

Some Methodological Problems in the Study of Multigenerational Mobility¹

Richard Breen

Department of Sociology and Nuffield College, Oxford

Abstract

A number of recent studies by sociologists have sought to discover whether a person's status (typically their social class, education or SES) is directly affected by the status of their grandparents, once the effects of parents' status are controlled. The results have been ambiguous, with some studies finding a direct effect of grandparents on their grandchildren, while others find no effect. I use causal graphical methods to demonstrate some of the methodological problems that occur in trying to identify this direct effect and I offer some suggestions as to how they might be addressed.

Introduction

A growing number of studies in sociology and economics seek to determine whether characteristics of grandparents affect grandchildren's outcomes, net of, or controlling for, the influence of parental factors. The outcome in question is usually a measure of occupation, such as class or socio-economic status, of income or earnings, or of education. There is some scepticism among economists about whether these studies have shed much light on the question (Solon 2017) and sociological analyses of transmission over three generations have reached conflicting conclusions, with some claiming that grandparents' characteristics have no direct effect on their grandchildren; others claiming that they do; and yet others arguing that there is an effect, but only under particular circumstances or for certain groups (see the comprehensive review by Anderson, Sheppard and Monden, 2017).

In this note I argue that there are some methodological problems that have gone largely unrecognised by sociologists working in this area and which make the identification of a direct effect of grandparents on grandchildren very challenging. If

¹ I thank the ESR editor and reviewers for their comments. Thanks to John Ermisch, Ian Lundberg, Kenneth MacDonald and participants at workshops in Nuffield College in 2017 and 2018 for their helpful suggestions.

we cared only about establishing an association between grandparents and their grandchildren, net of parents, the critiques I shall make would not apply. But studies in this area write of the “effect” of grandparents on grandchildren, suggesting a desire to do more than document associations. Nevertheless, the arguments I advance here, while they mostly serve to show the problems involved in making causal claims, can also be taken as suggestions about how an association may arise and thus might help in interpreting it. The three problems I discuss are: collider bias in estimates of the direct effect of grandparents on grandchildren; conditioning on children in the parental generation; and conditioning on 'mechanisms' that are invoked to explain the circumstances under which grandparents directly affect their grandchildren.

Graphical Models

I use a graphical approach.² The graphical approach has advantages of clarity, because graphs are usually more transparent than equations, and generality, in that they do not make any parametric assumptions. If a causal effect is identified in a graph it is non-parametrically identified, and its identification does not depend on any assumptions about the functional form of the relationship between cause and effect. In the directed acyclic graphs (DAGs) used in this paper, nodes (vertices) represent variables and arrows (directed edges) indicate hypothesised causal effects. DAGs can include both observed and unobserved variables: the latter are shown in circles. A variable in a box has been conditioned on.

Three-generation models

Almost all studies of grandparental effects estimate the relationship between an outcome for grandchildren and one or more measures of parents and grandparents. The simplest design of this kind is found in, for example, Erola and Moisio (2007), Chan and Boliver (2013) and Hertel and Groh-Samberg (2014), all of whom examine the relationships between grandparental, parental and child social class. Other studies, such as Møllegaard and Jaeger (2015) and Bol and Kalmijn (2015), consider multiple dimensions of grandparental and parental status (economic, social and cultural capital in the former case) and their relationship to one or more grandchild outcomes. In one of the earliest studies of this kind, Warren and Hauser (1997) used

² Accessible introductions to graphical models are Elwert 2013; Pearl, Glymour and Jewell 2016.

multiple measures of parents and grandparents to predict grandchildren's education and occupational status. The majority of studies also include other characteristics of the parents that are not included for the grandparents: Ziefle (2016), for example, examined the relationship between the education of grandparents, parents and children but also included controls for parental social class, occupational status, family income and so on.

The goal of these studies is to determine whether there is a direct effect of grandparental measures on grandchildren's outcomes, once parental measures are controlled for. The DAG in Figure 1 shows the simplest case. Here the direct effect of grandparents' status, X , on grandchild's outcome, Y , is the causal effect that does not go through parents' status, Z . As noted earlier, usually X and Z measure the same characteristic in the two generations. The direct effect is the estimand of interest. For ease of exposition it is labelled b in Figure 1, though, because we are concerned with non-parametric identification, it is not necessarily a single number (such as a linear regression coefficient). The indirect effect is the product of the causal paths from grandparent to parent and parent to child (a times c). The total effect is the unconditional effect of X on Y . In a linear model the total effect is the sum of the direct and indirect effects ($b + a$ times c), but this need not hold in a non-linear model and so does not necessarily hold in the DAG.

[Figure 1 here]

The conditioning set defines the direct effect: in Figure 1 the conditioning set is Z and so the direct effect is the effect of X on Y that does not go through Z . The possibility that the effect of X on Y might also be mediated by another parental variable, say $Z2$, explains why other parental measures are sometimes included in the model, as shown in Figure 2. In this case the direct effect is the effect of X on Y that does not pass through either $Z1$ or $Z2$.

[Figure 2 here]

There are two problematic issues in estimating the direct effect. The first is the possibility that there are unmeasured factors, W , that are causes of X and Y : these would confound the direct grandparent – grandchild relationship. The second is the possibility of unmeasured factors, U , causing Z and Y : these would confound the direct parent – child relationship. Both are shown in Figure 3.³

[Figure 3 here]

It is relatively easy to think of candidates for U : if Z and Y represented parental and child educational attainment their relationship would partly depend on genes, on cultural factors such as values and beliefs that shape parental education and are passed to children, material factors such as wealth, and common environmental influences related to where parents and children live. It is less easy to think of candidates for W since these are sources of a grandparent—grandchild correlation independent of parents. Indeed, economists commonly use grandparental status as an instrument for parental status, so assuming that no W s exist and that there is no direct effect of X on Y . In any event, and without prejudging the issue, our attention focuses on the problems caused by unobserved factors U – not least because these seem to have largely gone unnoticed.

Although U straightforwardly confounds the direct parent – child relationship, it also leads to bias in the estimate of the direct grandparent – grandchild effect. This is because, to estimate the direct effect, b , we must condition on Z , but Z is a collider on the path from X to U , as revealed by the arrows entering it from both X and U . A collider that is not controlled, or conditioned on, blocks the path between its sources, but, once a collider is conditioned on, that path is opened (Elwert and Winship 2014). In the graph shown in Figure 3,⁴ conditioning on Z opens a path from X to Z to U to Y .

³ When a DAG includes unobserved variables a single such variable, like U in Figure 3, may stand for a set of unobserved variables with arrows representing causal relationships among them (for example, $Z \leftarrow U1 \rightarrow U2 \rightarrow Y$ where both $U1$ and $U2$ are unmeasured) as long as the direction of the arrows from the unobserved to the observed variables (Z and Y in this case) is preserved and there are no unmeasured colliders blocking the path from Z to Y (as in, for example, $Z \leftarrow U1 \rightarrow U3 \leftarrow U2 \rightarrow Y$, where $U3$ is unmeasured).

⁴ This graph, and the others in this paper, could be made more elaborate through the introduction of more unmeasured variables and more arrows between variables, but for the expositional purposes of this paper making the graph more complicated would not be helpful.

The resulting estimate suffers from “collider bias”.⁵ If there were no unmeasured confounders, W , we could estimate the total effect of X on Y without bias. But to identify the direct effect we would also have to block the biasing path through the collider, Z . We could do this by trying to measure U and then conditioning on it. Of course, since U probably represents numerous unmeasured causes of Z and Y , conditioning would have to be exhaustive to remove all the bias. Partial conditioning might or might not reduce the bias, as we can see if we consider two confounders with opposite signed effects on Y . Conditioning on only one of them could easily increase the absolute bias. Adding other parental measures (as in Figure 2) will not help unless they can be considered to be causes of Z , and may even worsen the bias because these additional variables might also be colliders.

Neighbourhoods provide a concrete example of a confounder that, thus far, has been omitted from studies of the direct effect of grandparents on grandchildren. The outcomes for children and the predictors measured for parents and grandparents, such as education, social class and various forms of capital and resources, are all likely to be affected to some extent by neighbourhoods, and neighbourhood location may be transmitted from one generation to the next (for example, Sharkey and Elwert 2011; the same argument could be applied to regions: see Chetty *et al.* 2014). The situation is shown in Figure 4, where G represents the neighbourhood factors relevant for X , Z and Y . Here the path from $Z \leftarrow Gp \rightarrow Gc \rightarrow Y$ equates to the confounding shown by U in Figure 3; the path from $X \leftarrow Gg \rightarrow Gp \rightarrow Gc \rightarrow Y$ equates to the confounding shown by W in Figure 3. The path from $Y \leftarrow Gg \rightarrow Gp \rightarrow Z$ indicates confounding of the effect of grandparents on parents. Collider bias in the estimate of b arises from conditioning on Z thus opening the path $X \rightarrow Z \leftarrow Gp \rightarrow Gc \rightarrow Y$. If X , Z and Y represent educational attainment, and we assume that the effect of G has the same sign for parents and children, Gp positively affects Gc , and ordinary least squares is used, then the collider bias will lead the estimate of b to be downwardly biased (see the appendix for the

⁵ There is a growing literature on how this problem affects studies dealing with the mediation of causal effects (see, for example, VanderWeele 2016). Imagine X is a randomized treatment (administering of a particular drug) and we would like to know how much of its effect on Y (survival to one year) is not mediated via Z (a subsequent condition partially caused by treatment). Although X is assigned randomly, Z is not, and if the determinants of Z (other than X) are associated with Y , estimates of the direct and indirect effects of X on Y will be biased. The problem seems not to be known by analysts of three-generational transmission with the exception of Song (2016), but in her analysis she explicitly assumes that there are no omitted confounders of the mediators and the outcome (2016: 1915).

proof). This may help explain cases in which no direct effect of grandparents on grandchildren is found. In general however, it will be impossible to know the magnitude of such bias and there may be offsetting biases from different sources as I show in the appendix. But insofar as variables such as those mentioned above have not been included in studies of multigenerational mobility there exists the likelihood that direct effect estimates are biased.

[Figure 4 here]

Conditioning on Children

Economists sometimes extrapolate from the conventional intergenerational income elasticity to a multi-generational elasticity. The parent-child intergenerational income elasticity (IGE) is the estimated β from the following log-linear regression:

$$\log(y_c) = \alpha + \beta \log(y_p) + \varepsilon \quad (1)$$

with y representing income of the child, c , and parents, p . The n -generational elasticity is then β^{n-1} . This was the method used by Becker and Tomes (1986: S1) to underpin their claim that: “Almost all earnings advantages and disadvantages of ancestors are wiped out in three generations”.

This extrapolation rests on two assumptions: that intergenerational transmission follows a first-order Markov process, and that the process has a constant transition rate. We have already seen that, even with data from three generations, it is difficult to determine whether the direct effect of grandparents on grandchildren is absent; thus it is difficult to test the first-order Markov assumption. To test the assumption of a constant transition rate we might consider comparing the effect of X on Z (labelled a in Figure 1) with the effect of Z on Y (c in Figure 1). However, even putting aside the problem of unmeasured confounders of these relationships, a and c represent different effects. This is because the effect of parents on children is estimated for all children, irrespective of whether they have children or not. But, the effect of grandparents on parents is estimated only for those children of grandparents who are

parents themselves. In other words, the estimate of a conditions on generation 2 having children; the estimate of c does not condition on generation 3 having children.

Figure 5 shows the consequences of conditioning on parents in the middle generation. To focus on the issues that arise, the confounders, U , are omitted from Figure 5: we are therefore ignoring the biases already discussed. Figure 5 reflects the belief that Z has a causal effect on whether or not someone has a child, denoted by the variable *Child*, which, for simplicity, distinguishes between those who do ($Child = 1$) and those who do not ($Child = 0$) have children. V represents the other determinants of whether a person becomes a parent. These are assumed to be independent of X though this is not necessary for what follows.

[Figure 5 here]

Child is a collider and, by conditioning on it, a path is opened from X to V (shown by the dotted line in Figure 5) and thence to Z . This path biases the estimate of the direct effect of X on Z relative to what would have obtained had the parental status of the second generation not been conditioned on (and which is not done in estimating the direct effect of Z on Y).⁶ There are circumstances in which conditioning on *Child* can also bias the direct effect of X on Y . Suppose that V affected not only *Child* but also Y . This could occur if V included unmeasured characteristics of a spouse (such as education) that affected the likelihood of having a child and affected children's outcomes, Y (if there were any children). This is shown in Figure 6. Then there is a biasing path from Z to *Child* to V to Y that has been induced by conditioning on the collider, *Child*. Biases in estimates of either or both of the $X - Z$ and $Z - Y$ effects will invalidate comparisons between them.

[Figure 6 here]

Conditioning on Joint Survival

⁶ If we had assumed that V depended on X then the bias would still be present. In this case there would be two arrows linking X and V : a directed arrow from X to V representing the causal effect of X on V and the dashed line representing the relationship induced by conditioning on the collider, *Child*. Both of these paths would introduce bias because neither would link X to Z via V if *Child* had not been conditioned on.

Several studies have sought to show the circumstances under which grandparents will have a direct effect on their grandchildren. It has been suggested that when they live in the same household or in close proximity, grandparents will have a greater impact (Ferguson and Ready 2011, Bol and Kalmijn 2016) and that the effect will be larger when the grandparents' and grandchildren's lives overlap (Braun and Stuhler 2016). Zeng and Xie (2014) consider both these mechanisms. Empirically these hypotheses are usually addressed by adding interaction terms: that is, conditioning on joint survival or co-residence to allow the direct effect of X on Y to differ accordingly. But conditioning on variables can sometimes create more problems that it solves.

Assume that the effect of grandparents on grandchildren is proportional to the number of years they are alive together. Denote this overlap by the variable S . To test our hypothesis we condition on S . But this introduces a quite subtle bias. Grandparents and grandchildren are likely to overlap a greater number of years if the grandparents had their children when they were young, and if their children did the same. We denote age at first childbirth by E . But age at first childbirth depends on social class or SES (the things that typically make up X and Z). And we also know that the age at which a mother had her first child is consequential for the education (and probably other outcomes) of her children (McLanahan 2004). Putting all this together leads to the DAG in Figure 7.

[Figure 7 here]

Figure 7 shows that S is a collider on the path between Eg and Ep . Eg and Ep depend on X and Z , respectively. Eg is consequential for Z and Ep for Y . Because S is a collider, conditioning on it opens the path linking Eg and Ep (shown as a dotted line) and this opens a path from X to Eg to Ep to Y . This is a further source of bias in the estimate of the direct effect of X on Y .

To make things more concrete: imagine X , Z and Y all represent socio-economic status. Lower SES women tend to start having children when they are younger, and so conditioning on joint survivorship will bias the sample towards lower-SES families. If the fertility – SES relationship were to change, estimates of the direct effect would

likely also change, but this would not be because the true effect itself had changed but because we would be estimating the effect on a sample with a different distribution of SES.⁷

Limitations

I have focused on studies that measure characteristics of grandparents, parents and children. I have not addressed an alternative, though less common, approach that exploits the difference in the correlation of an outcome among siblings and among cousins to infer the existence and magnitude of a grandparental (or even great-grandparental) effect (Jaeger 2012; Hällsten 2014; Knigge 2016). These studies are reviewed and criticised by Lundberg (2018).

Even among the multi-generational mobility studies of the kind I have considered, the problems discussed here are not exhaustive. Economists have different criticisms (Solon 2017, Stuhler 2012) and there are well known issues concerning such things as which parent and which grandparents should be included and, if more than one, how the information about them should be combined. I have ignored other problems too, such as measurement error, despite the possibility that they are also present.⁸

A potential problem that has not been discussed here or anywhere else is the following. Analyses of grandparental effects are often motivated by testing claims made in other papers about whether such an effect exists or not. But the data in the different studies may have been collected using different sampling schemes. For example, Chan and Bolliver (2013) motivate their test of grandparental effects by referring to Warren and Hauser's (1997) finding, using US data, that there are none; Chan and Bolliver, using British data, arrive at the opposite conclusion. But, even putting aside the country differences, the data used by Warren and Hauser (the Wisconsin Longitudinal Study) samples individuals born in a specific year and collects

⁷ Conditioning on joint survivorship may introduce other biases too. Within families, the lives of earlier born children are more likely to overlap with the lives of their grandparents and, if birth order matters to Y , this is another source of bias. On the other hand, higher SES people live longer, and so conditioning on overlapping lives may be leading to the selection of higher SES grandparents who nevertheless started childbearing when they were young.

⁸ Two early and apparently little known papers (MacDonald 1974; Ridge 1974) provide cogent discussion of several relevant issues in studying social mobility over three generations.

information about them, their parents, and their children whereas the National Child Development Study, used by Chan and Bolliver, samples individuals born in a specific month and year, and collects information about them, their parents, and their grandparents. So, in this design, compared with the WLS design, there is no variation in birth years of the child generation but greater variation in birth years of parents and much greater variation in the birth years of grandparents.⁹ This may or may not be consequential for the determination of whether there is a direct grandparental effect, but it seems never to have been considered.

Solutions

A first and important step towards dealing with the problems I have discussed is to consider how the relationship between grandparent and grandchild might have come about – including explanations that do not assume the relationship is causal. Not only will this suggest a range of interpretations of the relationship it may also point to ways in which, through conditioning or otherwise, some of these interpretations may be tested. DAGs provide an efficient and intuitive way of doing this.

In some specific cases, solutions are available. The bias that arises in Figure 7, for example, could be dealt with by controlling for age at first birth among grandparents and/or parents. *Eg* and/or *Es* would thus be observed and, since neither is a collider, controlling for one or both would block the path from *X* to *Eg* to *Ep* to *Y*. Figure 7 is a particular case of the situation in which a variable (*S* in this case) is believed to mediate the grandparental effect. It will normally be helpful to include in the model variables that are determinants of both the mediator and the outcome, just as, in many circumstances, we control for variables that are determinants of “treatment” and outcome.

Whether the second problem considered here, namely the different samples used to estimate the grandparent – parent and parent – child relationships, can be addressed depends on the structure of the data. If, as in the Warren and Hauser (1997) study, information is collected on one or more specific birth cohorts, some members of which

⁹ Liu (2018) uses the Framingham Heart Study to obtain data on three generations and here, unlike both WLS and NCDS, the original sampling is on the grandparental generation.

become parents, it may be possible to estimate the relationship between grandparents and their children as well as between grandparents and their children who also became parents. In this way estimates of the relationship would be made more comparable to those involving parents and children.¹⁰ However, if, as in Chan and Bolliver's (2013) study, children are sampled and information is added on their parents and grandparents, it is difficult to see any adequate solution. Because the parents are drawn from a range of birth cohorts it will be difficult to establish who the equivalent non-parents should be and, even if this could be done, we would still have no data about them and their parents.

The first problem discussed here, of collider bias in the estimation of direct effects, involves unobserved variables, and so it is impossible to know the magnitude or direction of the bias (if any) in any specific study. We can say something about the bias under certain assumptions, but these include assumptions about unobserved relationships and so cannot be tested. For example, in a linear model specified according to Figure 3, and assuming all the arrows (including those from unobserved variables) had the same sign, confounding and collider bias would have different signs and so would offset each other. Confounding would impart upward bias while conditioning on the collider, Z , would lead to downward bias (see the appendix for the proof). But this is not necessarily the case in general (for example, in non-linear models).

There are two main approaches to the issue, primarily found in the epidemiological literature. The first is tests of the sensitivity of estimates of a direct effect to different degrees of collider bias: Tingley *et al* (2014) have developed an R package for this purpose. The second is an instrumental variable approach. Consider a linear model based on Figure 3, but ignoring the confounder W . Define \hat{Z} as the predicted value of Z from a regression of Z on X and R , where R is an instrument for Z . Assuming that X has a constant (non-heterogeneous) effect, a regression of Y on X , conditioning on \hat{Z} , will yield consistent and asymptotically unbiased estimates of that direct effect. The IV approach was used by Robins and Greenland (1992) and has recently been developed

¹⁰ This would also be possible with population register data.

by, among others, Frölich and Huber (2014). Given an instrument, R , for Z one could also adopt a control function approach (Wooldridge 2015). In this case, let \hat{U} be the residual from the regression of Z on X and R ; \hat{U} is then included as a control variable in the regression of Y on X and Z that is used to estimate the direct effect.

Conclusion

I have discussed three problems that seem to have gone unrecognised in the burgeoning literature on the analysis of grandparental effects. Indeed, one of these problems – bias in the estimation of a direct effect by conditioning on a mediator variable that may be a collider – seems not to be recognised in the wider community of quantitative sociologists, even those who are aware of how causal estimates can be confounded by unobserved variables. Yet the problem is likely to occur in a wide range of settings where sociologists want to decompose the effect of a variable of interest into direct and indirect effects (for example, in mobility studies that seek to separate the effect of social origins on destination into the indirect effect, mediated by education, and the direct effect, net of education). It is impossible to know the extent of bias in estimates of the direct effect of grandparents on grandchildren in the published literature, but, insofar as these studies have omitted variables that affect both parents and children, such as genes, culture, ethnic group, neighbourhood or region and so forth, it seems highly probable that they suffer from collider bias.

I have suggested some solutions to the problems I identified but perhaps the most valuable approach, in these as in any other analyses, is to consider possible biases and interpret ones results in the light of them.

Appendix: Bias in direct grandparental effects in a linear model of three generation mobility.

In a linear model version of the DAG shown in Figure 3, in which all the arrows have the same sign, confounding and collider bias will have different signs and so will offset each other. To see this consider Figure A1. U is an unmeasured variable that affects the mediator, Z , and the outcome, Y , and W is an unmeasured variable that affects X and Y . We want to know the bias in the estimate of the direct effect if X on Y (that is, conditioning on Z).

[Figure A1 here]

Assuming that all the variables have a standard deviation of 1, then, using the rules of path analysis and the relationship between marginal and conditional regression parameters, we have:

$$\hat{\beta}_{YX} = \beta + \alpha\gamma + rs$$

Here β is the true direct effect of X on Y and $\hat{\beta}_{YX}$ indicates an estimated coefficient from the regression of Y on X .

The conditional regression coefficient is:

$$\begin{aligned}\hat{\beta}_{YX|Z} &= \frac{\beta + \alpha\gamma - (\gamma + pq + \alpha\beta + \alpha rs)\alpha}{1 - \alpha^2} \\ &= \beta + rs - \frac{\alpha pq}{1 - \alpha^2}\end{aligned}$$

The bias in the conditional estimate is thus equal to confounding bias (from W) given by rs and collider bias (due to U) and equal to $-\frac{\alpha pq}{1 - \alpha^2}$. So, in this case, the two forms of bias will offset each other. However, we cannot know the extent of this offsetting, of

course, because we cannot estimate p , q , r or s . Indeed, the only parameter in Figure A1 that is identified is α .

References

Anderson, Lewis, Paula Sheppard and Christiaan Monden. 2018 Grandparent effects on educational outcomes – a systematic review. *Sociological Science* 5: 114-42.

Becker, Gary S. and Nigel Tones. 1986. Human Capital and the Rise and Fall of Families. *Journal of Labor Economics*, 4: S1–S39.

Bol, T. and M. Kalmijn. 2016. Grandparents' resources and grandchildren's schooling: Does grandparental involvement moderate the grandparent effect? *Social Science Research*. 55: 155–70.

Braun, Sebastian. T., and Jan Stuhler. 2016. The Transmission of Inequality Across Multiple Generations: Testing Recent Theories with Evidence from Germany. *The Economic Journal*, 128: 576-611.

Brumback, Babette, Miguel A. Hernán, Sebastien J. P A . Haneuse, and James M. Robins. 2004. Sensitivity analyses for unmeasured confounding assuming a marginal structural model for repeated measures. *Statistics in Medicine* 23: 749-67.

Chan, Tak Wing and Vikki Bolliver. 2013. The Grandparents Effect in Social Mobility: Evidence from British Birth Cohort Studies. *American Sociological Review*. 78: 662-78.

Chetty, Raj, Nathaniel Hendren, Patrick Kline and Emmanuel Saez. 2014. Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States. *The Quarterly Journal of Economics* 129: 1553-1623.

Elwert, Felix. 2013. Graphical Causal Models. Chapter 13 in S.L. Morgan (ed.), *Handbook of Causal Analysis for the Social Sciences*. Springer.

Elwert, Felix and Christopher Winship. 2014. Endogenous Selection Bias: The Problem of Conditioning on a Collider. *Annual Review of Sociology* 40: 31-50.

Ferguson, J.L and D.D. Ready. 2011. Expanding notions of social reproduction: Grandparents' educational attainment and grandchildren's cognitive skills. *Early Childhood Research Quarterly*. 26: 216–26.

Frolich, Markus and Martin Huber. 2014. Direct and Indirect Treatment Effects: Causal Chains and Mediation Analysis with Instrumental Variables. IZA Discussion Paper No. 8280.

Hällsten, Martin. 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: 19-33.

Jaeger, Mads Meier. 2012. The extended family and children's educational success. *American Sociological Review* 77: 903-22.

Knigge, Antonie. 2016. Beyond the parental generations: The influence of grandfathers and great-grandfathers on status attainment. *Demography* 53: 1219-44.

Liu, Hexuan. 2018. Social and Genetic Pathways in Multigenerational Transmission of Educational Attainment. *American Sociological Review*, 83:278-304.

Lundberg, Ian. 2018. Measuring multigenerational mobility: What can we learn from cousin correlations? Princeton: Office of Population Research.

MacDonald, Kenneth. I. 1974. Downwardly Mobile Mothers: And Other Interaction Effects. Pages 73-84 in J.M.Ridge (ed.) *Mobility in Britain Reconsidered*, Oxford Studies in Social Mobility, Oxford University Press.

McLanahan, Sara. 2004. Diverging Destinies: How Children Are Faring Under the Second Demographic Transition. *Demography* 41: 607-27.

Pearl, Judea, Madelyn Glymour and Nicholas P. Jewell. 2016. *Causal Inference in Statistics: A Primer*. Wiley.

Ridge, John M. 1974. Three Generations. Pages 47-72 in J.M.Ridge (ed.) *Mobility in Britain Reconsidered*, Oxford Studies in Social Mobility, Oxford University Press.

Robins, James M. and Sander Greenland. 1992. Identifiability and Exchangeability for Direct and Indirect Effects. *Epidemiology*, 3: 143-55.

Sharkey, Patrick and Felix Elwert. 2011. The Legacy of Disadvantage: Multigenerational Neighborhood Effects on Cognitive Ability. *American Journal of Sociology* 116: 1934-81.

Solon, Gary. 2017. What Do We Know So Far about Multigenerational Mobility? IZA Discussion Paper, 10623.

Song, Xi. 2016. Diverging Mobility Trajectories: Grandparent Effects on Educational Attainment in One- and Two-Parent Families in the United States. *Demography* 53: 1905-32.

Stuhler, Jan. 2012. Mobility Across Multiple Generations: The Iterated Regression Fallacy. IZA Discussion Paper, 7072.

Tingley, Dustin, Teppei Yamamoto, Kentaro Hirose, Luke Keele and Kosuke Imai. 2014. mediation: R Package for Causal Mediation Analysis. *The Journal of Statistical Software* 59: 1- 38.

VanderWeele, Tyler J. 2016. Mediation Analysis: A Practitioner's Guide. *Annual Review of Public Health*, 37: 17-32.

VanderWeele, Tyler J. and Onyebuchi A. Arah. 2011. Bias Formulas for Sensitivity Analysis of Unmeasured Confounding for General Outcomes, Treatments, and Confounders. *Epidemiology*, 22: 42-52.

Warren, John R. and Robert M. Hauser. 1997. Social Stratification across Three Generations: New Evidence from the Wisconsin Longitudinal Survey. *American*

Sociological Review. 62:561–72.

Wooldridge, Jeffrey M. 2015. Control Function Methods in Applied Econometrics. *The Journal of Human Resources* 50: 420 – 45.

Zeng, Z. and Y. Xie. 2014. The Effects of Grandparents on Children's Schooling: Evidence From Rural China. *Demography*. 51: 599–617.

Figure 1: Mobility over 3 Generations

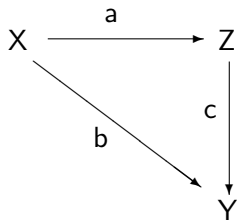


Figure 2: Multiple Parental Measures

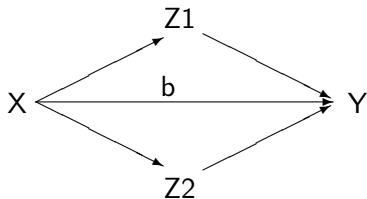


Figure 3: Two Sources of Bias

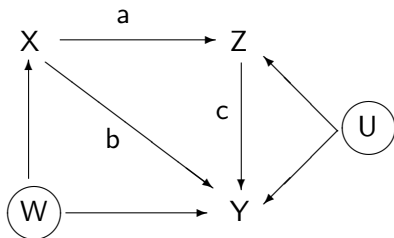


Figure 4: Mobility over 3 Generations with confounders

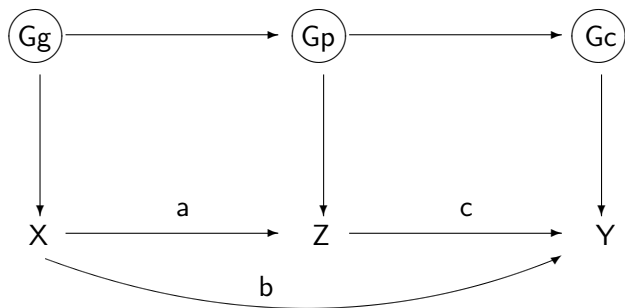


Figure 5: Conditioning on Children

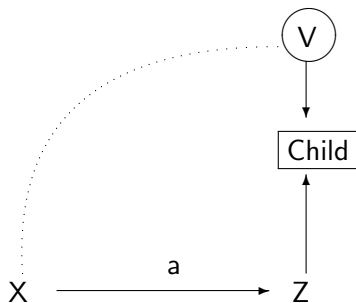


Figure 6: Conditioning on Children, bias in $Z - Y$

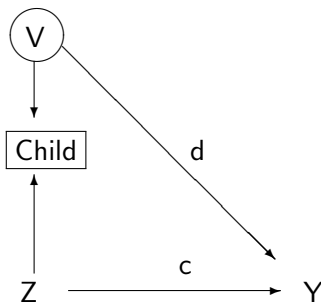


Figure 7: Conditioning on Surviving

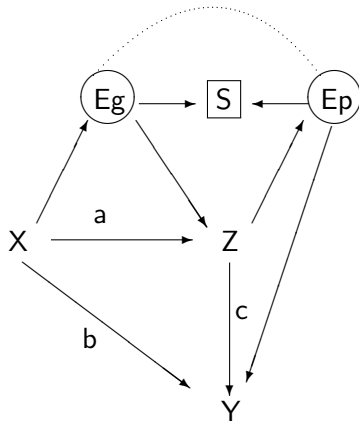


Figure A1: Offsetting Biases

