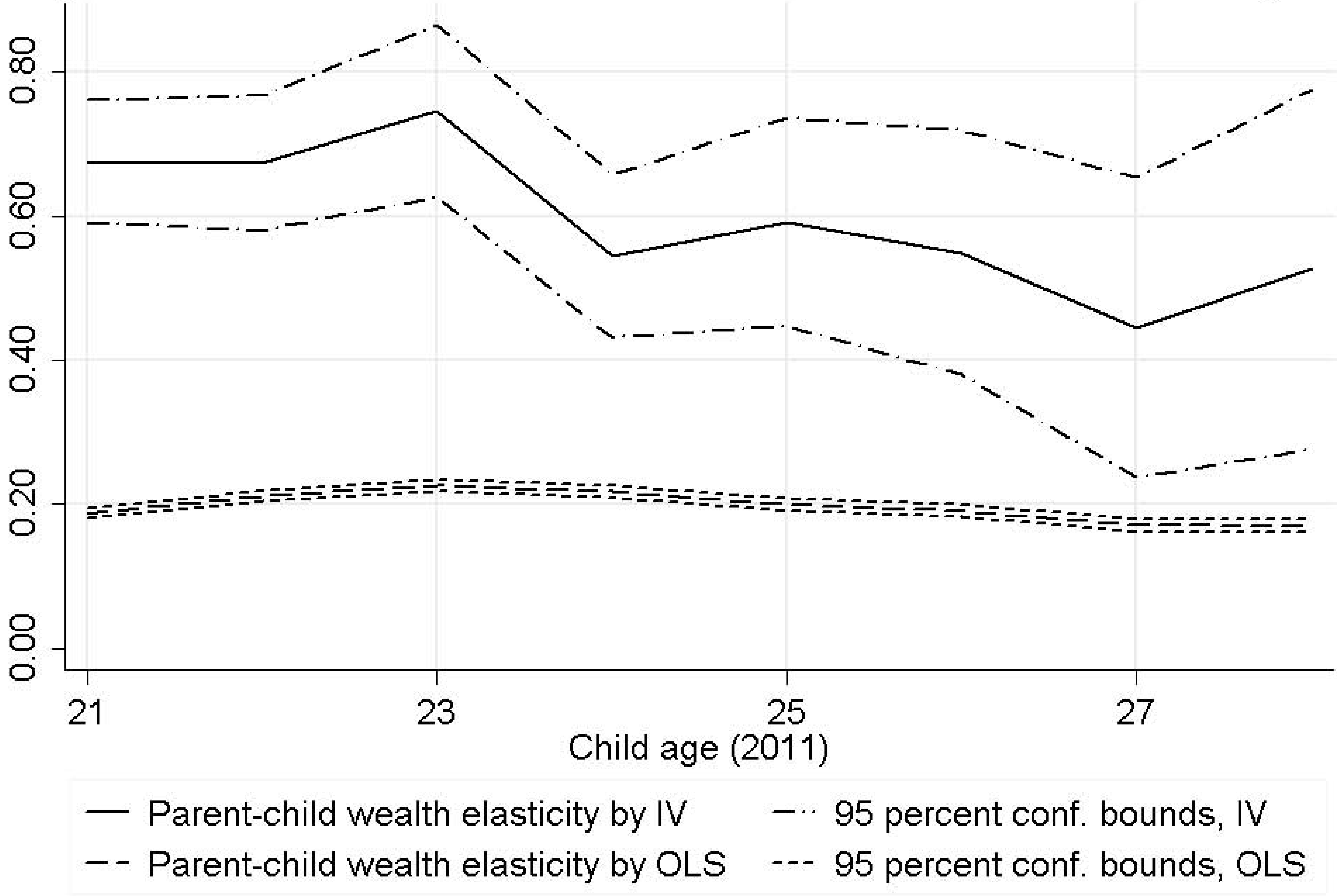

Figure 2b. 2SLS and OLS estimates of the IWE as a function of parental age.