Steady-State Assumptions in Intergenerational Mobility Research

Martin Nybom[*], Jan Stuhler[†]

March 7, 2019

1 Introduction

Research on on the transmission of socio-economic status from parents to their children gained prominence in the second half of the 20th century, first in sociology and subsequently in economics. Today, the literature on intergenerational mobility is integral to the social sciences and draws significant interest from policy circles. Much of the early work in economics was theoretical in nature, relating intergenerational transmission mechanisms to steady-state levels of inequality and intergenerational mobility (e.g. Becker and Tomes, 1979; Loury, 1981). In the following decades the center of gravity shifted to empirical estimation of descriptive measures of intergenerational mobility across countries and time, as well as causal identification of intergenerational effects.[1]

Tony Atkinson contributed to many different aspects of this intergenerational literature, including its welfare foundations, the measurement of mobility, data collection, and the analysis of mobility in the United Kingdom. Many of these contributions were well recognized, and had a profound influence on subsequent work on these topics. A prime example is his series of work on the social welfare foundations of mobility measurement, which includes Atkinson (1981), reprinted as Atkinson (1983), and Atkinson and Bourguignon (1982). Atkinson's proposed welfare function, a generalization of the cross-sectional (or one-period) case discussed in Atkinson (1970) to the intergenerational (or multi-period) context, has served as a foundation for subsequent work on the topic, and for the derivation of various "mobility dominance results" (see Atkinson and Bourguignon 1982). In other influential work, Atkinson (1980) noted that short-run income measures were subject to transitory fluctuations, which would attenuate estimates of the intergenerational associations in more permanent measures of living standards. The empirical relevance of this insight was later demonstrated by Solon (1992) and Zimmerman (1992). Refined with the life-cycle consideration developed in Jenkins (1987), it still plays an

[*]Institute for Evaluation of Labour Market and Education Policy (IFAU), SOFI and UCLS (email: martin.nybom@ifau.uu.se)

[†]Universidad Carlos III de Madrid and SOFI (email: jan.stuhler@uc3m.es)

[1]See the reviews by Black and Devereux (2011) and Jäntti and Jenkins (2014).

important role in the measurement of income mobility today.

However, not all of Atkinson's contributions had such direct influence on the subsequent literature. In work that turned out to become much less well known, Atkinson and Jenkins (1984) studied the importance of steady-state assumptions in the context of intergenerational mobility research. While noting that such assumptions can help to identify the parameters of a structural intergenerational equation system in settings with incomplete data, they also show that they are purchased at a cost. This research, however, received little attention at the time and is perhaps even less known about among mobility scholars today. Most theoretical and empirical work has continued to examine the relationship between different transmission mechanisms and implied steady-state levels of intergenerational mobility, explicitly or implicitly assuming that the distributional moments (such as variances and covariances) of income and other variables are at their long-run steady-state values. Only few studies reflect on the transition paths between those steady states (e.g. Davies et al., 2005).

We believe that Atkinson and Jenkins' insights have great relevance for intergenerational research today. We therefore review the role of steady-state assumptions in the literature, show how their arguments relate to more recent strands of the literature, and extend them in different directions. In Section 2, we review the main insights of Atkinson and Jenkins (1984) with respect to the identification of parameters in a structural intergenerational equation system. The scenario considered by Atkinson and Jenkins was one of *incomplete data*, in which a variable was potentially observable, but not in fact observed in the actual data under consideration. In such settings, researchers often assume that the variable's distributional moments are in a steady-state equilibrium, allowing them to replace the missing moments from other sources. Atkinson and Jenkins noted that there is no good reason for that assumption in intergenerational data spanning across *generations*, and illustrated the potential consequences of its violation. We review these arguments and illustrate them using Swedish register data, and in the context of rising income inequality in many developed countries.

Atkinson and Jenkins did not argue against the use of steady-state assumptions – while considering them to be a poor substitute for the collection of better data, they did acknowledge their potential usefulness. However, they advised to use them only with caution, after careful consideration of the structure of the particular model. Unfortunately, this advice seems to have rarely been heeded. In our review, we point at two factors that may have limited the influence of this argument on subsequent work. First, Atkinson and Jenkins made their argument at a time when researchers were limited to small survey data sets and therefore frequently forced to combine multiple data sources. Second, they placed their arguments in the context of simultaneous-equations models. While being the cornerstone of inequality analysis in much of the second half of the 20th century, the approach has been less popular in recent decades.

While the insights from Atkinson and Jenkins (1984) may therefore not apply directly, we argue that they do extend to the current literature. In empirical work, key variables are still frequently missing or not well observed, motivating the use of two-sample instrumental variable methods or correction factors imported from external sources (Björklund and Jäntti, 1997). On the theoretical side, the switch from "mechanical" to "economic" models of intergenerational transmission along the lines of Becker and Tomes (1979) did not reduce our dependence on steady-state assumptions.[2] Arguably, this dependency has become even less transparent, as empirical moments and variables tend to be less closely linked in calibration or estimation. As in the earlier literature, the role of steady-state assumptions is rarely evaluated and often not made explicit.

In Section 3, we extend Atkinson and Jenkins' insights to a more severe type of data limitation, in which one variable of the transmission system is *inherently* unobservable for *any* generation. This latent-factor scenario has played a central role in early sociological work (see Goldberger 1972) as well as in standard economic models (such as Becker and Tomes 1979). It has gained empirical interest in the recent literature on *multi*generational persistence, as it would rationalize the observation that status correlations decay more slowly across generations than a naive extrapolation from parent-child correlations would suggest (see Clark and Cummins 2012, Stuhler 2012, Clark 2014, Lindahl et al. 2015, Anderson et al. 2018). The observation of data across more than two generations may also aid in the identification of a transmission model, including its inherently unobserved component (Becker and Tomes 1979, Goldberger 1989). Applying this idea to the recent multigenerational literature, Braun and Stuhler (2018), Neidhöfer and Stockhausen (forthcoming), and Adermon et al. (2018) find comparatively high rates of persistence in the latent factor, suggesting that inequalities are more persistent than parent-child correlations in observable status would reveal.[3] However, this interpretation relies, again, critically on steady-state assumptions.

We argue that steady-state assumptions play a more critical role in this latent-factor than the incomplete-data scenario considered by Atkinson and Jenkins (1984). In the latter, the imposition of steady-state assumptions is only a makeshift measure to achieve identification using data sources that happen to be incomplete, with the collection of richer data being the obvious alternative. But in a scenario with inherently unobservable variables, better data will not be a solution. The steady-state assumption is then less dispensable, and Atkinson and Jenkins' advice to carefully consider their implications particularly relevant. Against this background, we discuss a potential trade-off in the use of "horizontal moments" (such as sibling and cousin correlations) and "vertical moments" (such as parent-child of grandparent-grandchild correlations), which we expect to play an important role in future research on the topic. A horizontal perspective has advantages, as with respect to the measurement and the number of moments that can be considered (Björklund et al. 2009, Hällsten 2014, Adermon et al. 2016, Collado et al. 2018). We note that horizontal moments are also less sensitive to steady-state assumptions, because they can be measured at the same time and within the same generation.

[2] Studies such as Becker and Tomes (1979) or Loury (1981) emphasized the dynamics of *individual incomes* within families in response to individual shocks, but were fairly silent about the transitional dynamics of (functions of) population moments, such as the intergenerational elasticity of income, in response to structural changes.

[3] One can think of this latent factor as a broad measure of underlying ability or *potential* socio-economic status, although it is not always clearly defined.

Finally, in Section 4 we study how deviations from the steady-state assumption affect the *interpretation* of empirical evidence. Research on mobility trends over time or differences across countries often makes an implicit steady-state assumption when interpreting the evidence at hand, associating variation in mobility with variation in contemporaneous policies or institutions. But as we show in Nybom and Stuhler (2014), changes in *previous* generations can be important determinants of contemporaneous shifts in mobility as well. Recent changes in inequality and mobility may be less closely related to concurrent changes in the economic environment than commonly thought. Moreover, mobility trends generated by structural changes can be non-monotonic, complicating their interpretation further. A static interpretation may thus lead to inaccurate conclusions if the system is in fact not in steady state.

Transitional dynamics are of particular importance in intergenerational research since even a single transmission step – one generation – corresponds to a long time period. Different types of shocks to the economic environment can trigger different transitional dynamics, which in turn suggests that the dynamic pattern can be quite informative about the economic environment. This insight is in line with Davies et al. (2005), who note that the observation of mobility trends may help us to distinguish between alternative causes of rising cross-sectional inequality. The emphasis on transitional dynamics seems also to be in line with earlier work of Atkinson, who for example considered them to be a key component in research on growth models: "*If we throw away information about the time dimension, we are reducing still further our limited understanding of the relationship between these models and the real world*" (Atkinson 1969). While Atkinson made this statement in the context of the growth literature (see Aaberge et al. 2017), we believe that a more explicit focus on transitional dynamics and non-stationarity would also be fruitful for research on the intergenerational transmission of socio-economic status.

2 Steady-State Assumptions in Intergenerational Research

Atkinson and Jenkins (1984) illustrate how steady-state assumptions can be used to identify the parameters of a simultaneous-equations system, as well as the consequences when those steady-state assumptions fail to hold. While the simultaneous-equations approach is less frequently used today, their insights extend, directly or indirectly, to many other settings in the intergenerational context, where key variables are frequently missing, poorly observed or inherently unobservable. In this section, we review and discuss the main insights of Atkinson and Jenkins (1984) using a simple intergenerational model. We then extend the discussion to the recent literature on *multi*generational mobility, which estimates the strength of latent and higher-order transmission mechanisms with data that span more than two generations.

2.1 Identification of an Intergenerational Model

We assume a simple *causal* model of income determination and intergenerational transmission in the tradition of the simultaneous-equations approach developed by Conlisk (see Conlisk, 1969, 1974a), as also considered by Atkinson and Jenkins (1984). While such a model does not explicitly address optimizing behavior as in Becker and Tomes (1979) or Solon (2004), the mechanistic pathways represented by the structural equations can be derived from an underlying utility-maximization framework (see Goldberger, 1989). The model has two structural equations for income and education (or alternatively, ability), denoted y_t and e_t for generation t, which are determined recursively across generations according to

$$y_t = \gamma_t y_{t-1} + \rho_t e_t + u_t \tag{1a}$$

$$e_t = \lambda_t e_{t-1} + v_t. \tag{1b}$$

The parameters $\{\gamma_t, \rho_t, \lambda_t\}$ are assumed to be non-negative and (to ensure stability) strictly smaller than one.[4] The stochastic terms $\{u_t, v_t\}$, which can be interpreted as labor-market and human-capital luck, are assumed uncorrelated with each other, with the other variables, and over time, and have constant zero means. We assume that y_t and e_t are measured as trendless indices with constant mean zero, such that we do not need to include constants. Hence, own income is assumed to be determined by parental income, own education, and by chance, while own education is assumed to be influenced by parental education and by chance. This model is a simplified version of the model considered by Atkinson and Jenkins (1984), in which also child education depends on parental income, to facilitate parts of the presentation. Another deviation from their model is the addition of time subscripts t to the parameters, thus potentially allowing for generation-specific strengths of the different transmission mechanisms.

However, we start out by dropping the t-subscripts on the parameters, in order to briefly review the main insights from Atkinson and Jenkins (1984). We thus assume that the structural parameters of the system are all time invariant. If $\sigma_{y_t y_t}$ denotes the variance of y_t, $\sigma_{y_t y_{t-1}}$ the covariance of y_t and y_{t-1}, and so on, this simple model implies the following covariances and variances in observed variables y_t and e_t, which one can use to solve for the unkown parameters γ, ρ, λ:

$$\sigma_{e_t e_t} = \lambda^2 \sigma_{e_{t-1} e_{t-1}} + \sigma_{v_t v_t} \tag{2a}$$

$$\sigma_{y_t y_t} = \gamma^2 \sigma_{y_{t-1} y_{t-1}} + \rho^2 \lambda^2 \sigma_{e_{t-1} e_{t-1}} + 2\gamma\rho\lambda \sigma_{y_{t-1} e_{t-1}} + \rho^2 \sigma_{v_t v_t} + \sigma_{u_t u_t} \tag{2b}$$

$$\sigma_{e_t y_t} = \gamma\lambda \sigma_{y_{t-1} e_{t-1}} + \rho\lambda^2 \sigma_{e_{t-1} e_{t-1}} + \rho \sigma_{v_t v_t} \tag{2c}$$

$$\sigma_{e_t e_{t-1}} = \lambda \sigma_{e_{t-1} e_{t-1}} \tag{2d}$$

$$\sigma_{y_t y_{t-1}} = \gamma \sigma_{y_{t-1} y_{t-1}} + \rho\lambda \sigma_{y_{t-1} e_{t-1}} \tag{2e}$$

$$\sigma_{e_t y_{t-1}} = \lambda \sigma_{e_{t-1} y_{t-1}} \tag{2f}$$

$$\sigma_{y_t e_{t-1}} = \gamma \sigma_{y_{t-1} e_{t-1}} + \rho \sigma_{e_t e_{t-1}}. \tag{2g}$$

The model is overidentified, with seven equations and only five unknowns.[5] If y_{t-1}, y_t, e_{t-1},

[4] Jenkins (1982) discusses stability conditions for systems of stochastic linear difference equations with constant coefficients, Conlisk (1974b) derives stability conditions for systems with random coefficients.

[5] In addition to λ, ρ, and γ, the variances of the error terms $\sigma_{v_t v_t}$ and $\sigma_{u_t u_t}$ are also unknown. The model in

and e_t were all observed for several families, the estimation of equations (1a) and (1b) would be straightforward. The need for imposing steady-state assumptions arise when any of those variables are unobserved. Atkinson and Jenkins (1984) consider a particular scenario, in which these variables are potentially observable, but at least one variable is missing in the data at hand: "*Our particular interest is in the situation where a full set of observations is not available. By this we do not mean that certain variables are inherently unobservable – a case which has received considerable attention ... Rather, the variables are potentially observable, but are not in fact observed in the particular data set that is being used.*"

In particular, consider a case of incomplete data in which income (y_t), education (e_t), and parental education (e_{t-1}) are all observed, but parental income (y_{t-1}) is either unobserved or poorly observed (i.e. the observed measure of y_{t-1} is contaminated by significant measurement error). This case is still relevant today, as few data sets contain *good* proxies for both offspring and parent lifetime incomes, and intergenerational measures can be very sensitive to measurement error in income (Nybom and Stuhler, 2017). For illustrative purposes we here only focus on the identification of γ, the (conditional) intergenerational income coefficient. The parameter can be written in terms of population moments:

$$\gamma = \frac{r_{y_t y_{t-1}} - r_{y_t e_t} r_{y_{t-1} e_t}}{1 - r^2_{y_{t-1} e_t}} \left(\frac{\sigma_{y_t}}{\sigma_{y_{t-1}}} \right), \tag{3}$$

where we use the notation r_{XY} for the correlation between two variables X and Y, and σ_X for the standard deviation of X. The details of the derivations are outlined in Appendix A.1. At this stage, γ depends on three population moments that are unobserved: $r_{y_t y_{t-1}}$, $r_{y_{t-1} e_t}$, and $\sigma_{y_{t-1}}$. However, since the model is overidentifed, we can use the additional, observed moments to substitute out both $r_{y_t y_{t-1}}$ and $r_{y_{t-1} e_t}$, which yields:

$$\gamma = \frac{r_{y_t e_{t-1}} - r_{y_t e_t} r_{e_{t-1} e_t}}{r_{y_{t-1} e_{t-1}} (1 - r^2_{e_{t-1} e_t})} \left(\frac{\sigma_{y_t}}{\sigma_{y_{t-1}}} \right). \tag{4}$$

2.2 Identification With Steady-State Assumptions

The estimation of these types of parameters from simultaneous-equations models was a cornerstone in analyses of inequality and intergenerational mobility in much of the second half of the 20th century (see e.g. Conlisk 1969, Joreskog 1970, Goldberger 1972, Conlisk 1974a, Goldberger 1979, Goldberger 1989). As intergenerational data at the time did not include all variables of interest, however, researchers typically relied on the assumption that the distributional moments of those variables were in steady state. This allows the researcher to make further progress by deducing some of the moments without directly observing the variables. For example, if incomes of the parental generation are unobserved, a steady-state assumption allows the researcher to replace moments of the parental income distribution with moments of the corresponding distribution of their children. Moreover, the available data can be augmented by a second and closely

Atkinson and Jenkins (1984) has six unknowns.

related assumption that empty cells of the variance-covariance matrix can be filled using evidence from *external* data sources, potentially covering entirely different populations. As noted by Atkinson and Jenkins (1984), the latter assumption hardly makes much sense without first making a general steady-state assumption.

For example, Conlisk (1969) imports external evidence on correlation coefficients in order to overcome the unavailability of parent-child data. Similarly, Bowles and Nelson (1974) draw on multiple different data sets in order to estimate an intergenerational model of IQ, schooling, and income. However, prior to Atkinson and Jenkins (1984), the critical role of these assumptions, and the consequences when the distribution is not in fact in steady state, had not received much attention. Steady-state assumptions also played an important role in related fields, such as contemporary macroeconomics and the literature on genetic heritability. For example, Gimelfarb (1981) notes that the assumption that assortative correlations remain constant across generations has been controversial in population genetics.

While the expression in equation (4) still depends on two moments involving y_{t-1}, steady-state assumptions can ensure identification even if that variable is not observed. In particular, it might be reasonable to assume that $r_{y_t e_t} = r_{y_{t-1} e_{t-1}} = r^*_{y,e}$ and $\sigma_{y_t} = \sigma_{y_{t-1}} = \sigma^*_y$, i.e. that the intragenerational income-education correlation and the standard deviation of income are both constant across generations. Under these assumptions, the *steady-state estimator* is equal to the sample analog of

$$\gamma_{ss} = \frac{r_{y_t e_{t-1}} - r^*_{ye} r_{e_{t-1} e_t}}{r^*_{ye}(1 - r^2_{e_{t-1} e_t})}. \tag{5}$$

The corresponding steady-state estimators of ρ and λ are shown in Appendix A1.[6]

To illustrate how this steady-state estimator may perform in practice, we make use of administrative data on log lifetime incomes and years of education for Swedish parents and children (see data description in Appendix A2). Applying the steady-state estimator above provides the estimate $\hat{\gamma}_{ss} = 0.15$.[7] However, using actual data on parental income (y_{t-1}), so that we do not need to rely on the steady-state assumptions, provides the substantially lower estimate $\hat{\gamma} = 0.07$. While the steady-state assumption can thus appear very useful at first sight – it allows us to identify the conditional relationship between child and parental income despite the latter being unobserved – our example illustrates that this workaround can give quite inaccurate results in practice.

One might also ask more generally how the error that arises from the steady-state assumption depends on the different moments. For our particular example, we can therefore consider the derivative of γ_{ss} in equation (4) with respect to the unobserved moments $r_{y_{t-1} e_{t-1}}$ and $\sigma_{y_{t-1}}$,

[6] Under these specific steady-state assumptions, we can thus replace $r_{y_{t-1} e_{t-1}}$ with $r_{y_t e_t}$ so that γ_t can be estimated using data only on y_t, e_t, e_{t-1}. Such information is often available in survey data sets of generation t, which asks retrospectively about e_{t-1}.

[7] These and the other calculations in the paper are provided only as illustrations of the pitfalls of relying on stead-state assumptions. We therefore do not provide a full set of statistics for the results, including leaving out standard errors.

respectively. To simplify interpretation we normalize the derivatives as elasticities:

$$\left.\frac{\partial\gamma_{ss}/\gamma_{ss}}{\partial r_{y_{t-1}e_{t-1}}/r_{y_{t-1}e_{t-1}}}\right|_{r_{y_{t-1}e_{t-1}}=r^*_{ye}} = -\frac{(r_{y_te_{t-1}} - r^*_{ye}r_{e_{t-1}e_t})r^*_{ye}}{r^2_{y_{t-1}e_{t-1}}(1-r^2_{e_{t-1}e_t})\gamma_{ss}} \tag{6}$$

$$\left.\frac{\partial\gamma_{ss}/\gamma_{ss}}{\partial\sigma_{y_{t-1}}/\sigma_{y_{t-1}}}\right|_{\sigma_{y_{t-1}}=\sigma^*_y} = -\frac{r_{y_te_{t-1}} - r^*_{ye}r_{e_{t-1}e_t}}{r^*_{ye}(1-r^2_{e_{t-1}e_t})\gamma_{ss}}\left(\frac{\sigma^*_y}{\sigma_{y_{t-1}}}\right)^2, \tag{7}$$

where $r_{y_te_t} = r^*_{y,e}$ and $\sigma_{y_t} = \sigma^*_y$. Plugging in the estimates from our empirical example above we find that a percentage deviation from steady state in the income-education correlation would approximately change the estimate of γ_{ss} by -0.21 percent. A corresponding percentage change in the standard deviation of y_{t-1} would result in a change of about -0.57 percent. Moreover, we know that various measures of income dispersion can change quite substantially across generations (e.g. Piketty and Saez 2003). Deviations between the assumed steady-state value and the actual variance of income can therefore lead to potentially large biases.

Atkinson and Jenkins (1984) make a couple of more general points regarding these types of steady-state estimators. First, even if the researcher has the ability to observe all the necessary moments to apply a steady-state estimator like the one above – either if the model is overidentifed (as in our example) or by importing such evidence from some other data source – it is not possible to say if $\gamma_{ss} > \gamma$ or $\gamma_{ss} < \gamma$.[8] Even if it is possible to posit that a moment is transitioning towards its steady state from a specified direction, e.g. $r_{y_t,e_t} > r_{y_{t-1},e_{t-1}}$, one can have either $\gamma_{ss} > \gamma$ or $\gamma_{ss} < \gamma$, depending on other parameter values. Secondly, they note that the relationship between the "steady-state bias", e.g. $\gamma_{ss} - \gamma$, and the quantitative inaccuracy of the steady-state assumption, e.g. $r_{y_t,e_t} - r_{y_{t-1},e_{t-1}}$, is not necessarily monotonic. In fact, the bias $\gamma_{ss} - \gamma$ can change sign multiple times along $r_{y_t,e_t} - r_{y_{t-1},e_{t-1}}$. It is therefore not possible to draw general conclusions about the direction of the potential bias introduced by the steady-state assumption. Atkinson and Jenkins (1984) therefore argue that " *... the steady-state assumption should be used with considerable caution, and only after careful consideration of the structure of the particular model in question*".

2.3 Implications and Extensions

Atkinson and Jenkins (1984) made their arguments at a time in which researchers were limited to small survey data sets, frequently forcing them to combine multiple data sources to estimate a parameter of interest. This need is reduced in many contemporary data sources, which may include variables such as education and income for two or more generations.[9] Moreover, interest in the simultaneous-equation approach considered by Atkinson and Jenkins has waned over the last decades. Many factors are thought to contribute to the transmission of socio-economic

[8] Atkinson and Jenkins (1984) also discuss the case in which the components of the steady-state estimator are not all available to the researcher either, such that e.g. γ_{ss} cannot be point identified. Instead a range of γ_{ss} is identified, varying with the value of the unobserved moment.

[9] However, the quality of those observations is still frequently poor. In such settings, external moments may be used to correct for measurement error, instead of outright replacing a moment that is unobserved in the main data. See e.g. Haider and Solon (2006) for an example of how external moments can be imported.

status from parents and children, and these factors tend to be correlated – parents with high income tend to also have higher education, wealth, cognitive and non-cognitive skills, and so on. Omitted-variable bias is therefore a major concern, making it difficult to place an unequivocal causal interpretation on any given simultaneous-equation model. For the identification of specific causal links, the literature has shifted towards more targeted research designs.[10]

For these reasons, the illustrations in Atkinson and Jenkins (1984) may appear less relevant and be less cited today. But abstracting from the simultaneous equation context, the article makes a more general point, namely that transitional dynamics are central in a literature in which steady-state assumptions relate to two or more *generations*. This general idea remains under-appreciated, as the steady-state assumption is directly relevant for many other strands of the intergenerational literature. And just as in the simultaneous equation context, its importance is frequently ignored. In Section 3, we illustrate its role in the rapidly expanding literature on mobility across multiple generations. In Section 4, we illustrate that transitional dynamics do not only impede the identification of causal parameters, but also affect the interpretation of the large descriptive literature on intergenerational mobility differences across time, space, and groups.

In these two cases, the role of steady-state assumptions has as not yet received much attention. In other cases, they feature under a different label. An example is the two-sample instrumental variables (TSIV) estimator, which is frequently relied on for countries with less rich income data. The relation between income and education or other parental characteristics are estimated in a first ("auxiliary") sample. These estimates are then used to impute parental income in a second ("main") sample, in which this variable is not directly observed (e.g Björklund and Jäntti, 1997). A necessary assumption for consistency of the TSIV-estimator is "random sampling from *the same population*" (Angrist and Krueger, 1995). This is reminiscent of the second type of steady-state assumption discussed by Atkinson and Jenkins (1984), i.e. the assumption that empirical moments from other data sets can be used to replace unobserved moments in the main data set.

3 Multigenerational and Latent Transmission

The previous section considered the case of incomplete data, in which a variable was not (or not well) observed for one particular generation. Now imagine a more severe type of data limitation, in which a key variable is inherently unobservable for *any* generation. For example, the variable e_t in equations (1a) and (1b) could represent an individual's latent ability in a broad sense. Such latent factors played a central role in early sociological research, as well as standard economic models (Goldberger 1972). For instance, Becker and Tomes (1979) assume that a person's socio-economic success depends on his or her *endowments,* which may represent a great variety of unobserved cultural and genetic attributes. Latent characteristics received less

[10] For example, Holmlund et al. (2011) review how the causal effect of parental schooling on child outcomes can be estimated by considering twins, adoptees, or external reforms in the educational system.

attention in empirical work, which instead focused on dependence in observable characteristics (e.g. the intergenerational elasticity of *income*) or specific causal channels (e.g. the causal effect of parent on child education).[11]

That key variables might be inherently unobservable has been instrumental for the rise of the literature on *multigenerational* mobility, which has grown rapidly in recent years. In this literature, researchers consider other types of relatives than just parents and children, with data spanning three or more generations.[12] Much of this work focuses on descriptive questions, in particular on the observation that socio-economic inequalities are more persistent across generations than a Markov interpretation of the conventional parent-child associations would suggest. For example, differences in the average socioeconomic status of *surnames* have been shown to correlate across centuries (e.g. Clark 2014, Barone and Mocetti 2016). The transmission of latent advantages from parents to children could rationalize this finding.[13]

Moreover, as already noted by Becker and Tomes (1979) and Goldberger (1989), the observation of data across more than two generations may help us to identify the parameters of a transmission model – including its inherently unobserved components. As an illustration, consider again the transmission system

$$y_t = \gamma_t y_{t-1} + \rho_t e_t + u_t \tag{8a}$$

$$e_t = \lambda_t e_{t-1} + v_t. \tag{8b}$$

but now assume that the variable y_t represents *education* (or some other observed measure of human capital), while the variable e_t represents an individual's *latent endowment* that is inherently unobserved.[14] To simplify notation, standardize e_t and y_t to have variance one for all t.[15] If we assume that the system is in steady state with time-invariant parameters, we can write the coefficients from a regression of child on parent (β_{-1}) or grandparent (β_{-2}) outcomes as

$$\beta_{-1} = \gamma + \rho\lambda Cov(y, e) \tag{9a}$$

$$\beta_{-2} = \gamma\beta_{-1} + \rho\lambda^2 Cov(y, e) \tag{9b}$$

allowing us to express λ as a function of γ, $\frac{\beta_{-2}-\gamma\beta_{-1}}{\beta_{-1}-\gamma} = \lambda$.

[11] An exception is the literature on sibling correlations. The family background captured by sibling correlations accounts for all factors shared by siblings, including latent factors that are orthogonal to the observed socioeconomic status of parents (see Jäntti and Jenkins, 2014).

[12] Hällsten (2014) provides a comprehensive summary of the early literature. Recent examples include Clark (2014), Clark and Cummins (2014), Lindahl et al. (2015), Braun and Stuhler (2018), Adermon et al. (2018), Solon (2018), Olivetti et al. (2018), Neidhöfer and Stockhausen (forthcoming), Adermon et al. (2016), Modalsli (2016), Barone and Mocetti (2016) or Clark (2017).

[13] See Clark (2014) and Stuhler (2012). A related literature on the role of more distant relatives, such as grandparents, provides an alternative rationalization of the same finding (see for example Mare 2014, Pfeffer 2014, Zeng and Xie 2014, Ferrie et al. 2016, Knigge 2016, Anderson et al. 2018, Breen 2018).

[14] We need to resort to education as our socio-economic outcome in order to facilitate our empirical illustrations below. In particular, we do not have access to income data of sufficient quality for analyses of three generations.

[15] This standardization is without loss of generality. If the parameters of the model are time variant, standardization rescales them accordingly. In particular, λ_t represents then the correlation between e_t and e_{t-1}.

Table 1: Vertical, horizontal and within-person correlations in education

	Vertical			Horizontal	
	1971 cohort	1971 cohort	1951 cohort	1970-75 cohorts	1970-75 cohorts
Panel A: Vertical and horizontal correlations					
Education	Father	Grandfather	Father	Sibling	Cousin
	0.365	0.177	0.374	0.474	0.223
Panel B: Within-person correlations					
Cognitive skills	0.560		0.490	0.563	0.560
Noncognitive skills	0.356		0.247	0.351	0.353
N	12,268		18,358	13,869	19,526

Note: Panel A reports the Pearson correlation coefficient in years of schooling for vertical kins (father-son and paternal grandfather-grandson) and horizontal kins (male siblings and cousins). Panel B reports the within-person correlation between years of schooling and cognitive or noncognitive test scores from the Swedish military enlistment.

The recent literature is motivated by an even simpler model (e.g. Clark, 2014; Clark and Cummins, 2014; Clark, 2017), in which socio-economic outcomes are assumed to have no direct causal effect ($\gamma = 0$) and are instead indirectly transmitted via the latent factor ($\lambda > 0$ and $\rho > 0$). In this *latent factor model*, λ is immediately identified by the ratio (see Braun and Stuhler, 2018; Adermon et al., 2018; and Neidhöfer and Stockhausen, forthcoming)[16]

$$\frac{\beta_{-2}}{\beta_{-1}} = \lambda. \tag{10}$$

We provide an illustration here based on the same Swedish data also used in the previous section, considering the 1971 birth cohort (see description in Appendix A2). As reported in columns (1) and (2) of Table 1, $\hat{\beta}_{-1} \approx 0.365$ for father-son pairs and $\hat{\beta}_{-2} \approx 0.177$ in grandfather-grandson pairs, implying an estimate of $\hat{\lambda} = 0.177/0.365 = 0.485$. Estimates of λ from other birth cohorts are similar.

However, these results rely again critically on steady-state assumptions.[17] In the pure latent factor model ($\gamma = 0$) with time-variant parameters, the ratio of parent-child and grandparent-child correlations identifies instead

$$\frac{\beta_{-2}}{\beta_{-1}} = \lambda_{t-1} \frac{\rho_{t-2}}{\rho_{t-1}} \tag{11}$$

where $\rho_t = Corr(y_t, e_t)$ is the correlation between the socio-economic outcome y_t and the latent endowment e_t (i.e. the square root of the share of outcome variance explained by the latent variable) in generation t. Intuitively, the strength of grandparent-child and parent-child correlations may differ because endowments were imperfectly transmitted from grandparent to the parent generation ($\lambda_{t-1} < 1$) *or* because the degree to which these endowments determine socio-

[16] With the observation of additional generations, more detailed models can be identified. For example, the parameters γ and therefore λ in (1a) and (1b) are identified once a fourth generation becomes available (derivation upon request).

[17] Goldberger (1989) notes that a steady-state assumption is required, but his criticism revolves around other identifying assumptions in Becker and Tomes' model.

economic outcomes may have changed over time.[18] The same argument applies when comparing other moments that relate to different generations, such as parent-child and sibling correlations.

To illustrate this problem, Table 1 reports the father-son correlation in education, and the within-person correlation between education and cognitive and noncognitive test scores for different generations. Columns (1) and (3) compare the 1971 cohort to a cohort born twenty years earlier (i.e. to their potential parent generation).[19] The father-son correlation in educational attainment remains remarkably stable, with 0.374 for the earlier vs. 0.365 for the later cohort (Panel A). However, the correlation between education and cognitive or noncognitive skills (from military enlistment tests) does not remain stable (Panel B). The correlation increases by more than 10 percent for cognitive skills and by more than 40 percent for noncognitive skills between the two cohorts, suggesting that education might have been a substantially better proxy for other advantages in the later cohort. The ratio ρ_{t-2}/ρ_{t-1} could therefore be quite different from one, leading to substantial biases in estimates of λ based on the ratio $\frac{\beta_{-2}}{\beta_{-1}} = \lambda$. This bias could in principle be corrected for by using estimates of ρ_{t-1} and ρ_{t-2}, as in equation (11), if one has access to a good proxy for latent endowments for multiple generations.[20]

Alternatively, the problem can potentially be addressed if additional moments become available; for example, Braun and Stuhler (2018) show that λ_{t-1} is identified if an additional autocorrelation across four generations is observed. But deviations from steady-state will be more difficult to address in more complicated models, such as those also allowing for direct transmission mechanisms (e.g. $\gamma \neq 0$). Steady-state assumptions may therefore play an even more critical role in the multigenerational literature than for the cases studied by Atkinson and Jenkins (1984). First, the problem is not restricted to some data set or settings in which a required moment happens to not be observed, but applies to any attempt at identifying latent factor models from multigenerational data. Second, we may have to impose steady-state assumptions across more than two generations, making such assumptions even less plausible. We also face a trade-off in that the observation of additional generations may help identification of more detailed models, but only if we are willing to impose increasingly strong steady-state assumptions.

3.1 Identification using Horizontal Kinships

To reduce our reliance on the steady-state assumption, a promising approach may be to use *horizontal* (e.g. siblings and cousins) instead of *vertical* (e.g. parents and grandparents) kinship moments for identification (Collado et al., 2018). To illustrate the potential, assume that the errors in equations (8a) and (8b) are not correlated between siblings and consider a pure latent

[18] A partial test can of course be conducted by comparing the variances of y_t across generations, i.e. to check whether $Var(y_t) \approx Var(y_{t-1}) \approx Var(y_{t-2})$, or to compare the parent-child correlations across generations. But while these tests may indicate deviations from steady state, they would not say all that much about whether $r(y_t, e_t)$ remains constant over generations, which is the crucial requirement for identification of λ_{t-1}.

[19] We consider the 1951 cohort because this is the first cohort for which cognitive and noncognitive test scores are observed.

[20] For example, if we assume that ρ_{t-1} is 40 percent larger than ρ_{t-2} (corresponding roughly to the increase in the correlation between education and noncognitive skills reported in Table 1) then a bias-corrected estimate of λ would be 1.4 x 0.485 = 0.679. If we instead use the correlation between education and cognitive skills for the bias-correction, the estimate would be lower at about 1.10 x 0.485 = 0.535.

factor model in which $\gamma = 0$.[21] In that model, the sibling correlation equals $\beta_{sibling} = \rho_t^2 \lambda_t^2$, while the cousin correlation equals $\rho_t^2 \lambda_t^2 \lambda_{t-1}^2$, so that their ratio identifies λ,

$$\left(\frac{\beta_{cousin}}{\beta_{sibling}} \right)^{\frac{1}{2}} = \lambda_{t-1}. \tag{12}$$

In contrast to identification from vertical moments as in equation (11), identification from horizontal moments is not sensitive to time-variation in the correlation between the socio-economic outcome y_t and the latent endowment e_t. Intuitively, if socio-economic status is measured in the same generation and around the same time, we do not have to worry that its relation with the latent factor may change over time.

Columns (4) and (5) in Table 1 provide an example. Considering birth cohorts 1970-75, the correlation in years of schooling between brothers is 0.474, while the correlation between (male) cousins is 0.223. An estimate of λ based on equation (12) is therefore $\hat{\lambda} = (0.223/0.474)^{1/2} = 0.686$. Panel B reports the within-person correlation between years of schooling and the cognitive and noncognitive test scores. The correlations between education and cognitive or noncognitive test scores are nearly the same in the brother and cousin sample. So a key advantage of the horizontal compared to the vertical approach is therefore that outcomes can be measured for the same birth cohorts, reducing our reliance on steady-state assumptions.

While this horizontal approach depends less on steady-state assumptions, they still affect the interpretation of the results. For example, expression (12) illustrates that in our simple model, the cousin-to-sibling ratio identifies the parent-child transmission in the latent factor in generation $t-1$, i.e. in the generation that links the cousins. More generally, kinship correlations are a function of both current and past transmission mechanisms, and can therefore be subject to different time trends. These issues become clearer if we consider not only the steady-state implications, but also the off-steady-state dynamics of an intergenerational model. The next section provides further examples. The recent literature has employed other workarounds to identify λ in the latent-factor model. For example, instead of using multiple generations of the same socio-economic outcome variable, Vosters and Nybom (2017) use multiple proxy variables for the unobserved endowment within a single generation. In this case, the identification does not explicitly depend on steady-state assumptions, but rather on a set of measurement error assumptions. Clark (2014), Clark and Cummins (2014), and Clark (2017) estimate λ from data aggregated at the surname level, with identification relying on an exclusion restriction on the relation between surnames and socio-economic outcomes. The literature has therefore proposed very distinct identification strategies for intergenerational models with latent transmission mechanisms. An interesting task for future research is therefore to compare the proposed methods and their associated identifying assumptions, and to characterize their respective strengths and

[21] It is in fact plausible that the errors are correlated, i.e. that siblings share common influences over and above the shared parental influence. Collado et al. (2018) consider a model that allows for intergenerational and shared sibling influences in both observable and latent variables, and estimate it from both horizontal and vertical moments.

weaknesses. The role of steady-state assumptions, their plausibility, and the degree to which they can be relaxed will be an important part of that debate.

4 The Transitional Dynamics of Intergenerational Mobility

The issues considered by Atkinson and Jenkins (1984) also affect the interpretation of descriptive evidence on intergenerational mobility, although in ways that may not be immediately obvious from the discussion in Section 2. While Atkinson and Jenkins show that failure of the steady-state assumption impedes identification of *structural* parameters, one may also ask how shifts in the structural parameters affect *descriptive* measures of intergenerational mobility and inequality (see Nybom and Stuhler, 2014). This question is in many ways the mirror-image of the structural perspective in Atkinson and Jenkins (1984) or Conlisk (1969; 1974a).

To illustrate the basic idea, we here focus on the intergenerational elasticity of income (IGE), the most popular descriptive measure in the economics literature.[22] With $y_{i,t}$ denoting the (log) lifetime income of the offspring in generation t of family i and $y_{i,t-1}$ the (log) lifetime income of the parent, the IGE is defined as the slope coefficient in the linear regression

$$y_{i,t} = \alpha_t + \beta_t y_{i,t-1} + \epsilon_{i,t}. \tag{13}$$

The IGE β_t captures to what degree percentage differences in parental income on average transmit to the next generation, with a low IGE indicating high mobility. It can be expressed as a function of the transmission system discussed in the previous sections by plugging equations (1a) and (1b) into the linear slope coefficient β_t, such that

$$\beta_t = \frac{Cov(y_t, y_{t-1})}{Var(y_{t-1})} = \gamma_t + \rho_t \lambda_t \frac{Cov(e_{t-1}, y_{t-1})}{Var(y_{t-1})}. \tag{14}$$

The intergenerational elasticity β_t thus depends on current transmission mechanisms (the parameters γ_t, ρ_t and λ_t), and also the covariance between income and education in the parent generation. Intuitively, if income and educational advantages are concentrated in the *same* families, mobility will be low (the IGE will be high). The covariance $Cov(e_{t-1}, y_{t-1})$ is in turn determined by past transmission mechanisms, and thus past values of $\{\gamma_t, \rho_t, \lambda_t\}$.[23] The IGE and other mobility measures are therefore a function of both *current* and *past* transmission mechanisms. Following a permanent structural change (such as an increase in the returns to education), the IGE will transition towards a new steady-state mobility level. As is evident

[22] In recent years it has become increasingly popular to estimate the (Spearman) intergenerational correlation in income ranks (e.g. Chetty et al., 2014a), with the main motivation often being related to measurement concerns.

[23] One can iterate equation (14) backwards to express β_t solely in terms of parameter values (see Nybom and Stuhler, 2014),

$$\beta_t = \gamma_t + \rho_t \lambda_t \rho_{t-1} + \rho_t \lambda_t \left(\sum_{r=1}^{\infty} \left(\prod_{s=1}^{r} \gamma_{t-s} \lambda_{t-s} \right) \rho_{t-r-1} \right),$$

where it is assumed that the process is infinite.

from equation (14), the covariance $Cov(e_{t-1}, y_{t-1})$ is a key determinant of the transition path.[24] Nybom and Stuhler (2014) show that the transition can be long-lasting, with the IGE shifting across multiple generations in response to events that occurred in a particular time period.

This observation has a number of implications. The IGE may change in generation t even if the transmission system has not changed – intergenerational mobility trends may reflect events in the parent generation, instead of events in the offspring generation. Conversely, the IGE may remain stable in generation t even if the transmission system has changed – if trends in the parent generation counteract changes in the transmission system in the current generation. The implications of equation (14) are most obvious with respect to the interpretation of mobility trends over time, but it also has implications for other types of questions. For example, two populations that are currently subject to the same transmission mechanisms can still differ in their levels of mobility, as current mobility also depends on the joint distribution of income and endowments in the parent generation.

The exact shape of the transition path and the magnitudes of the dynamic shifts in the IGE can be analyzed under specific model assumptions. But it is instructive to first demonstrate the existence of such dynamics in actual data. Figure 1 provides an illustration based on our data from Sweden. Subfigure 1a plots the intergenerational elasticity of income for fathers and sons across birth cohorts in Sweden.[25] The elasticity is comparatively stable, increasing slightly between the 1955 birth cohort and those born in the 1960s, before decreasing again in the early 1970s. Similarly, stability of the IGE has also been observed for recent cohorts in the U.S. A common interpretation of such patterns is that the degree of intergenerational transmission has remained stable.

However, the fact that the IGE has remained stable does not necessarily imply that the transmission system has remained stable. Subfigure 1b plots $\frac{Cov(e_{t-1}, y_{t-1})}{Var(y_{t-1})}$, i.e. the slope coefficient from a regression of father's years of schooling on father's log income. The coefficient is declining substantially over the analysis period, from around 3 in cohorts born in the late 1950s to less than 2 in the 1975 cohort. As shown, this drop is due to a substantial drop in the corresponding correlation coefficient – father's education and income became decreasingly correlated over the analysis period. From equation (14), this decrease would decrease the intergenerational elasticity of income, ceteris paribus.

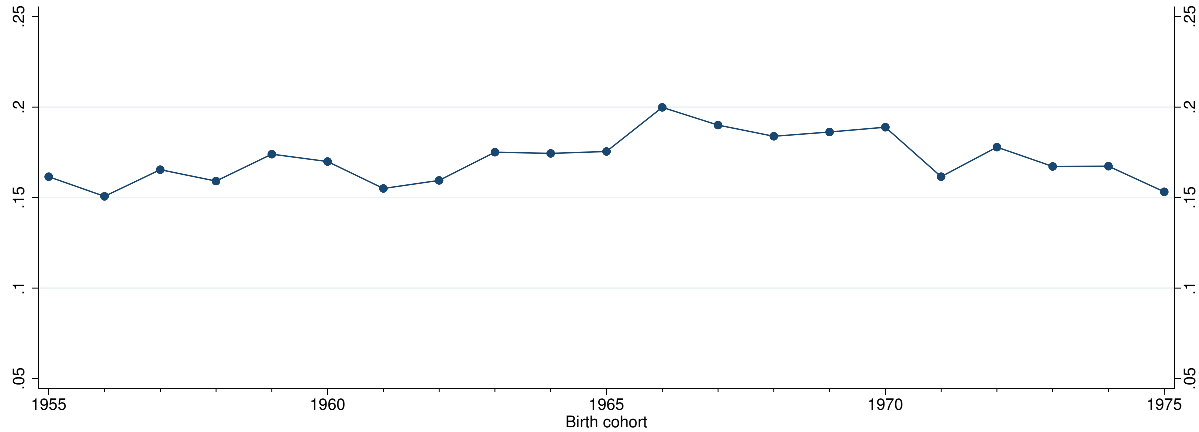
The fact that the intergenerational elasticity has remained stable *despite* the decreasing concentration of income and education in the parent generation therefore suggests that the intergenerational transmission system has not remained stable in generation t – some of the coefficients γ_t, ρ_t or λ_t must instead have increased over time. These observations suggest that intergenerational advantages became more strongly transmitted between the 1955 to 1975 birth

[24] Possible co-movements of the variance $Var(y_{t-1})$ obviously also matter. In fact, if the intergenerational shifts in $Var(y_{t-1})$ and $Cov(e_{t-1}, y_{t-1})$ are exactly proportional, the IGE will be stable.

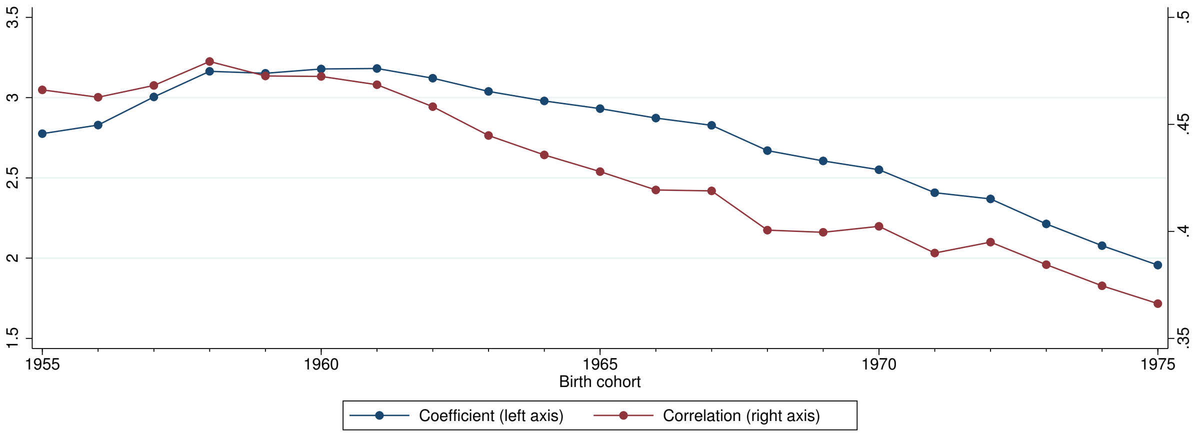
[25] To keep the age at measurement constant across birth cohorts, incomes are measured as five-year averages at age 30-34 for sons and at age 45-49 for fathers. The estimated elasticity is downward biased because these income measures are an imperfect proxy of lifetime income. See Nybom and Stuhler (2017) for a discussion of attenuation and life-cycle bias in the Swedish context.

Figure 1: Intergenerational trends (1955-1975 cohorts)

(a) The intergenerational elasticity of income



(b) The correlation between father's education and income



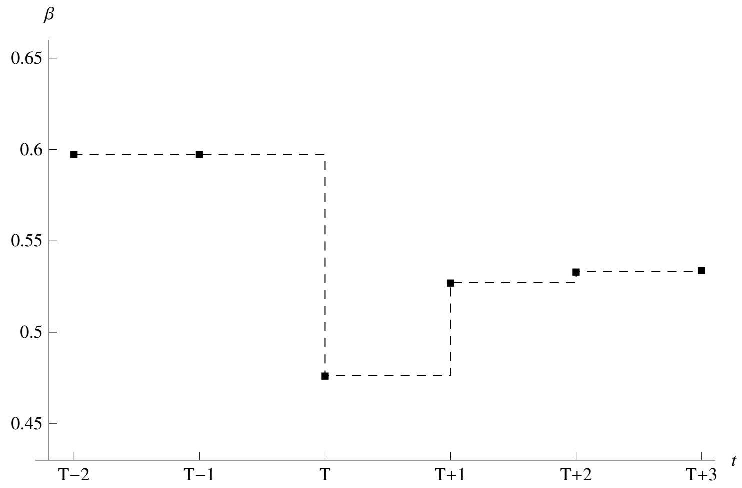
Note: Sub-figure (a) plots the intergenerational elasticity of income for fathers and sons, separately for each birth cohort. Sub-figure (b) plots the coefficient in a regression of years of schooling on log income for the fathers of the corresponding birth cohorts of sons. Source: Swedish register data.

cohorts in Sweden, but that the IGE does not reflect this shift because it was counterbalanced by the decreasing concentration of those advantages in the parent generation.

Nybom and Stuhler (2014) explore such off-steady state dynamics in a simulation, given assumptions about initial conditions and parameter values. A central result is that the transition paths will in some cases be non-monotonic. Figure 2 shows the response to a particular type of structural change, in which the IGE follows a non-monotonic path towards its new steady state. In this case we assume that the direct effect of parental income diminishes ($\gamma_1 > \gamma_2$), while education or skills are instead more strongly rewarded ($\rho_1 < \rho_2$); or, in other words, in generation T the economy becomes more *meritocratic*.[26] Since endowments are imperfectly transmitted within families ($\lambda < 1$) and explain only part of the variation in parental income, the diminishing importance of parental income dominates in generation T and mobility increases.

[26] See Nybom and Stuhler (2014) for a set of similar numerical exercises and a motivation for the choice of parameter values.

Figure 2: A declining impact of parental income and increasing returns to education



Note: Numerical examples of trends in the intergenerational elasticity: in generation T the impact of parental income γ declines from $\gamma_1 = 0.4$ to $\gamma_2 = 0.2$ while the returns to endowments ρ increase from $\rho_1 = 0.5$ to $\rho_2 = 0.7$ (assuming $\lambda = 0.6$).

But after this first affected generation, $Cov(e_{t-1}, y_{t-1})$ responds – in this case it increases – and intergenerational mobility tends to decrease again.[27] Intuitively, a change towards a more meritocratic society increases the long-run correlation between endowments and income, thereby *decreasing* mobility after the first affected generation. Figure 2 illustrates that this transition can be fairly long-lasting; mobility shifts become visually insignificant only in the third generation, or, in terms of calendar years, more than half a century after the structural change.

Times of change may therefore be times of high mobility, and negative mobility trends that are observed empirically may in fact stem from *gains* in mobility in the past. Nybom and Stuhler (2014) also note that the transition paths of mobility and cross-sectional measures of inequality interact in interesting ways, and that allowing for dynamics in the dispersion of income and endowments complicates the dynamics in the IGE further.[28] This last point relates directly to the original idea in Atkinson and Jenkins (1984): imposing assumptions of constant distributional moments, or steady states, simplifies the analysis while also potentially leading to important differences in actual results.

The presence of off-steady-state dynamics affects therefore the *interpretation* of empirical estimates. They are of particular importance for the recent literature studying time trends in intergenerational mobility, including how such trends differ between groups and how they relate

[27] Mobility shifts in the first affected generation by $\Delta\beta_T = (\gamma_2 - \gamma_1) + (\rho_2 - \rho_1)\lambda Cov(e_{T-1}, y_{T-1})$, while the second-generation shift is given by $\Delta\beta_{T+1} = \rho_2\lambda\Delta Cov(e_T, y_T)$. Additional dynamics also tend to arise from changes in the variance of income, which we assumed away here.

[28] The analysis can be further generalized by allowing for multiple dimensions of human capital (or endowments), thus replacing the e_t above with a vector whose elements can have different returns. It can be shown that changes in relative returns to endowments then tend to lead to relative gains and losses that generate *transitional* mobility, which is similar to the intuition for the case in Figure 2.

to policy changes.[29] That the off-steady-state dynamics can be important is understood in so far as previous research demonstrates that intergenerational correlations are much more stable over time than regression coefficients (Hertz et al., 2008).[30]

However, the discussion above hopefully illustrates that the issue of non-stationarity goes beyond simple scalings of the variance or other adjustments for distributional changes across generations. Systems of intergenerational transmission are inherently dynamic and observed time trends can be either a consequence of recent policy changes or a repercussion from changes in the more distant past (or a combination of both). For example, a central concern in much of the recent research on mobility trends, as well as in the public debate, is that mobility has declined in conjunction with the recent rise in income inequality. A host of potential explanations, including educational expansion, rising returns to education, or immigration, have been proposed in both research and in the public debate. Common to all is that they tend to relate trends to *recent* events that directly affect current cohorts, which makes sense if the results are interpreted from a steady-state perspective. Allowing for non-stationarity – or an off-steady-state perspective – would instead suggest that the explanations might instead be found in the more distant past.

For this reason, we suggest that researchers consider both contemporary and more distant events when attempting to explain time trends in intergenerational mobility. In particular, the concentration of income and other characteristics in the *parent* generation is a useful statistic for the interpretation of trends in the child generation. For example, we may consider the correlation between income and education among parents, as plotted in Subfigure 1b. This analysis can be extended if the data contain more information about parents. For example, Markussen and Røed (2017) show that the correlation between economic success and cognitive ability remained stable among Norwegian parents during their analysis period. This observation suggests that the parent-child association in Norway changed because of recent events, which directly affected the child generation, not because of distant events that affected the child generation via their parents. The implications also extend to analyses of causal effects of policy and institutional reforms. As Figure 2 illustrates, a reform that affects a specific transmission mechanism may have quite different consequences for aggregate intergenerational mobility in the short and in the long run. For example, evidence from Finland indicates a large positive effect of the introduction of comprehensive schooling on intergenerational income mobility (Pekkarinen et al., 2009). It is possible that this mobility gain is partly temporary and that the intergenerational elasticity will again increase in subsequent generations.[31] Other research has estimated a substantial mobility-

[29] Hertz (2007), Lee and Solon (2009) and Chetty et al. (2014a) find no evidence of a major trend in the second half of the 20th century in the US, but cannot reject more gradual changes over time. In contrast, Levine and Mazumder (2007) as well as Aaronson and Mazumder (2008) argue that mobility has fallen in recent decades. A decline has also been found for the UK (Blanden et al., 2004; Nicoletti and Ermisch, 2007), while mobility increased in the Nordic countries (Bratberg et al., 2007; Pekkala and Lucas, 2007; Björklund et al., 2009; Pekkarinen et al. forthcoming).

[30] Chetty et al. (2014b) extends this idea further, arguing that measures of rank correlation are even more stable, as they are invariant to any changes in marginal distributions over generations.

[31] In addition, the large effect in the first affected generation may be partially due to changes in the variance of incomes. As the variance stabilizes in subsequent generations, the IGE is likely to exhibit a more mechanical rebound to pre-reform levels (see also Nybom and Stuhler, 2014).

enhancing effect of subsidized child care (Havnes and Mogstad, 2015). A similar story is possible here – subsequent dynamics may well lead to a rebound of the intergenerational income elasticity closer towards pre-reform levels. In order to say more about the likely long-run effects of reforms like these, explicit consideration of off-steady-state dynamics can be fruitful.

5 Concluding Remarks

Tony Atkinson made several distinct contributions to research on the intergenerational transmission of economic status, including work on its measurement and welfare foundations. While many of these contributions had a profound impact, not all of his work had a direct influence on the subsequent literature. One example is his work, together with Stephen Jenkins, on the role of steady-state assumptions in the context of intergenerational mobility research (Atkinson and Jenkins, 1984). The objective of this paper was to review their arguments and to show that they have great relevance for intergenerational research today.

Departing from Atkinson and Jenkins (1984), we discussed their main insights for the identification of parameters in a structural intergenerational equation system. The situation they considered was one of *incomplete data*, in which a variable was unobserved in a certain data set (but potentially observable). It is then convenient to assume that the unobserved variable's distributional moments are in steady state, such that the missing data moments (e.g. variance of parental income) can be retrieved from other sources (e.g. variance of child income). Atkinson and Jenkins noted that this assumption may be particularly implausible in intergenerational data spanning across *generations*, and illustrated the potential consequences of its violation.

We argue that their core arguments extend to more recent work on multigenerational data across more than two generations, which can be used to identify structural relationships between both observable and unobservable variables. Multigenerational models have, for example, recently been used to identify the intergenerational persistence of latent status or ability. By extending Atkinson and Jenkins' arguments, however, one may ask if this line of work is not even more susceptible to the same critique, as data span not only two but three or sometimes even four generations. Assuming stationary data moments or constant transmission mechanisms in this setting appears implausible, and we found that the deviations from steady-state can potentially be large. A potential solution, as we illustrate, is to instead consider horizontal kinships, such as associations between siblings and cousins, in order to identify the exact same transmission parameters. This is a potentially useful way forward, although other caveats apply.

Steady-state assumptions are often implicitly invoked when interpreting descriptive measures of intergenerational mobility, such as mobility trends over time. The existing literature relates observed trends in mobility (or the lack of such trends) to recent changes in the economic environment. But if the transmission system is not in steady state, mobility trends may also reflect past structural changes in the transmission system. We demonstrate the potential role of these off-steady-state dynamics in Swedish registry data. While the intergenerational elasticity of income remains stable for males born between 1955 and 1975, the concentration of income and education among their parents has in fact increased over this period. A stable intergenerational measure may thus hide important changes in intergenerational transmission that become more apparent when explicitly studying deviations from the steady-state assumption.

Steady-state assumptions are unattractive in a literature in which a single "period" represents an entire generation. Still, steady-state assumptions have remained popular in practice, as they greatly simplify the estimation of intergenerational models and the interpretation of descriptive findings. As we showed here, it is possible to reduce the dependence on such assumptions. And a more explicit consideration of transitional dynamics may turn out to be key for understanding recent descriptive findings, such as on the variation in mobility across countries and over time.

References

AABERGE, R., F. BOURGUIGNON, A. BRANDOLINI, F. H. G. FERREIRA, J. C. GORNICK, J. HILLS, M. JÄNTTI, S. P. JENKINS, E. MARLIER, J. MICKLEWRIGHT, AND B. NOLAN (2017): "Tony Atkinson and his Legacy," *Review of Income and Wealth*, 63(3), 411–444.

AARONSON, D., AND B. MAZUMDER (2008): "Intergenerational Economic Mobility in the United States, 1940 to 2000," *Journal of Human Resources*, 43(1).

ADERMON, A., M. LINDAHL, AND M. PALME (2016): "Dynastic Human Capital, Inequality and Intergenerational Mobility," Discussion paper, mimeo, Uppsala University.

ADERMON, A., M. LINDAHL, AND D. WALDENSTRÖM (2018): "Intergenerational Wealth Mobility and the Role of Inheritance: Evidence from Multiple Generations," *The Economic Journal*, 128(612), F482–F513.

ANDERSON, L. R., P. SHEPPARD, AND C. W. S. MONDEN (2018): "Grandparent Effects on Educational Outcomes: A Systematic Review," *Sociological Science*.

ATKINSON, A., AND F. BOURGUIGNON (1982): "The Comparison of Multi-Dimensioned Distributions of Economic Status," *Review of Economic Studies*, 49(2), 183–201.

ATKINSON, A. B. (1969): "The Timescale of Economic Models: How Long is the Long Run?," *Review of Economic Studies*, 36(2), 137–152.

——— (1970): "On the Measurement of Inequality," *Journal of Economic Theory*, 2(3), 244–263.

——— (1980): "On Intergenerational Income Mobility in Britain," *Journal of Post Keynesian Economics*, 3(2), 194–218.

——— (1981): "The Measurement of Economic Mobility," in *Essays in Honor of Jan Pen*, ed. by P. J. Eigjelshoven, and L. J. van Gemerden. Het Spectrum, Utrecht.

——— (1983): "The Measurement of Economic Mobility," in *Social Justice and Public Policy*, ed. by A. B. Atkinson, chap. 3, pp. 61–76. MIT Press, Cambridge, MA.

ATKINSON, A. B., AND S. JENKINS (1984): "The Steady-State Assumption and the Estimation of Distributional and Related Models," *Journal of Human Resources*, 19(3), 358–376.

BARONE, G., AND S. MOCETTI (2016): "Intergenerational Mobility in the Very Long Run: Florence 1427-2011," Temi di discussione (Economic working papers) 1060, Bank of Italy, Economic Research and International Relations Area.

BECKER, G., AND N. TOMES (1979): "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *The Journal of Political Economy*, 87(6), 1153–1189.

Björklund, A., and M. Jäntti (1997): "Intergenerational Income Mobility in Sweden Compared to the United States," *American Economic Review*, 87(5), 1009–18.

Björklund, A., M. Jäntti, and M. J. Lindquist (2009): "Family Background and Income During the Rise of the Welfare State: Brother Correlations in Income for Swedish Men born 1932-1968," *Journal of Public Economics*, 93(5-6), 671 – 680.

Black, S. E., and P. Devereux (2011): "Recent Developments in Intergenerational Mobility," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 4A. Elsevier.

Blanden, J., A. Goodman, P. Gregg, and S. Machin (2004): "Changes in Intergenerational Mobility in Britain," in *Generational Income Mobility in North America and Europe*, pp. 122–46. Cambridge University Press.

Bowles, S., and V. I. Nelson (1974): "The "Inheritance of IQ" and the Intergenerational Reproduction of Economic Inequality," *The Review of Economics and Statistics*, 56(1), pp. 39–51.

Bratberg, E., Ø. A. Nilsen, and K. Vaage (2007): "Trends in Intergenerational Mobility across Offspring's Earnings Distribution in Norway," *Industrial Relations: A Journal of Economy and Society*, 46(1), 112–129.

Braun, S. T., and J. Stuhler (2018): "The Transmission of Inequality Across Multiple Generations: Testing Recent Theories with Evidence from Germany," *The Economic Journal*, 128(609), 576–611.

Breen, R. (2018): "Some Methodological Problems in the Study of Multigenerational Mobility," *European Sociological Review*, p. jcy037.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014a): "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States," *Quarterly Journal of Economics*, 129(4), 1553–1623.

Chetty, R., N. Hendren, P. Kline, E. Saez, and N. Turner (2014b): "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility," *American Economics Review: Papers and Proceedings*, 104(5), 141–47.

Clark, G. (2014): *The Son Also Rises: Surnames and the History of Social Mobility*. Princeton University Press.

——— (2017): "Nature Versus Nurture in Social Outcomes. A Lineage Study of 66,000 English Individuals, 1750-2016," Discussion paper, mimeo.

Clark, G., and N. Cummins (2012): "What is the True Rate of Social Mobility? Surnames and Social Mobility, England 1800-2012," Unpublished working paper.

——— (2014): "Intergenerational Wealth Mobility in England, 1858–2012: Surnames and Social Mobility," *The Economic Journal*, 125(582), 61–85.

Collado, M. D., I. Ortuño-Ortin, and J. Stuhler (2018): "Kinship Correlations and Intergenerational Mobility," Discussion paper, Working Paper.

Conlisk, J. (1969): "An Approach to the Theory of Inequality in the Size Distribution of Income," *Economic Inquiry*, 7(2), 180–186.

——— (1974a): "Can Equalization of Opportunity Reduce Social Mobility?," *The American Economic Review*, 64(1), pp. 80–90.

——— (1974b): "Stability in a Random Coefficient Model," *International Economic Review*, 15(2), pp. 529–533.

Davies, J. B., J. Zhang, and J. Zeng (2005): "Intergenerational Mobility under Private vs. Public Education," *Scandinavian Journal of Economics*, 107(3), 399–417.

Ferrie, J., C. Massey, and J. Rothbaum (2016): "Do Grandparents and Great-Grandparents Matter? Multigenerational Mobility in the US, 1910-2013," Working Paper 22635, National Bureau of Economic Research.

Gimelfarb, A. (1981): "A General Linear Model for the Genotypic Covariance Between Relatives Under Assortative Mating," *Journal of Mathematical Biology*.

Goldberger, A. S. (1972): "Structural Equation Methods in the Social Sciences," *Econometrica*, 40(6), 979–1001.

——— (1979): "Heritability," *Economica*, 46(184), pp. 327–347.

——— (1989): "Economic and Mechanical Models of Intergenerational Transmission," *The American Economic Review*, 79(3), pp. 504–513.

Haider, S., and G. Solon (2006): "Life-Cycle Variation in the Association between Current and Lifetime Earnings," *American Economic Review*, 96(4), 1308–1320.

Hällsten, M. (2014): "Inequality across three and four generations in Egalitarian Sweden: 1st and 2nd cousin correlations in socio-economic outcomes," *Research in Social Stratification and Mobility*, 35(0), 19 – 33, Inequality Across Multiple Generations.

Havnes, T., and M. Mogstad (2015): "Is Universal Child Care Leveling the Playing Field?," *Journal of Public Economics*, 127, 100 – 114, The Nordic Model.

Hertz, T. (2007): "Trends in the Intergenerational Elasticity of Family Income in the United States," *Industrial Relations*, 46(1), 22–50.

Hertz, T., T. Jayasundera, P. Piraino, S. Selcuk, N. Smith, and A. Verashchagina (2008): "The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends," *The B.E. Journal of Economic Analysis & Policy*, 7(2), 1–48.

Holmlund, H., M. Lindahl, and E. Plug (2011): "The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods," *Journal of Economic Literature*, 49(3), 615–51.

Jäntti, M., and S. P. Jenkins (2014): "Income Mobility," in *Handbook of Income Distribution*, ed. by A. B. Atkinson, and F. Bourguignon, vol. 2. Elsevier.

Jenkins, S. (1987): "Snapshots versus Movies: 'Lifecycle biases' and the Estimation of Intergenerational Earnings Inheritance," *European Economic Review*, 31(5), 1149–1158.

Jenkins, S. P. (1982): "Tools for the Analysis of Distributional Models," *The Manchester School of Economic & Social Studies*, 50(2), 139–50.

Joreskog, K. G. (1970): "A General Method for Analysis of Covariance Structures," *Biometrika*, 57(2), 239–251.

Knigge, A. (2016): "Beyond the Parental Generation: The Influence of Grandfathers and Great-grandfathers on Status Attainment," *Demography*, pp. 1–26.

Lee, C.-I., and G. Solon (2009): "Trends in Intergenerational Income Mobility," *The Review of Economics and Statistics*, 91(4), 766–772.

Levine, D. I., and B. Mazumder (2007): "The Growing Importance of Family: Evidence from Brothers' Earnings," *Industrial Relations: A Journal of Economy and Society*, 46(1), 7–21.

Lindahl, M., M. Palme, S. Sandgren Massih, and A. Sjögren (2015): "Long-term Intergenerational Persistence of Human Capital: An Empirical Analysis of Four Generations," *Journal of Human Resources*, 50(1), 1–33.

Loury, G. C. (1981): "Intergenerational Transfers and the Distribution of Earnings," *Econometrica*, 49(4), 843–67.

Mare, R. D. (2014): "Multigenerational aspects of social stratification: Issues for further research," *Research in Social Stratification and Mobility*, (35), 121–128.

Markussen, S., and K. Røed (2017): "Egalitarianism under Pressure: Toward Lower Economic Mobility in the Knowledge Economy?," IZA Discussion Papers 10664, Institute for the Study of Labor (IZA).

Modalsli, J. H. (2016): "Multigenerational persistence: Evidence from 146 years of administrative data," Discussion paper, mimeo.

Neidhöfer, G., and M. Stockhausen (forthcoming): "Dynastic Inequality Compared: Multigenerational Mobility in the United States, the United Kingdom, and Germany," *Review of Income and Wealth*, 0(0).

Nicoletti, C., and J. Ermisch (2007): "Intergenerational Earnings Mobility: Changes across Cohorts in Britain," *The B.E. Journal of Economic Analysis & Policy*, 7(2), Article 9.

Nybom, M., and J. Stuhler (2014): "Interpreting Trends in Intergenerational Mobility," Working Paper Series 3/2014, Swedish Institute for Social Research.

——— (2017): "Biases in Standard Measures of Intergenerational Income Dependence," *Journal of Human Resources*, 52(3), 800–825.

Olivetti, C., M. D. Paserman, and L. Salisbury (2018): "Three-generation mobility in the United States, 1850–1940: The role of maternal and paternal grandparents," *Explorations in Economic History*.

Pekkala, S., and R. E. B. Lucas (2007): "Differences across Cohorts in Finnish Intergenerational Income Mobility," *Industrial Relations: A Journal of Economy and Society*, 46(1), 81–111.

Pekkarinen, T., K. G. Salvanes, and M. Sarvimäki (forthcoming): "The Evolution of Social Mobility. Norway over the 20th Century," *Scandinavian Journal of Economics*.

Pekkarinen, T., R. Uusitalo, and S. Kerr (2009): "School Tracking and Intergenerational Income Mobility: Evidence from the Finnish Comprehensive School Reform," *Journal of Public Economics*, 93(7-8), 965–973.

Pfeffer, F. T. (2014): "Multigenerational Approaches to Social Mobility. A Multifaceted Research Agenda," *Research in Social Stratification and Mobility*, 35, 1–12.

Piketty, T., and E. Saez (2003): "Income Inequality In The United States, 1913-1998," *The Quarterly Journal of Economics*, 118(1), 1–39.

Solon, G. (1992): "Intergenerational Income Mobility in the United States," *American Economic Review*, 82(3), 393–408.

——— (2004): "A Model of Intergenerational Mobility Variation over Time and Place," *Generational Income Mobility in North America and Europe*, pp. 38–47.

——— (2018): "What Do We Know So Far about Multigenerational Mobility?," *The Economic Journal*, 128(612), F340–F352.

Stuhler, J. (2012): "Mobility Across Multiple Generations: The Iterated Regression Fallacy," IZA Discussion Papers 7072, Institute for the Study of Labor (IZA).

VOSTERS, K., AND M. NYBOM (2017): "Intergenerational Persistence in Latent Socioeconomic Status: Evidence from Sweden and the United States," *Journal of Labor Economics*, 35(3), 869 – 901.

ZENG, Z., AND Y. XIE (2014): "The Effects of Grandparents on Children's Schooling: Evidence From Rural China," *Demography*, 51(2), 599–617.

ZIMMERMAN, D. J. (1992): "Regression toward Mediocrity in Economic Stature," *American Economic Review*, 82(3), 409–29.

Appendix

A1 Derivations for Section 2.1

We can solve for the model's parameters using its implied covariances and variances listed in equations (2a)-(2g). One should note that, as in Atkinson and Jenkins (1984), this simultaneous-equations model is overidentified – we have five unknowns and seven equations – and we thus have multiple solutions to the system of equations. We proceed by combining equations (2e) and (2f) which gives:

$$\gamma = \frac{\sigma_{y_{t-1}y_t} - \rho\lambda\sigma_{y_{t-1}e_{t-1}}}{\sigma_{y_{t-1}y_{t-1}}} = \frac{\sigma_{y_{t-1}y_t} - \rho\sigma_{y_{t-1}e_t}}{\sigma_{y_{t-1}y_{t-1}}}. \quad \text{(A1)}$$

Using equations (2a) and(2c) gives:

$$\rho = \frac{\sigma_{e_t y_t} - \gamma\lambda\sigma_{y_{t-1}e_{t-1}}}{\sigma_{e_t e_t}} = \frac{\sigma_{e_t y_t} - \gamma\sigma_{y_{t-1}e_t}}{\sigma_{e_t e_t}}, \quad \text{(A2)}$$

where we again used equation (2f) in the second step. We can then use these last two expressions (equations (A1) and (A2)) to write γ and ρ in terms of observable population moments:

$$\gamma = \frac{\sigma_{y_{t-1}y_t}\sigma_{e_t e_t} - \sigma_{y_t e_t}\sigma_{y_{t-1}e_t}}{\sigma_{y_{t-1}y_{t-1}}\sigma_{e_t e_t} - \sigma^2_{y_{t-1}e_t}} \quad \text{(A3)}$$

$$\rho = \frac{\sigma_{y_{t-1}y_{t-1}}\sigma_{e_t y_t} - \sigma_{y_{t-1}e_t}\sigma_{y_{t-1}y_t}}{\sigma_{y_{t-1}y_{t-1}}\sigma_{e_t e_t} - \sigma^2_{e_t y_{t-1}}}. \quad \text{(A4)}$$

So far these expressions are identical to those in Atkinson and Jenkins (1984), despite our slightly different and simpler model. In addition, we get our corresponding expression for λ directly from equation (2f). One can then express these in terms of correlations rather than covariances:

$$\lambda = \frac{\sigma_{y_{t-1}e_t}}{\sigma_{e_{t-1}y_{t-1}}} = \frac{r_{y_{t-1}e_t}}{r_{e_{t-1}y_{t-1}}}\left(\frac{\sigma_{e_t}}{\sigma_{e_{t-1}}}\right) \quad \text{(A5)}$$

$$\rho = \frac{\sigma_{y_{t-1}y_{t-1}}\sigma_{e_t y_t} - \sigma_{y_{t-1}e_t}\sigma_{y_{t-1}y_t}}{\sigma_{y_{t-1}y_{t-1}}\sigma_{e_t e_t} - \sigma^2_{e_t y_{t-1}}} = \frac{r_{y_t e_t} - r_{y_t y_{t-1}}r_{y_{t-1}e_t}}{1 - r^2_{y_{t-1}e_t}}\left(\frac{\sigma_{y_t}}{\sigma_{e_t}}\right) \quad \text{(A6)}$$

$$\gamma = \frac{\sigma_{y_{t-1}y_t}\sigma_{e_t e_t} - \sigma_{y_t e_t}\sigma_{y_{t-1}e_t}}{\sigma_{y_{t-1}y_{t-1}}\sigma_{e_t e_t} - \sigma^2_{e_t y_{t-1}}} = \frac{r_{y_t y_{t-1}} - r_{y_t e_t}r_{y_{t-1}e_t}}{1 - r^2_{y_{t-1}e_t}}\left(\frac{\sigma_{y_t}}{\sigma_{y_{t-1}}}\right), \quad \text{(A7)}$$

where we used the notation r_{XY} for the correlation between two variables X and Y, and σ_X for the standard deviation of X. Our focus is on the case in which parental income y_{t-1} is unobserved. We restrict our attention to γ and ρ again and note that these depend on three moments involving y_{t-1}. But since this model (as the one in Atkinson and Jenkins (1984)) is overidentified, one can use the previously unused correlation in equation (2g) together with equations (A6) and (A7) to substitute out $r_{y_t y_{t-1}}$:

$$\rho = \frac{r_{y_t e_t} r_{e_{t-1} y_{t-1}} - r_{y_t e_{t-1}} r_{e_t y_{t-1}}}{r_{e_{t-1} y_{t-1}} - r_{y_{t-1} e_t} r_{e_{t-1} e_t}} \left(\frac{\sigma_{yt}}{\sigma_{et}}\right) \tag{A8}$$

$$\gamma = \frac{r_{y_t e_{t-1}} - r_{y_t e_t} r_{e_{t-1} e_t}}{r_{e_{t-1} y_{t-1}} - r_{y_{t-1} e_t} r_{e_{t-1} e_t}} \left(\frac{\sigma_{yt}}{\sigma_{yt-1}}\right) \tag{A9}$$

However, we have yet another unused equation in the case of our simplified model. We can therefore also substitute out $r_{y_{t-1} e_t}$, using $\lambda = r_{y_{t-1} e_t} = r_{e_{t-1} e_t} r_{e_{t-1} y_{t-1}}$, which we get from equations (2d) and (2f). We now have:

$$\rho = \frac{r_{y_t e_t} - r_{y_t e_{t-1}} r_{e_{t-1} e_t}}{1 - r^2_{e_{t-1} e_t}} \left(\frac{\sigma_{yt}}{\sigma_{et}}\right) \tag{A10}$$

$$\gamma = \frac{r_{y_t e_{t-1}} - r_{y_t e_t} r_{e_{t-1} e_t}}{r_{e_{t-1} y_{t-1}} (1 - r^2_{e_{t-1} e_t})} \left(\frac{\sigma_{yt}}{\sigma_{yt-1}}\right), \tag{A11}$$

where ρ does not depend on the unobserved y_{t-1} at all, while γ depends on two moments involving y_{t-1}. However, the counterpart of those moments can be observed for generation t and it is therefore possible to adopt a couple of steady-state assumptions to ensure identification. In particular, we might be willing to assume that $r_{y_t, e_t} = r_{y_{t-1}, e_{t-1}} = r^*_{y,e}$ and $\sigma_{yt} = \sigma_{yt-1} = \sigma^*_y$, i.e. that the intragenerational income-education correlation and the standard deviation of income are both constant across generations. Under these assumptions, we can take as the steady-state estimator:

$$\gamma_{ss} = \frac{r_{y_t e_{t-1}} - r^*_{ye} r_{e_{t-1} e_t}}{r^*_{ye} (1 - r^2_{e_{t-1} e_t})}. \tag{A12}$$

Thus, under these steady-state assumptions we are capable of identifying both γ and ρ. Note that the identification of λ follows directly from equation (2d), such that

$$\lambda = \frac{\sigma_{e_t e_{t-1}}}{\sigma_{e_{t-1} e_{t-1}}} = r_{e_{t-1} e_t} \left(\frac{\sigma_{e_t}}{\sigma_{e_{t-1}}}\right) \tag{A13}$$

and, with the additional (but in this case unnecessary) steady-state assumption that $\sigma_{e_t} = \sigma_{e_{t-1}}$, this reduces to $\lambda = r_{e_{t-1} e_t}$.

A2 Data Description

Our empirical illustrations are based on a 35 percent random sample of the Swedish population born between 1932 and 1967. Using information based on population registers, we can link these sampled individuals to their biological parents and children. We then individually match data on demographic characteristics based on bi-decennial censuses starting from 1960, as well as education and income data stemming from official registers. For males born 1951-1980 we also match data on cognitive and noncognitive test scores from the mandatory military enlistment held by the Swedish War Archives.

Educational registers were compiled in 1970, 1990 and about every third year thereafter, containing detailed information on each individual's highest educational attainment. We consider the highest schooling level recorded across these years, and translate it into years of education, with 7 years for the old compulsory school being the minimum, and 20 years for a doctoral degree the maximum. Education data in 1970 is available only for those born 1911 and later. As the data are collected from official registers there are no standard non-response problems.

The income data stem from official tax declaration files and are held by Statistics Sweden. We construct measures of long-run income based on age-specific averages of annual incomes, which are observed for the years 1968-2007. We use total (pre-tax) income, which is the sum of an individual's labor (and labor-related) earnings, early-age pensions, and net income from business and capital realizations. We adjust all incomes for inflation using the CPI. Incomes for parents are necessarily measured at a later age than incomes for their offspring, which may bias estimates. Specifically, we construct five-year averages of annual incomes measured between age 45 and 49 for fathers, and between age 30 and 34 for sons. To limit the influence of outliers, we winsorize these averages at the 1st and 99th percentile of their respective birth cohort.

To these data we add military enlistment test scores. Complete information from the draft is available for males who were drafted between 1969 and 2000. During these years, almost all males went through the draft procedure at age 18 or 19, and enlistment scores are available for 88-95 percent of each birth cohort. The data include an overall measure of cognitive skill and a corresponding measure of overall non-cognitive skill. The overall cognitive score is based on four sub-tests measuring: inductive skill (or reasoning); verbal comprehension; spatial ability; and technical understanding. Overall cognitive skill is reported on an integer Stanine scale, which varies from one to nine.[32] The evaluation of non-cognitive ability is based on a procedure that was adopted in 1969 and consists of a 25-minute interview with a certified psychologist. As a basis for the interview, the psychologist had information on the results and the cognitive tests, the results on various physical tests, school grades, and answers from a questionnaire about families, friends and hobbies. The interview as such was centered around a number of pre-specified behavioral topics. Based on the interview, the draftee gets an overall score on an integer Stanine scale. The overall score reflects social maturity, psychological energy (e.g., focus and perseverance), intensity (e.g., activation without external pressure), and emotional stability (e.g., tolerance to stress).

For the analysis in Section 2, we consider (log) incomes and years of education for males born 1968-1970 and their fathers. For the analyses in Section 3 we consider males born between 1951 (the earliest cohort for which cognitive and noncognitive test scores are available) and 1975, as well as the fathers and paternal grandfathers of the 1971 cohort. For the analyses in Section 4 we consider males born between 1955 and 1975, as well as their fathers. To ensure that incomes are observed at the same age range, we here consider father-son pairs in which the father was between 25 and 30 years old at the birth of the child. We verified that this restriction has only

[32]The Stanines are (approximately) normally distributed with a mean of 5 and a standard deviation of 2.

a small effect on our estimates in those cohorts in which the age restrictions were not necessary.